2000
Solid-State Sensors, Actuators, and Microsystems Workshop

Greeting from the Chair

Committees

Table of Contents

Author Index

Keyword Index

Copyright

www.hh2000.org

2000
Solid-State Sensors, Actuators, and Microsystems Workshop

Hilton Head Island, South Carolina • June 4 - 8, 2000

If you have questions regarding the installation, please contact:

The Printing House
Phone: +1-608-873-4500
Hours: Monday through Friday, 8 am - 5 pm CST
E-mail: graphics@printinghouseinc.com

Greetings from the General Chairman

On behalf of the organizing committee and the technical program committee, it is my pleasure to welcome you to the 2000 Hilton Head Solid State Sensor and Actuator Workshop. This is the ninth in a biannual series of regional workshops that started in 1984, and we hope that it has now established itself as one of the major meetings in the field of microfabricated or micromachined devices and systems.

One of our goals has always been to preserve the workshop atmosphere of this meeting, in spite of the increasing number of attendees. We cannot, of course, recreate the intimacy when the workshop had less than one hundred attendees, and it was possible to interact with everyone; that is an inevitable consequence of the growth of this field. But many of the features that have always characterized the Hilton Head Workshops have been maintained: no parallel sessions, ample time for questions during the oral presentations, and free afternoons to allow direct interaction between participants. In addition to the oral presentations and contributed posters, which are highly competitive and selected from a large number of high-quality abstracts, we have repeated the open poster session, which gives everyone an opportunity to present their work. This new poster session was very successful during the 1998 workshop, and we look forward to the 2000 version.

The strength of this workshop is the dedication and enthusiasm of its participants, and we can all take pride in the progress of our field. Sixteen years ago, we started mostly with solid-state sensors and actuators, which we were hoping would take over the world. Today, there are many success stories, including pressure sensors, accelerometers, and ink-jet print heads. Paradoxically, it is often when a device becomes widely accepted in the marketplace, and thus achieves the real-world impact that we all aspire to, that it is no longer discussed in conferences such as this one. The Hilton Head Workshop has always tried to look ahead, to the products of the future, and we hope that among the contributions in this workshop are some pioneering papers that will initiate new directions or applications. Compared to a decade ago, the level of ambition of many projects in our field has increased greatly. We no longer think in terms of sensor devices, but rather of complete microfabricated systems.

This workshop is built on the volunteer efforts of many individuals, who contribute their time and talent. We would like to thank all the members of the Technical Program Committee for reading and evaluating the several hundred abstracts we received, and Prof. Marty Schmidt, the Technical Program Chairman, for organizing the creation of the program of the workshop. Local arrangements were handled by Prof. Mark Allen. Dave Eddy oversaw the production of the digest and the CD-ROM, and Joe Giachino, as always, was in charge of finances. Funding for student travel was arranged by Prof. Mehran Mehregany, and we thank Bill Tang and Abe Lee from DARPA, and the Transducer Research Foundation, for providing the funding that allows many graduate students to present their work here.

For the first time, the Hilton Head Workshop has made use of a professional organizer to handle all the interactions with participants and authors, and thereby allow Marty Schmidt and myself to concentrate our efforts on the scientific goals of the workshop. We are therefore particularly grateful to Katharine Cline, of Preferred Meeting Management, for her dedication in making this workshop possible.

I hope you will all enjoy your workshop!
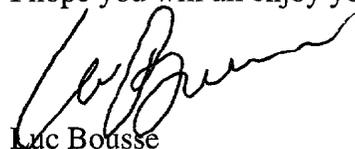
Luc Bousse
General Chairman

2000 Solid-State Sensor & Actuator Workshop

Organizing Committee

General Chairman

Luc J. Bousse
Caliper Technologies Corporation

Technical Program Chairman

Martin A. Schmidt
Massachusetts Institute of Technology

Treasurer

Joseph M. Giachino
Visteon

Publications

David S. Eddy
Delphi Automotive Systems

Local Arrangements

Mark G. Allen
Georgia Institute of Technology

Program Meeting Arrangements

Robert J. Huber
University of Utah

Student Travel Grants

Mehran Mehregany
Case Western Reserve University

Technical Program Committee

Deborah Charych
Chiron Corporation

Philip Congdon

Leslie Field
Agilent Technologies

John R. Gilbert
Microcosm Technologies, Inc.

Jiri Janata
Georgia Institute of Technology

Hal Jerman
Seagate Research

Thomas W. Kenny
Stanford University

Dorian Liepmann
University of California at Berkeley

Marc Madou
Ohio State University

Carlos H. Mastrangelo
University of Michigan

Chip Spangler
Visteon

Solid-State Sensor and Actuator Workshop
June 4 – 8, 2000
Crowne Plaza Resort
Hilton Head Island, South Carolina

Program and Table of Contents

Sunday, June 4

6:00 to 9:00 pm   Registration and Welcome Reception

Monday, June 5

7:15 to 7:45 am   Breakfast

7:45 to 8:15 am   Welcome and Introduction – Luc J. Bousse, Chairman

10:15 to 10:40 am   Break

Session 2 – MEMS Technology and Circuits
Session Chair – Hal Jerman

Tuesday, June 6

Session 4 – Fluid Microactuators
Session Chair – Carlos Mastrangelo

Wednesday, June 7

Session 5 – Microfluidics
Session Chair – Deborah Charych

Contributed Posters (Monday June 5)

Contributed Posters (continued)

Contributed Posters (continued)

Contributed Posters (continued)

Contributed Posters (continued)

Contributed Posters (continued)

TECHNICAL PAPERS

Solid-State Sensors, Actuators, and Microsystems
Workshop

2000
Hilton Head Island, South Carolina
June 4 - 8

DESIGN OF AFM CANTILEVERS FOR COMBINED THERMOMECHANICAL DATA WRITING AND READING

William P. King, Thomas W. Kenny, and Kenneth E. Goodson
Department of Mechanical Engineering, Stanford University
Stanford, California, 94305-3030
**Graham Cross, Michel Despont, Urs Dürig, Mark I. Lutwyche, Hugo Rothuizen,
Gerd K. Binnig, and Peter Vettiger**
Micro- and Nanomechanics Group, IBM Research, Zurich Research Laboratory
CH-8803 Rüschlikon Switzerland

ABSTRACT

A Joule-heated atomic force microscope (AFM) cantilever writes a data bit by scanning over a polymer substrate. Heat and force applied to the polymer cause it to soften and flow, yielding an indentation of radius near 50 nm. Data is read by measuring the change in temperature of the heated cantilever as it follows the contour of an existing bit. Heat conduction governs the ultimate performance of a thermomechanical data storage device. The present work develops a finite-difference simulation of single-cantilever thermal and electrical behavior. Simulation results indicate design tradeoffs in power requirements, data writing speed, and data reading sensitivity. Scaling of single-cantilever predictions allows prediction of power requirements and ultimate data rates of cantilever arrays. The design tool predicts operating points for the present array cantilever, as well as for a proposed array cantilever. The framework established here considers for the first time comprehensive thermal, mechanical, and system-level requirements for data writing and reading in a thermomechanical data storage system.

INTRODUCTION

In today's magnetic data storage, the superparamagnetic effect limits the stability of the smallest magnetic domain, and will ultimately limit room-temperature magnetic storage technology at a data density of approximately 40 Gb/in² [1]. Several promising alternative data storage technologies that could surpass this limit involve the use of scanning-probe devices [2,3]. Thermomechanical data storage is one such technology [4-7].

Data Writing   **Data Reading**

"0" Higher Temperature   "1" Lower Temperature
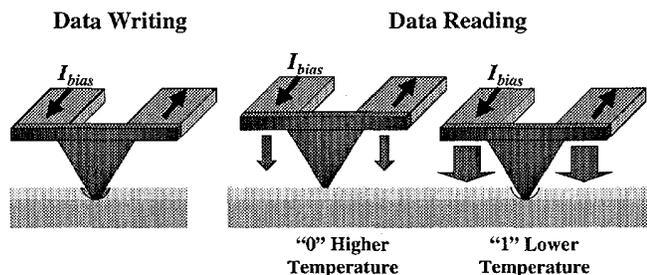
Figure 1. Thermomechanical data bit writing and thermal data reading with a Joule-heated AFM cantilever.

A Joule-heated AFM cantilever writes a data bit by scanning over a polymer substrate. Heat and force applied to the polymer cause it to soften and flow, thus writing a data bit, as shown in Fig. 1 [4-7]. Thermomechanical data storage with an electrically active silicon cantilever was first demonstrated to write data bits in a

thick polycarbonate substrate with a cantilever thermal time constant of approximately 0.4 ms [4]. Silicon cantilevers with small solid-state Joule heaters near the cantilever tip were fabricated specifically for application to thermomechanical data storage [5]. This allowed cantilever optimization, which reduced the cantilever heating time into the sub-microsecond range [6].

Binnig et al. [7] wrote very small thermomechanical data bits on the order of 40 nm in diameter, shown in Fig. 2, in a PMMA layer of 40 nm thickness [7]. The very thin polymer layer confines the polymer melted-region to dimensions on the order of the polymer thickness, and limits tip penetration and therefore data bit size.

150 nm ——   **40 nm** ——
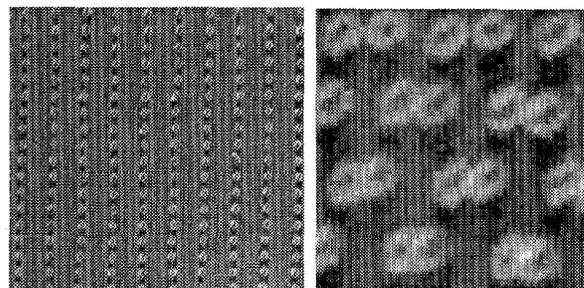
Figure 2. Tracks of thermomechanically written data bits (left) and bits written with high periodicity (right). The bit diameter and pitch of 40 nm corresponds to a data density of 0.4 Tb/in² [7].

Binnig et al. [7] also showed that existing pits can be located using the temperature signal from a heated cantilever. As the cantilever tip enters a previously written data bit, the reduction in thermal impedance between the cantilever and the substrate causes a lower temperature rise of the cantilever, as shown in Fig. 1. Measurement of the cantilever electrical resistance, which corresponds to cantilever temperature, determines the cantilever–substrate distance. The images in Fig. 2 were recorded by this thermal reading process.

In order to compete with traditional data storage technology, thermomechanical data storage must offer significant improvements over traditional data storage technology not only in data density but also in speed and power consumption. Array operation of heater-cantilevers allows higher data rates than a single cantilever configuration. Vettiger et al. [8] fabricated a 5×5 cantilever array, and used it to read topographical test structures with 200-nm features [9]. A 32×32 array of cantilevers was fabricated and wrote fields of data bits [10,11]. The most recent cantilever array is shown in Fig. 3.
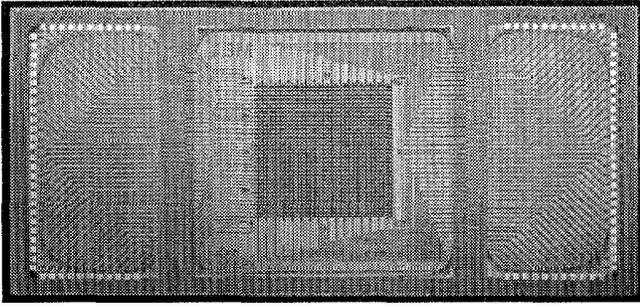
Figure 3. *32×32 (1024) array of cantilevers. From [12], © IEEE 1999.*

While much progress has been made to decrease both the time to write a data bit and the size of the bit written, no work has been done to optimize cantilevers for thermal data reading. There has also been no work to design cantilevers for explicit array operation, or to identify tradeoffs in cantilever design. The present work aims to understand the heating and cooling of single AFM heater-cantilevers by measuring cantilever heating and cooling and through detailed device simulations. An electrical and thermal simulation is developed for the operation of a single cantilever and verified by measurements. Scaling of the single-cantilever predictions explores cantilever design and determines the operating window of the present cantilever array and one possible alternative design. Data rate is a key figure of merit in data storage, as well as power consumption in a possible mobile computing application. The present work aims to frame device operation in terms of these parameters.

SINGLE-CANTILEVER MEASUREMENT AND SIMULATION

Chui et al. [13] modeled and measured thermal and electrical behavior of heater-cantilevers for the purpose of studying thermal runaway. This approach employed a finite-difference simulation of heating in the cantilever. The present work develops a similar model, which improves upon the previous approach by accounting for heat transfer in the air near the cantilever, and into the underlying substrate. This improvement allows accurate time resolution as well as prediction of thermal data reading sensitivity.

Measurements were made on a cantilever having a design similar to the cantilevers from the array in Fig. 3 to verify the simulation. The measurement cantilever has a thickness of 1 μm, a length of 50 μm, a width of 5 μm, and a heater area of 5 μm × 5 μm. The cantilever is boron-doped with 1×10^{17} cm^{-3} in the heater region and 1×10^{20} cm^{-3} in the legs, yielding a total room-temperature cantilever resistance of 924 Ω. The fabrication process for this cantilever is documented elsewhere [8]. The cantilever is placed in series with a sense resistor of 1 kΩ, and voltage pulses are applied to the circuit. The measurement of the voltage across the circuit and the sense resistor allow the cantilever resistance to be calculated. Measurements of the cantilever resistance are made during square voltage pulses of 1 μs, 10 μs, and 100 μs duration and a pulse amplitude of 1–15 V. Following the heating pulse, a 1 V bias is applied to the cantilever in order to measure cantilever cooling.

The measurements and predictions here are made for a stationary cantilever in air. The model accounts for heat generation in the cantilever and for conduction into the air and along the cantilever. Scaling arguments show that thermal

radiation and convection into the environment are small compared to conduction. The simulation parameters for the cantilever electrical properties and heater size geometry were used to fit the measured values. A model of the temperature-dependent intrinsic carrier generation in the cantilever is taken from Sze [14]. VLSI models calculate the temperature-dependent electrical resistivity of the doped silicon [15], and the thermal conductivity of the highly doped silicon cantilever is assumed to be 50 Wm^{-1}K^{-1} [16].

Figure 4 shows the measurement and prediction of the maximum cantilever temperature for a range of voltage pulses. The simulated temperature compares well with measurements. The measured cantilever cooling, not shown here, also agrees well with simulations. The longer pulses indicate the nonlinearity of the electrical properties of the cantilever at higher temperature.
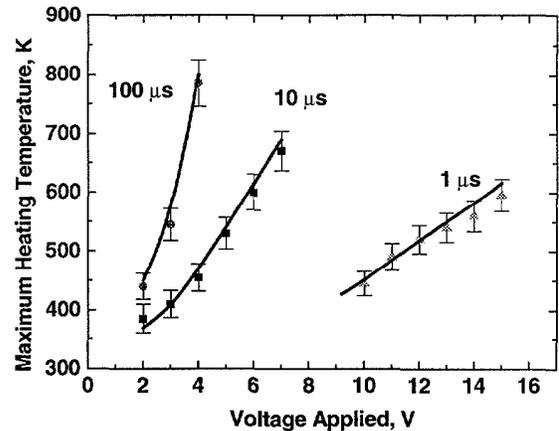
Figure 4. *Measurement and prediction of maximum temperature during heating pulses of duration 1 μs, 10 μs, and 100 μs. Data points represent measurements; solid lines represent prediction.*

SINGLE-CANTILEVER DESIGN AND ARRAY OPERATION

This section describes predictions made for the operation of a cantilever from the array shown in Fig. 3 as well as the operation of a proposed cantilever. An SEM image of a cantilever from the array is shown in Fig. 5, as well as a diagram of the proposed cantilever. The design analysis presented here aims to explore data rate and related power requirements.
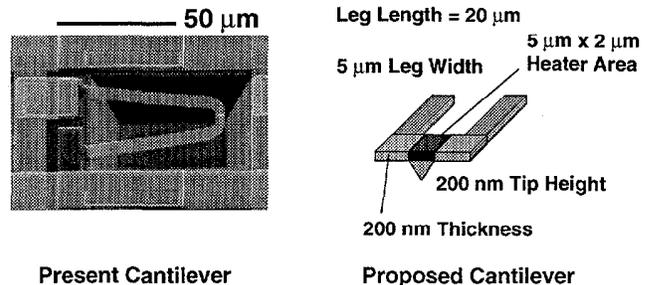
Figure 5. *SEM image of a cantilever from the array shown in Fig. 3, and a diagram of the proposed cantilever. The proposed cantilever could be fabricated with no additional fabrication requirements. The design details of both cantilevers pictured here are shown in Table 1.*

The simulations in this section employ the simulation tool used to make the predictions of cantilever heating shown in Fig. 4, with one modification: The simulations here model the cantilever near the polymer data surface, which is assumed to remain constant at room temperature. The maximum error induced by ignoring the temperature change in the substrate is estimated to be less than 5%.

A single cantilever operates most efficiently if it is heated only long enough to reach the thermomechanical writing temperature. The thermomechanical writing temperature is taken to be 350 °C, which is the measured writing temperature in thin PMMA [7,10]. In array operation, the cantilever can be selected again when it has cooled sufficiently, and moved to the appropriate position over the data media.

Table 1. Design details for present and proposed cantilevers.

	Present Cantilever	Proposed Cantilever
Heater Area	5 μm × 7 μm	2 μm × 5 μm
Thickness	500 nm	200 nm
Tip Height	500 nm	200 nm
Leg Width	10 μm	5 μm
Leg Length	50 μm	20 μm
Electrical Resistance At 25 °C	2120 Ω	2012 Ω
Resonant Frequency	220 kHz	330 kHz
Mechanical Limit on Repeated Data Writing	968 kHz	1.45 MHz

There are several constraints on cantilever design. It is desirable to minimize the power required to heat the cantilever with low operating voltage, which implies low electrical resistance. The electrical resistance of the cantilever heater region must however be much higher than any other element in the device. Successful thermal data reading has been accomplished for cantilevers having an electrical resistance in the range of 750 Ω to 4 kΩ. Reduction of the cantilever heater region size will generally reduce the cantilever thermal reading sensitivity. There is also a greater potential for electrical and thermal noise in cantilevers with very small heater regions.
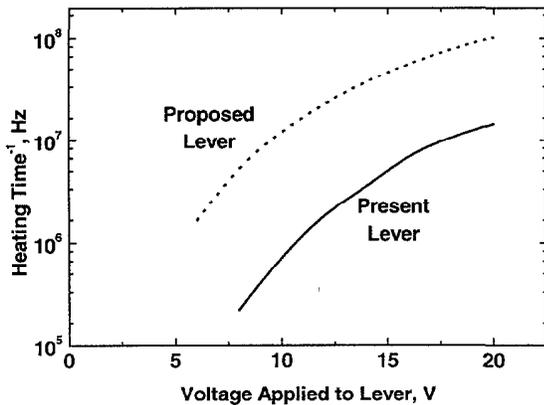
Figure 6. Predictions of the time required to reach the bit writing temperature as a function of voltage applied to the cantilever.

Figure 6 shows predictions for the time required for the cantilever to reach 350°C, the measured writing temperature in PMMA. The predictions indicate that the lowest applied voltage at which the present cantilever can induce a data bit in a PMMA polymer layer is close to 8 V, which agrees well with measurements.

Figure 7 shows the average power dissipated in the cantilever while writing a single data bit. These predictions show that the proposed cantilever does not have a significant reduction in power at a given applied voltage. However, predictions show that proposed cantilever could write bits at a lower minimum voltage and power than the present cantilever.
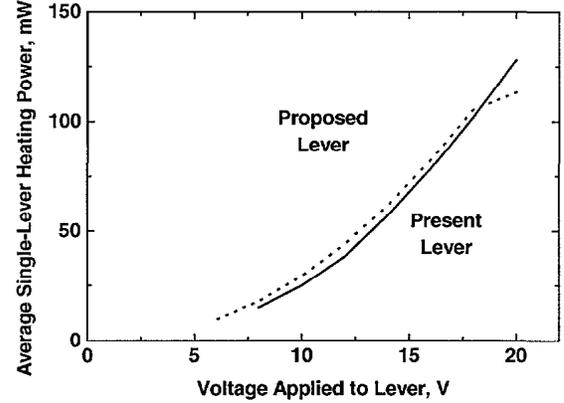
Figure 7. Predictions of average power required for a single cantilever to write a data bit as a function of voltage applied to the cantilever.

The thermal limit on the time in which a cantilever can repeatedly write data bits is governed by the time it takes for a cantilever to heat and then cool. The present work defines the cooling time as the time required for the cantilever to cool to one-half its maximum heating temperature. Figure 8 shows predictions of the cantilever total thermal cycle time as a function of heating power. Although higher power generally corresponds to faster operation, there is an upper limit on the thermal cycle time that corresponds to the finite heat capacity of the cantilever heater region. At longer heating times, heat diffuses farther along the cantilever, increasing the time required to then cool the cantilever. This explains the strong dependence of the thermal cycle time on cantilever heating power at longer heating times.

Another limit on the time in which a cantilever can write successive data bits is the cantilever mechanical resonant frequency. In general, it is desirable to have a cantilever resonant frequency that is greater than the bit writing frequency. King et al. [17] showed that this is in fact not a strict requirement, as cantilevers in the array are not free to vibrate. Instead, the cantilevers are continuously forced against the polymer data layer, and penetrate only until the tip reaches the substrate beneath the thin polymer data layer. A more appropriate mechanical resonance figure of merit is the first harmonic of a cantilever with a hinged end, rather than a free end. The first resonance of a cantilever with a hinged end is approximately 4.4 times greater than the first resonance of a cantilever with a free end. The resonant frequency of a freely vibrating cantilever from the present array is approximately 220 kHz, and that of the proposed cantilever is estimated at 330 kHz. The appropriate values for the mechanical limit of repeated writing are 968 kHz for the present cantilever and 1.45 MHz for the proposed cantilever.

The lowest possible power at which a thermomechanical data storage device could operate is of interest for mobile computing applications. At the lowest possible device power, each cantilever would be selected independent of other cantilevers, such that only one cantilever is heated at a time. In this case, the data storage device would operate at its lowest data rate.

The array in Fig. 3 is designed such that an entire row can read or write in concert, thus increasing data rate. The heating of an entire row of cantilevers at one time will increase the device power requirements. A possible data writing rate for multiplexed array operation is calculated in the following manner. The product of the cantilever heating time and the number of cantilever rows calculates the time for all the rows in the array to be selected once. The array cycle time must be greater than the total thermal cycle time of one row, as a cantilever must cool before it can move to a new location and write again. Another check must be made that the array cycle time is also greater than the cantilever mechanical response time. The number of cantilevers in the array is the number of bits that can be written in one array cycle.
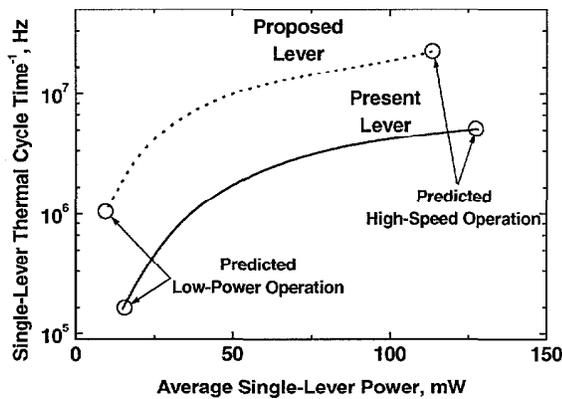
Figure 8. Single-cantilever total heating and cooling time as a function of applied power. The circles represent operating points for multiplexed cantilever arrays, listed in Table 2.

Table 2. Possible continuous-writing operation points for the present cantilever and proposed cantilever designs.

	Present Cantilever	Proposed Cantilever
Lowest Power	0.3 Mb/s	2.0 Mb/s
	15.0 mW	9.5 mW
Highest Speed (Multiplexed)	480 Mb/s	10.0 Gb/s
	4.0 W	13.0 W
	32×32 Array	100×100 Array

The circles in Fig. 8 are not data points, but represent possible operating points for continuous writing with an array of cantilevers. Table 2 describes these operation points in more detail. Two cases explore the boundaries of cantilever array operation, which are the highest possible data rate, and lowest possible power.

It is predicted that the present array could operate at its lowest possible power at 0.3 Mb/s. This operation point does not exploit array multiplexing, and therefore the single-cantilever power determines the device power. It is predicted that low-power

operation of an array of the proposed cantilevers would have a data rate of 2.0 Mb/s and a power requirement of 9.5 mW.

The highest writing rate of the present cantilever array is governed by the time required for cantilever heating at 480 Mb/s. Cantilever heating time also limits the fastest possible writing rate of a 100×100 array of proposed cantilevers at 10.0 Gb/s. Table 2 shows the increased power requirement of array operation. A 100×100 array consisting of the proposed cantilevers would have a device power requirement of 13 W. It is important that a thermomechanical data storage device operated in this manner is cooled such that the power dissipated in the device does not melt the polymer data layer.

The power requirements and data rates presented here are for continuous writing. Arrays of cantilevers could operate such that they are continuously heated, and therefore do not need to be heated from room temperature for every bit-writing event. Also, data reading requires less power. Therefore typical device power for combination writing and reading operation could be substantially lower than the power required for continuous writing. We envision adaptive operation of a thermomechanical data storage device that could actively control where it operates on the power-speed curve.

Thermal data-reading sensitivity is calculated as the change in steady-state thermal impedance for a small change in cantilever tip height. The maximum thermal reading signal is obtained when the cantilever is heated to the highest temperature at which the polymer data surface is not modified, which is approximately 325 °C [10]. The cantilever power at this operation point is approximately 5 mW. Predictions are shown in Fig. 8. The prediction of the sensitivity of the current cantilever array compares well with the measured value reported by Lutwyche et al. [10], also shown in Fig. 9.
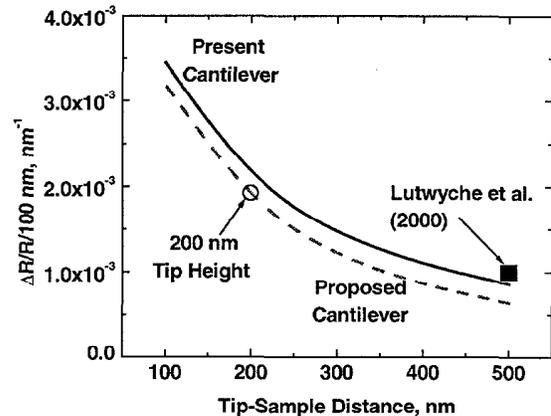
Figure 9. Sensitivity predictions for the current array cantilever and the proposed array cantilever. The sensitivity is calculated for reading a data bit of depth 100 nm. The data represents the operating point of Lutwyche et al. [10].

In general, shorter and thinner cantilevers with smaller tips are more sensitive. While the proposed cantilever is thinner than the current cantilever, the predicted sensitivity of the proposed cantilever is slightly less than the present cantilever at a given cantilever–sample distance. This is due to the much smaller heater area and narrower legs of the proposed cantilever. It is predicted that a tip height of 200 nm for the proposed cantilever will result in a sensitivity of almost double that of the present cantilever, as shown in Fig. 8.

Although the sensitivity calculation is made here at steady state, the result also indicates the transient value provided that two conditions are met: First, sufficient time must have elapsed for heat to diffuse across the cantilever–sample air gap, which is not necessarily the time required for the cantilever to reach steady state. For example, the time for heat to diffuse across a 200-nm air gap is approximately 5 ns. Secondly, the signal from the cantilever must be integrated over a sufficiently long time to achieve an acceptable signal-to-noise ratio. For transient sensing, we envision integrated read electronics fabricated directly on the cantilever array.

CONCLUSIONS

An electrical and thermal simulation of Joule-heated AFM cantilevers for thermomechanical data storage is developed and verified by comparison to measurements of cantilever heating. The simulation is then used to investigate the heating of the present array cantilever and a proposed array cantilever. Scaling of single-cantilever predictions allows prediction of array operation.

The framework established here considers for the first time comprehensive thermal, mechanical, and system-level requirements for data writing and reading in a thermomechanical data storage system. The approach developed here is valid for polymers of various melting temperatures, and for thermal data writing and reading in other media.

ACKNOWLEDGMENTS

REFERENCES

1. E. Grochowski and R. F. Hoyt, "Future Trends in Hard Disk Drives," *IEEE Transactions on Magnetics, 32,* 1850 (1996).

2. E. B. Cooper, S. R. Manalis, H. Fang, H. Dai, K. Matsumoto, S. C. Minnie, T. Hunt, and C. F. Quate, "Terabit-Per-Square-Inch Data Storage with the Atomic Force Microscope," *Applied Physics Letters, 75,* 3566 (1999).

3. H. J. Mamin, R. P. Ried, B. D. Terris, and D. Rugar, "High Density Data Storage Based on the Atomic Force Microscope," *Applied Physics Letters, 87,* 1014 (1999).

4. H. J. Mamin, "Thermal Writing Using a Heated Atomic Force Microscope Tip," *Applied Physics Letters, 69,* 433 (1996).

5. B. W. Chui, T. D. Stowe, T. W. Kenny, H. J. Mamin, B. D. Terris, and D. Rugar, "Low-Stiffness Silicon Cantilevers for Thermal Writing and Piezoresistive Readback with the Atomic Force Microscope," *Applied Physics Letters, 69,* 2767 (1996).

6. B. W. Chui, T. D. Stowe, Y. S. Ju, K. E. Goodson, T. W. Kenny, H. J. Mamin, B. D. Terris, R. P. Ried, and D. Rugar, "Low-Stiffness Silicon Cantilevers with Integrated Heaters and Piezoresistive Sensors for High-Density AFM Data Storage," *Journal of Microelectromechanical Systems, 7,* 69 (1998).

7. G. Binnig, M. Despont, U. Drechsler, W. Häberle, M. Lutwyche, P. Vettiger, H. J. Mamin, B. W. Chui, and T. W. Kenny, "Ultrahigh-Density Atomic Force Microscopy Data Storage with Erase Capability," *Applied Physics Letters, 76,* 1329 (1999).

8. P. Vettiger, J. Brugger, M. Despont, U. Drechsler, U. Dürig, W. Häberle, H. Rothuizen, R. Stutz, R. Widmer, and G. Binnig, "Ultrahigh Density, High-Data-Rate NEMS-Based AFM Data Storage System," *Microelectronic Engineering, 46,* 11 (1999).

9. M. Lutwyche, C. Andreoli, G. Binnig, J. Brugger, U. Drechsler, W. Häberle, H. Rohrer, H. Rothuizen, P. Vettiger, G. Yaralioglu, and C. Quate, "5 × 5 AFM Cantilever Arrays: A First Step Towards a Terabit Data Storage Device," *Sensors and Actuators A, 73,* 89 (1999).

10. M. I. Lutwyche, G. Cross, M. Despont, U. Drechsler, U. Dürig, W. Häberle, H. Rothuizen, R. Stutz, R. Widmer, G. K. Binnig, and P. Vettiger, "Millipede – A Highly-Parallel Dense Scanning-Probe-Based Data Storage System", *Technical Digest of IEEE International Solid-State Circuits Conference,* San Francisco, CA (2000), pp. 126-127.

11. P. Vettiger, M. Despont, U. Drechsler, U. Dürig, W. Häberle, M. I. Lutwyche, H. Rothuizen, W. Stutz, W. Widmer, and G. Binnig, "The Millipede: More Than One Thousand Tips for Future AFM Data Storage," *IBM Journal of Research and Development, in press.*

12. M. Despont, J. Brugger, U. Drechsler, U. Dürig, W. Häberle, M. Lutwyche, H. Rothuizen, R. Stutz, R. Widmer, H. Rohrer, G. K. Binnig, and P. Vettiger, "VLSI–NEMS Chip for AFM Data Storage," *Technical Digest 12th IEEE International Micro Electro Mechanical Systems Conference MEMS '99,* Orlando, FL, 01/99, IEEE, Piscataway (1999), pp. 564-569.

13. B. W. Chui, M. Asheghi, Y. S. Ju, K. E. Goodson, T. W. Kenny, and H. J. Mamin, "Intrinsic-Carrier Thermal Runaway in Silicon Microcantilevers," *Microscale Thermophysical Engineering, 3,* 217 (1999).

14. S. M. Sze, *Physics of Semiconductor Devices,* Wiley, New York (1981).

15. S. Selberherr, "Process and Device Modeling for VLSI," *Microelectronics Reliability, 24,* 225 (1984).

16. M. Asheghi, K. Kurabayashi, K. E. Goodson, R. Kasnavi, and J. Plummer, "Thermal Conduction in Doped Single-Crystal Silicon Films," *Proceedings of the 33rd National Heat Transfer Conference HTD99-186,* Albuquerque, NM, 8/99, American Society of Mechanical Engineers, New York (1999), pp. 1-13.

17. W. P. King, J. G. Santiago, T. W. Kenny, and K. E. Goodson, "Modeling and Prediction of Sub-Micrometer Heat Transfer during Thermomechanical Data Storage," *Proceedings of the ASME IMECE MEMS-Vol. 1,* Nashville, TN, 11/99, American Society of Mechanical Engineers, New York (1999), pp. 583-588.

A FLAT HIGH-FREQUENCY SCANNING MICROMIRROR

Robert A. Conant, Jocelyn T. Nee, Kam Y. Lau, and Richard S. Muller
Berkeley Sensor & Actuator Center, University of California, Berkeley
Berkeley, CA 94720-1774

ABSTRACT

We present the Staggered Torsional Electrostatic Combdrive (STEC) fabrication process that creates thick mirrors with high-torque actuators, thereby allowing high-speed optical scanning with large angular deflections. We have demonstrated a 550 μm-diameter mirror capable of scanning an optical angle of 24.9° at its 34 kHz resonant frequency with less than 30 nm static and dynamic deformation. The optical resolution is 350 pixels – very near the diffraction-limited resolution of 355 pixels with 655 nm wavelength light.

INTRODUCTION

Scanning mirrors are used in many applications: barcode readers, laser printers, confocal microscopes, and fiber-optic network components. There are significant limitations to the performance of macro-scale scanners; in particular, their scanning speed, power consumption, cost, and size often preclude their use in portable systems. In addition, higher-frequency optical scanning could enable new applications that are not practical with conventional scanning mirrors, such as raster-scanning projection video displays [1], and would significantly improve the performance of scanning mirrors in existing applications, such as laser printers. MEMS optical scanners promise to enable these new applications, and dramatically reduce the cost of optical systems.

However, previously demonstrated MEMS mirrors have not been able to simultaneously meet the requirements of high scan speed *and* high resolution. A plethora of micromirror designs have been presented, but none has been able to satisfy the potential of MEMS: a high-speed, high-performance scanning mirror. For example, surface-micromachined scanning mirrors actuated with electrostatic combdrives have been shown to operate at high scan speeds (up to 21 kHz), but static and dynamic mirror deformation limits the resolution to less than 20% of the diffraction-limited resolution. Magnetically actuated mirrors have been demonstrated with high speed and large amplitude [2], but have not demonstrated high resolution, and often require off-chip actuation.

In order to take advantage of the unique capabilities of micromachining technology, we have fabricated a new type of micromirror capable of high-speed, high-resolution scanning with low power consumption.

THEORY OF OPERATION

STEC Micromirror

Figure 1 shows a schematic of the Staggered Torsional Electrostatic Combdrive (STEC) micromirror, which is made from two layers of single-crystal silicon separated by a 1.7 μm-thick silicon dioxide layer. The mirror, torsion hinge, and moving comb teeth are in the top silicon layer, and the fixed comb teeth are in the bottom silicon layer. Applying voltage between the top and bottom layers attracts the moving comb teeth to the fixed comb teeth, thus exerting torque on the mirror and causing it to tilt. The torsion hinges, which are anchored to the bottom silicon layer, provide restoring torque. Similar torsional actuators have been

Mirror Structure

Actuated mirror
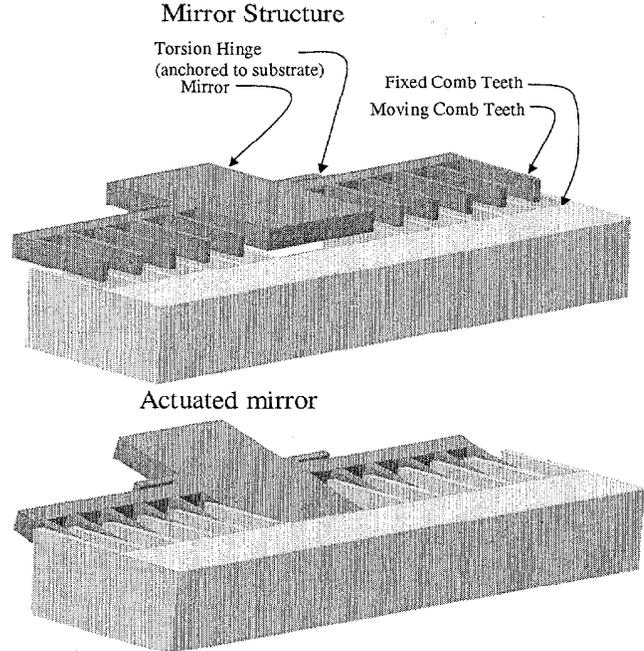
Figure 1. Schematic of a STEC micromirror. Applying a voltage between the top and the bottom layers generates a torque on the mirror, thereby causing it to tilt.

described earlier [3][4]. SEMs of the completed STEC micromirror are shown in Figure 2.

This torsional electrostatic combdrive offers several advantages over other electrostatic-actuator designs. First, the actuator applies torque to the mirror directly – there are no hinges to couple linear motion of an actuator into torsional mirror motion. This greatly simplifies the design of the structure, and makes post-fabrication assembly steps unnecessary.

Second, the actuator starts in an unbalanced state and is capable of static mirror positioning as well as resonant scanning. Previously demonstrated balanced torsional electrostatic actuators have been very promising for resonant operation [5], but are not capable of static mirror positioning.

Third, the torsional combdrive offers an advantage over gap-closing actuators because the energy density in the combdrive is higher than that in a gap-closing actuator, thereby allowing larger scan angles at high resonant frequencies.

Optical Resolution

The optical resolution – defined as the ratio of the optical-beam divergence and the mirror scan angle – is an essential performance metric for a scanning mirror. For a perfectly flat mirror under uniform illumination, the farfield intensity distribution is an Airy pattern, which has a full-width-half-max half-angle beam divergence α (the resolution criteria used for video displays [6]) given by

$$\alpha = \frac{1.03\lambda}{D} \qquad [1]$$
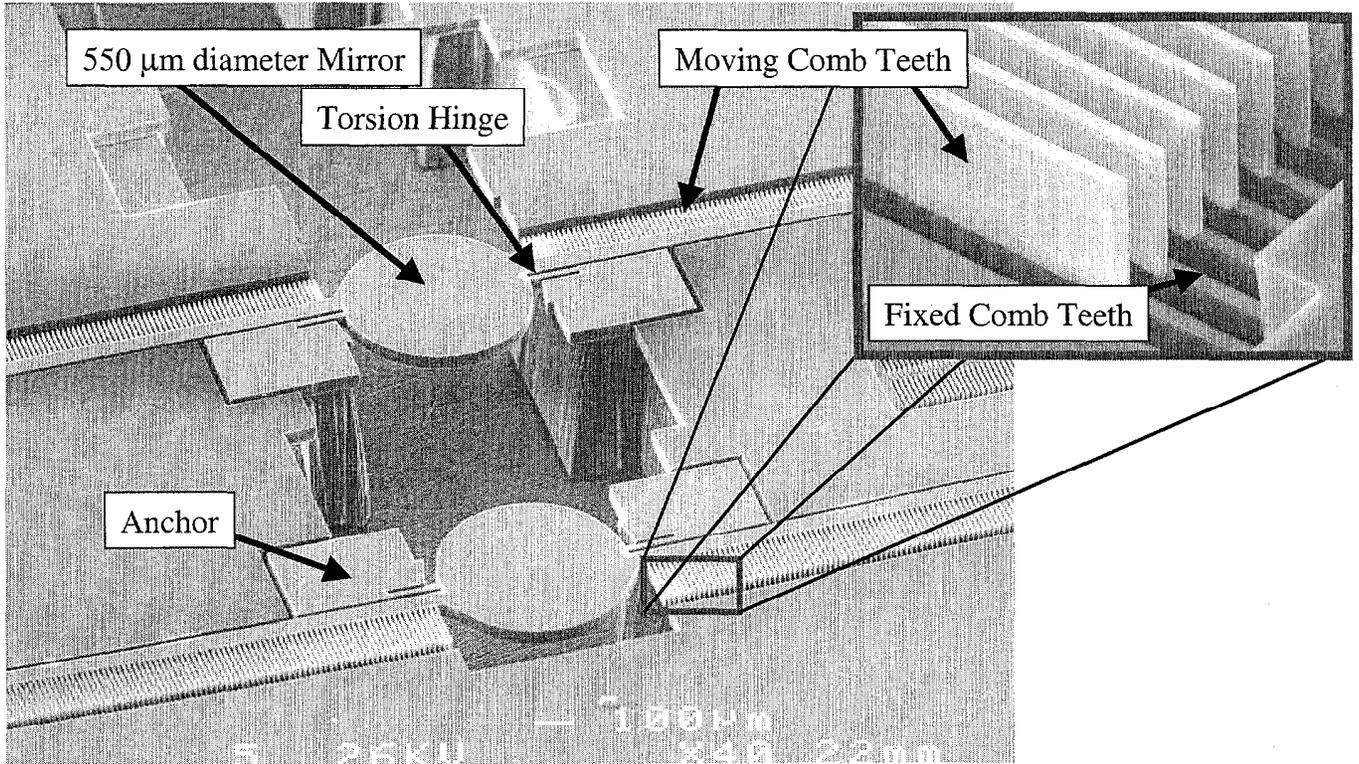
Figure 2. SEM of two STEC micromirrors. The SEM chuck is visible through the wafer below the mirror. The inset SEM in the upper right shows a close-up view of the fixed and moving combteeth.

where λ is the wavelength of the incident light, and D is the mirror diameter. The resulting optical resolution N is

$$N = \frac{4\theta}{\alpha} = \frac{4\theta\, D}{1.03\lambda} \qquad [2]$$

where θ is the mechanical half-angle mirror scan (the total optical scan is 4θ).

Dynamic mirror deformation can also contribute to beam divergence, thereby decreasing the optical resolution. For a mirror where the torsion hinge is the dominant compliance, the nonplanar surface deformation δ of a rectangular scanning mirror of half-length L with angular acceleration $(2\pi f)^2 \theta$ (where f is the scan frequency) is [7]

$$\delta = 0.183 \frac{\rho(1 - v^2)(2\pi f)^2 \theta}{E t^2} L^5 \qquad [3]$$

where ρ is the material density, v is Poisson's ratio, E is Young's modulus, and t is the mirror thickness.

The Rayleigh limit, the maximum amount of surface deformation tolerable without significant degradation in image quality, allows a peak-to-valley surface deformation of $\lambda/4$ [8]. For a 550 μm-long (275 μm-half-length) rectangular single-crystal-silicon mirror of thickness 50 μm, half-angle mechanical scan 6.25°, and resonant frequency 34 kHz, the calculated dynamic deformation is 8 nm – much lower than the Rayleigh limit for 655 nm light (164 nm). For comparison, a 550 μm-long surface-micromachined mirror of thickness 1.5 μm maintains the surface flatness within the Rayleigh limit only up to a frequency of 4.6 kHz.

FABRICATION

The STEC micromirrors are fabricated using deep reactive-ion etching and a bond-and-etchback process with a buried pattern. The process flow is shown in Figure 3.

First, a silicon wafer is oxidized in steam at 1000°C to grow 0.2 μm of thermal oxide. This wafer is patterned with the BURIED pattern, and 100 μm-deep trenches are etched into the wafer using an STS deep reactive-ion etcher to form the fixed combteeth. This wafer, along with another wafer having 1.5 μm of thermal oxide, is cleaned and the two wafers are bonded together with the fixed combteeth at the wafer interface. The bonded wafer pair is then annealed at 1100°C for one hour to increase the bond strength [9].

Next, the bonded wafer is ground and polished to leave a 50 μm-thick layer of silicon above the oxide interface. The resulting silicon-on-insulator (SOI) wafer is oxidized at 1100°C in a steam ambient to form a 1.1 μm-thick oxide layer on the top and bottom. Alignment between the subsequent patterns and the buried combteeth is achieved by etching a window in the top-layer silicon above the two dice used to align the mask to the wafer. These two dice are sacrificed to achieve the necessary alignment precision between the moving combteeth (on the top layer) and the fixed combteeth (on the bottom layer).

The FRONTSIDE pattern, which defines the moving comb teeth, the mirror, and the torsion hinges, is then patterned and etched into the top oxide layer (the pattern is etched into the silicon in a later step). The alignment of this step is critical because misalignment between the moving combteeth and the fixed combteeth can lead to instability in the torsional combdrive. We have typically been able to achieve better than 0.2 μm alignment between the BURIED pattern and the FRONTSIDE

Etch wafer1

Bond wafer1
and wafer2

Grind and
polish wafer2

Backside etch
wafer1

Pattern and
etch wafer1

Release

■ Silicon (wafer 1)
▨ Silicon (wafer 2)
▧ Oxide
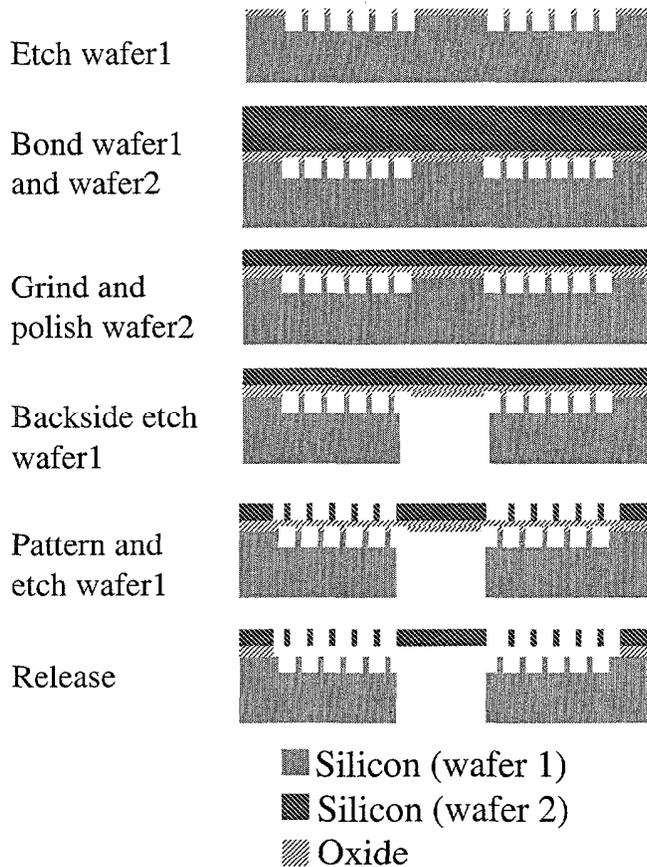
Figure 3. STEC micromirror fabrication process flow.

pattern using a GCA i-line wafer stepper, and this alignment error has not limited the STEC micromirror performance.

Next, the oxide and silicon on the backside of the bottom wafer are patterned with the HOLE layer, and the entire bottom wafer is etched to open an optical path underneath the micromirror. The silicon on the top layer is then etched using the previously patterned top oxide layer as an etch mask. The structure is released with a timed HF etch to remove the sacrificial oxide film below the combteeth and mirror. Finally, a 100 nm-thick aluminum film is evaporated onto the bottom of the mirror to increase the reflectivity for visible light to approximately 92% [10].

RESULTS

The STEC mirror excels in all of the critical performance criteria: cost, resolution, scan speed, scan repeatability, size, power consumption, and reliability. This section discusses measurements of four of these performance criteria for one STEC mirror design.

Resolution

The surface deformation of the micromirror was characterized using the stroboscopic interferometer reported in [11]. Figure 4 shows the dynamic deformation of the micromirror at the end of the scan where the dynamic deformation reaches its maximum. The total deformation is less than 30 nm, considerably below the Rayleigh limit, and does not significantly reduce the optical resolution.

Figure 5 shows a photograph of the scan of a STEC mirror with a 655 nm laser pulsed synchronously with the mirror scan to generate 43 spots. The mirror scan is sinusoidal, so the timing of

the laser pulses was adjusted across the scan to produce equally spaced spots. Figure 5 also shows a close-up image of three individual spots captured with a CCD camera. By fitting a Gaussian curve to each of these three spots, we were able to determine the full-width-half-max beam width and the distance between the spots. The spot size and separation at eight different regions across the scan give the measured total optical resolution of 350 pixels. The resolution of this 550 μm-diameter mirror with 24.9° optical scan and 655 nm laser light is near the diffraction-limited resolution of 355 pixels from Eq. [2].

Scan Speed

STEC micromirrors have been demonstrated with diameters of 550 μm and resonant frequencies up to 42 kHz – almost an order of magnitude faster than commercially available optical scanners. Larger STEC mirrors have also been fabricated (up to 2 mm) with lower resonant frequencies. Figure 6 shows the frequency response of a 550 μm-diameter micromirror.

The main limitation of macro-scale scanners comes from the dynamic deformation described by Eq.[3] – the dynamic deformation scales as the fifth power of the mirror length, so large mirrors scanning at high speeds will have considerable dynamic deformation. For example, a 10 mm-diameter, 1 mm-thick mirror with a mechanical scan of ±6.25° maintains less than 164 nm dynamic deformation (the Rayleigh limit for 655 nm light) up to a frequency of only 2.2 kHz. Large-scale mirrors cannot achieve the speeds demonstrated with the STEC micromirrors without severe dynamic deformation or very thick mirrors.

Power Consumption

High-speed scanners require more torque than low-speed scanners to reach the same scan angle. In order to generate the torque necessary for large angle, high-frequency operation of the STEC micromirror, we use fairly high voltages. The 550 μm-diameter mirror with resonant frequency 34 kHz requires a 171 Vrms input sine wave for a total optical scan of 24.9°. To simplify mirror testing and operation, we use a small (1 cm^3) 25:1 transformer, allowing us to use a conventional 0-10 V function generator to drive the scanning mirrors with a sinusoidal waveform of amplitude up to 250 V. The use of the transformer also provides efficient power conversion, so the power consumption of the entire system can be much lower than systems requiring high-voltage power supplies and opamps.

This power consumption is the sum of the power dissipation in the drive electronics and the power dissipated by air and material damping. The power consumption due to damping is

$$P = \frac{1}{2}b\theta^2\omega^2 = \frac{1}{2}\frac{k}{Q}\theta^2\omega \qquad [4]$$

where k is the torsional spring stiffness, b is the torque damping factor, θ is the mechanical scan half angle (the total optical scan is $\pm 2\theta_0$), ω is the resonant frequency, and Q is the resonant quality factor. For the 34 kHz 550 μm-diameter mirror scanning 25° optical (±6.25° mechanical), the calculated stiffness $k = 3.93 \times 10^5$ Nm/radian, the measured resonant quality factor $Q = 273$, so the power consumption due to damping from Eq. [4] is 0.18 mW. Vacuum packaging can be used to reduce the viscous damping, and thereby decrease the power consumption.

The measured power consumption is 6.8 mW, indicating that the majority of the power consumption is in charging and discharging the parasitic capacitance and losses in the transformer power conversion.
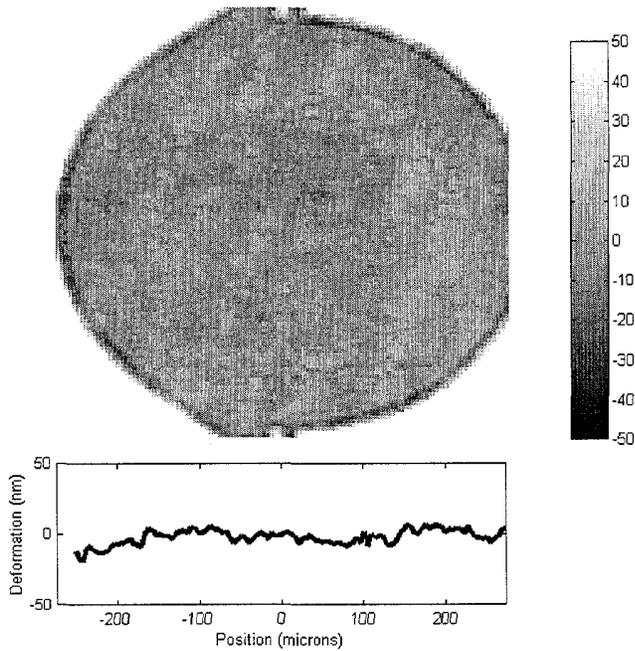
Figure 4. Static and dynamic deformation of the 550 μm-diameter mirror is less than 30 nm. The graph below the grayscale surface height map shows the measured mirror deformation along the center of the mirror perpendicular to the torsion hinges.

Figure 5. Optical scan of STEC micromirror operating at 34 KHz, with close-up image of spots and intensity distribution plot. The images recorded on the photograph and the CCD, which were red-on-black, were converted to black-on-white for this figure.

Reliability

The STEC micromirror is expected to be extremely reliable due to its simple structure. It is predicted that the failure point for the structure will be the torsion hinges (at the point of highest strain). The maximum strain in a 50 μm-thick, 15 μm-wide, 150 μm-long hinge (the hinge used for the 550 μm-diameter mirror with resonant frequency of 34 kHz) with a total scan of ±6.25° is approximately 1.8%. Mirrors have been operated at this level for over 200 million cycles without any noticeable degradation in performance, but this strain may be too high for acceptable long-term reliability. Future designs will use wider, longer hinges to reduce the maximum strain while retaining the same stiffness.

Figure 6: Frequency response of STEC micromirror.

CONCLUSIONS

The potential of micromachined mirrors over conventional scanning mirrors – high scan speed, small size, and low cost with diffraction-limited optical performance – has been demonstrated with the Staggered Torsional Electrostatic Combdrive (STEC) micromirror. The scan speed of the STEC micromirror is difficult to achieve with large-scale optical scanners, and exceeds the performance of previously demonstrated micromachined scanning mirrors.

REFERENCES

1. R.A. Conant, P.M. Hagelin, U. Krishnamoorthy, O. Solgaard, K.Y. Lau, R.S. Muller, "A Full-Motion Video Display Using Micromachined Scanning Micromirrors," proc. Transducers '99, Sendai, Japan (1999), pp. 376-379.
2. A. Garnier, T. Bourouina, H. Fujita, T. Hiramoto, E. Orsier, J-C. Peuzin, "Contactless actuationof bending and torsional vibrations for 2d-optical-scanner application," proc. Transducers '99, Sendai, Japan (1999), pp. 1876-1877.
3. Z.J. Yao, N.C. MacDonald, N.C. "Single crystal silicon supported thin film micromirrors for optical applications," Optical Engineering, 36 (no.5), pp. 1408-13 (1997).
4. J.-L.A Yeh, J. Hongrui; N.C. Tien, "Integrated polysilicon and DRIE bulk silicon micromachining for an electrostatic torsional actuator.", JMEMS, 8 (no.4), pp. 456-65 (1999)
5. H. Schenk, P. Durr, D. Kunze, H. Lakner, H. Kuck, "An electrostatically excited 2D-micro-scanning-mirror with an in-plane configuration of the driving electrodes," proc. MEMS 2000, Miyazaki, Japan (2000), pp. 473-478.
6. J. Hagerman, "Optimum spot size for raster-scanned monochrome CRT displays", Journal of the SID, 1 (no.3), pp. 367-369 (1993)
7. P.J. Brosen, "Dynamic Mirror Distortions in Optical Scanning," Applied Optics Vol. 11, No. 12, pp. 2987-2989, 1972.
8. D.J. Jacobs, Fundamentals of Optical Engineering, 1st ed. New York: McGraw-Hill, 1943.
9. W.P. Maszara, G. Goetz, A. Caviglia, and J.B. McKitterick, "Bonding of silicon wafers for silicon-on-insulator," Journal of Applied Physics, 64(10), pp. 4943-4950 (15 November 1988).
10. G.R. Elion, H.A. Elion, Electro-Optics Handbook, New York: Marcel Dekker, Inc., 1979.
11. M. Hart, R.A. Conant, K.Y. Lau, R.S. Muller, "Time-resolved measurement of optical MEMS, using stroboscopic interferometry," proc. Transducers '99, Sendai, Japan (June 1999), pp. 470-473.

A MICROMACHINED TORSIONAL MIRROR FOR TRACK FOLLOWING IN MAGNETO-OPTICAL DISK DRIVES

Joe Drake and Hal Jerman
Seagate Research
1870 Lundy Avenue, San Jose, CA 95131-1826

ABSTRACT

An electrostatically actuated polysilicon surface micromachined torsional mirror has been used as a fine actuator for track following in Optically Assisted Winchester (OAW™) magneto-optical disk drives. OAW™ technology combines a Winchester-style voice coil motor (VCM) as a coarse actuator with light delivery via optical fibers and a read/write head containing microfabricated optical elements, a coil, and a silicon chip containing the micromachined mirror. The mirror achieves a resonance of at least 35 kHz, an angular deflection range of +/- 2 degrees prior to electrostatic pull-in, and a pull-in voltage less than 150 Volts. Novel fabrication techniques allow for an RMS roughness of less than 3 nm on the reflecting surface of the mirror and mirror plate radius of curvature of greater than 10 cm over a temperature range of 5 to 55 C. Servo bandwidth for OAW disk drives has been demonstrated at 3-5 kHz while using the mirror for track following at greater than 35 thousand tracks per inch (ktpi).

INTRODUCTION

In the continuing effort to increase data storage density, the use of a dual stage servo system in hard disk drives has received considerable research attention. Several researchers [1, 2] have described the use of a piezoelectric push-pull actuator mounted between the coarse actuator arm and the head suspension to finely position the entire suspension and head assembly over the data tracks. Typically the reported servo bandwidth of such arrangements is on the order of a few kHz. Imamura, et. al. [3] demonstrated a micromachined suspension/slider concept wherein the entire slider is translated in the cross-track direction by a parallel plate electrostatic actuator. The reported bandwidth of this actuator was 34 kHz (without the slider present) with an actuation voltage of 50 Volts and a maximum displacement of 0.55 µm but the system servo bandwidth was not reported. Horsley, et. al. [4] also used a parallel plate electrostatic micromachined actuator for fine translation of the slider (± 1 µm) at ± 40 Volts. This actuator was constructed of polysilicon and the resulting servo bandwidth of the dual-stage system was 1.2-2.4 kHz.

Magneto-optical (MO) data storage presents the possibility of an elegant solution to the need for a second stage actuator. For MO data storage the fine actuator need only steer the focal point of the optical spot rendering unnecessary the movement of relatively massive components such as suspensions or sliders. Sekimura, et. al. [5] recognized this feature and described a bulk micromachined torsional mirror as a fine actuator mounted on an optical head in an MO data storage system. This mirror had an angular rotation range of ± 0.5 mrad (± 0.03 degrees) for an actuation voltage of 5 Volts and a bandwidth in excess of 50 kHz.

Optically Assisted Winchester (OAW™) [6] technology presents a data storage approach whereby the familiar and well-developed advantages of conventional Winchester style disk drives are integrated with optical elements that provide for MO data storage. In OAW™ technology, a rotary voice coil motor (VCM) actuator is used for coarse positioning of a suspension-mounted read/write head. Laser light is delivered to the head via an optical fiber mounted onto it (see Figure 1). A micromachined torsional mirror is mounted onto the trailing edge of the head and reflects the laser light that emanates from the fiber thus redirecting it normal to the MO storage media on the disk surface. Before reaching the disk, the laser spot is focussed through an aperture in a magnetic coil by a microfabricated-molded lens. By rotating about the axis defined by the torsional hinges, the reflective mirror plate translates the position of the focussed laser spot in the cross track direction thus providing the fine actuation for the dual-stage servo system (see Figure 2). Data is written to the MO media by heating the media to the Curie point with the laser spot while the coil-generated magnetic field is applied. Data is read back using the Kerr effect; laser light reflected off of a written spot in the media undergoes a slight rotation in its polarization compared to the reflection off of an unwritten spot. Thus the polarization of the reflected light determines the bit polarity.
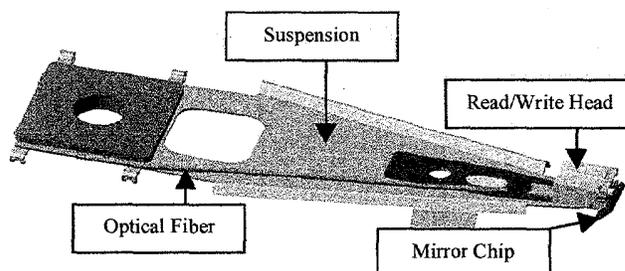
Figure 1. OAW™ Suspension assembly with read/write head and optical fiber. Micromachined mirror is fabricated onto the mirror chip (drawing courtesy of Mike Darling).
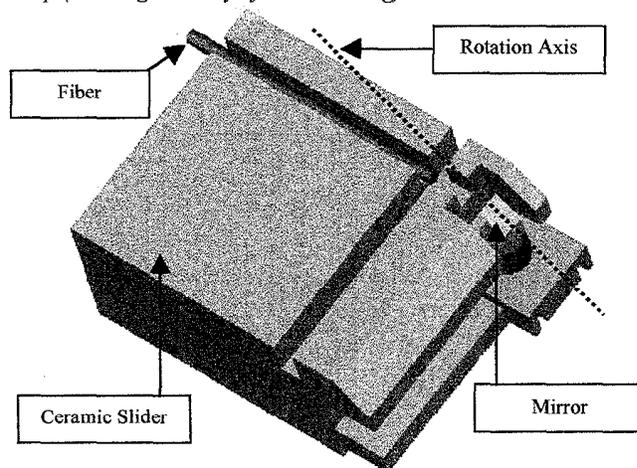
Figure 2. Schematic diagram showing how the torsional mirror rotation steers the optical spot. The supporting chip is stripped away in this view revealing the location of the mirror plate.

The use of the Winchester hard drive architecture allows for the presence of many such read/write heads in a single disk drive as well as the ability to store data on both sides of each disk. An optical switch [7] is used to select the active head by switching the beam from a single laser source to the optical fiber routed to the appropriate head.

TORSIONAL MIRROR DESIGN

Optical Properties

For optical efficiency, the reflecting surface must be smooth, flat, and highly reflective at the laser wavelength of 650 nm. To achieve optical smoothness, the polysilicon mirror plate is polished as described in the following fabrication sequence. A 100 nm thick Au layer affixed to the polysilicon mirror plate with a 5 nm thick layer of Cr is used as the reflector; the Au layer has a reflectivity of 95% at the 650 nm laser wavelength. In addition, the design of the OAW™ read/write head results in an optical spot 100 μm in diameter projected onto the mirror plate. The mirror plate is oversized a bit to allow for reasonable alignment tolerances and thus the plate width is 170 μm and its length in 220 μm.

Post-release flatness (radius of curvature greater than 10 cm) of the mirror plate is achieved by a long high-temperature anneal (5 hrs. at 1050 C). Stiffening ribs are fabricated integral to the mirror plate to prevent bimetallic distortion with temperature. Using the rib pattern as described in Figure 3 (perimeter of the mirror plate along with cross-bracing) wherein the ribs are 4 μm wide and 6 μm deep, center to edge distortion of the Cr-Au mirror plate is only 28 nm for a temperature excursion of 50 C.
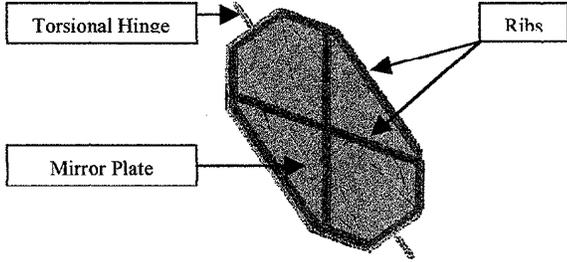
Figure 3. *Solid model of backside of torsional mirror plate showing presence of stiffening ribs.*

Electro-Mechanical Properties

The torsional mirror is driven using electrodes of polysilicon placed beneath the mirror plate. The two electrodes are on opposite sides of the rotational axis of the mirror plate. One of the electrodes is driven while the mirror plate and other electrode are grounded. The mirror plate rotates towards the driven electrode by a deflection angle θ, which is established by the equilibrium between the electrostatic driving torque, T_e, and the spring restoring torque, T_s, provided by the torsional hinges. These torques are described in the following equations:

$$T_s = k\theta \tag{1}$$

$$T_e = \frac{V^2}{2} \frac{\partial C}{\partial \theta} \tag{2}$$

where k is the torsional spring constant supplied by the two hinges, V is the voltage applied to the driven electrode, and C is the capacitance between the grounded mirror plate and the driven electrode. For the simple case of a rectangular mirror plate having rectangular drive electrodes of length a and width $b/2$ in static deflection, the capacitance of the mirror is given by:

$$C = \int_0^{\frac{b}{2}} \frac{\varepsilon_0 a \, dx}{g - \theta x} = -\frac{\varepsilon_0 a}{\theta} \ln\left(1 - \frac{b\theta}{2g}\right) \tag{3}$$

where ε_0 is the permittivity of free space and g is the air gap width of the undeflected mirror plate. Assuming $T_e = T_s$ and that there is no vertical translation of the mirror plate, Eqns. (1-3) can be solved analytically to determine the relationship between the actuation voltage and the mirror's deflection angle. Unfortunately this simple analytical approach does not correctly account for the non-rectangular shape of the mirror's plate and drive electrodes (see Fig. 3) nor for the presence of the stiffening ribs that renders the parallel plate capacitance analogy inappropriate. An alternative approach uses a finite element analysis (FEA) program (Pro/Mechanica™) to calculate the capacitance of the mirror plate as a function of angular rotation through the heat transfer analog for capacitance. For this situation it was found that the capacitance of the ribbed mirror could be empirically fit to the expression:

$$C = -A\frac{\varepsilon_0 a}{\theta} \ln\left(1 - B\frac{b\theta}{2g}\right) \tag{4}$$

where A and B are fitting parameters than depend upon the precise geometry of the mirror plate. For the mirror plate of Figure 3, $A=0.85$ and $B=1.31$. With this revised expression for capacitance, the relationship between actuation voltage and angular deflection becomes:

$$V = \sqrt{\frac{2k\theta^3}{\varepsilon_0 aA\left(\ln\left(1 - B\frac{b\theta}{2g}\right) + \frac{B\frac{b\theta}{2g}}{1 - B\frac{b\theta}{2g}}\right)}} \tag{5}$$

Electrostatic pull-in occurs for the torsional mirror when the nonlinear electrostatic torque overwhelms the available restoring torque supplied by the hinges thus limiting the angular deflection range. This condition can be derived from the equilibrium state wherein the angular derivatives of the electrostatic and restoring torques are equal or, equivalently, the angular derivative of Eqn. (5) vanishes. The calculation leads to the pull-in condition:

$$\alpha_{pi} = B\frac{b\theta_{pi}}{2g} = 0.44 \tag{6}$$

$$V_{pi} = 3.12\sqrt{\frac{k}{\varepsilon_0 aA}\alpha_{pi}^3\left(\frac{2g}{Bb}\right)^3} \tag{7}$$

where θ_{pi} and V_{pi} are the pull-in angle and voltage respectively. Eqns. (5-7) show that the presence of the stiffening ribs alters the available angular range and the necessary actuation voltage compared to the case of a flat rectangular mirror plate ($A=B=1$).

The system response of the mirror to an applied electrostatic torque, T_e, can be described by a classical mass-spring-damper model:

$$I\frac{d^2\theta}{dt^2} + t_d\frac{d\theta}{dt} + k\theta = T_e \tag{8}$$

where I is the mass moment of inertia of the mirror plate and t_d represents the squeeze-film damping torque provided by the air gap. This equation can be rewritten as:

$$\frac{d^2\theta}{dt^2} + \frac{\omega_0}{Q}\frac{d\theta}{dt} + \omega_0^2\theta = \frac{T_e}{I} \tag{9}$$

where Q represents the quality factor of this second order system and ω_0 is the natural frequency of the system as defined by:

$$\omega_0 = 2\pi f_0 = \sqrt{\frac{k}{I}} \qquad (10)$$

Combining Eqn. (6), (7), and (10) provides an expression that describes the linear relationship between actuation voltage and bandwidth for a desired angular deflection range of the mirror:

$$V_{pi} = 19.6 f_0 \sqrt{\frac{I}{\varepsilon_0 aA}} \theta_{pi}^{3} = 5.72 f_0 \sqrt{\frac{I}{\varepsilon_0 aA}\left(\frac{2g}{Bb}\right)^3} \qquad (11)$$

The mirror must be capable of steering the optical spot by ± 4 data tracks (about ± 2.8 μm) corresponding to a mirror angular deflection of ± 2 degrees. To provide some margin for error in the analytical model, we set the pull-in angle to be 2.6 degrees. The drive electrode dimensions are undersized from that of the mirror plate to allow for grounded landing pads beneath the perimeter stiffening ribs; thus a=200 μm and b=140 μm. Applying Eqn. (6) then suggests that the initial air gap must be at least 9.5 μm; a target gap of 10 μm was chosen for this device.

Servo system requirements dictate that the mirror have a resonance in excess of 25 kHz. Again using Pro/Mechanica™ FEA the resonant frequency can be calculated for the solid model given a polysilicon modulus of 160 GPa and a Poisson's ratio of 0.3. For the ribbed mirror plate geometry described previously and for torsional hinges 3.5 μm wide by 25 μm long and 2 μm thick, FEA predicts a torsional resonance of 38 kHz and a mass moment of inertia of 5×10^{-13} kg-mm^3. Thus a pull-in voltage on the order of 142 Volts is expected.

MIRROR FABRICATION

Fabrication of the micromachined mirror is described in the cross-sectional diagrams of Figure 4. To provide the high breakdown voltage required of the device, the silicon substrate is oxidized (3000 Å) and then a layer of LPCVD silicon nitride (2000 Å) is deposited on the oxide. ElectrodePoly (2500 Å) is then deposited and patterned to create interconnects, drive electrodes, and landing pads. A layer of PSG of a thickness equal to the ElectrodePoly layer thickness is then deposited and etched using a mask of opposite polarity to the ElectrodePoly mask. This reduces the topography of subsequent layers and consequently that of the mirror plate.

A thick (9-10 μm) spacer layer of PSG is deposited then trenches are etched into it. Two trench etch steps are done, the first to rib depth (6 μm) and then another to continue the trenches at the bond pads to the ElectrodePoly layer. All trenches are then filled with a 2.5 μm thick conformal layer of TrenchPoly, which leaves severe cusps at their tops. To eliminate these cusps a two-step chemical mechanical polish (CMP) is done using SDE3000 polysilicon slurry by Rodel. First the TrenchPoly is polished back to the top of the spacer PSG layer. Then another 2.5 μm thick polysilicon layer (MirrorPoly) is deposited, annealed, and polished back to 2.0 μm thick. The resulting surface of the MirrorPoly layer is smooth; RMS surface roughness is less than 3 nm.

Next a 7000 Å thick layer of Au with a 100 Å thick Cr adhesion layer is deposited and patterned to provide wire bond pads. A second Cr-Au layer is then deposited and patterned to create the reflector on the mirror plate. The MirrorPoly layer is then patterned to define the bond pads, mirror plate, and torsional hinges. Finally a HF release etch of the PSG spacer is done to free

the mirror plate. A diced mirror chip ready for attaching to the read/write head is shown in Figure 5.
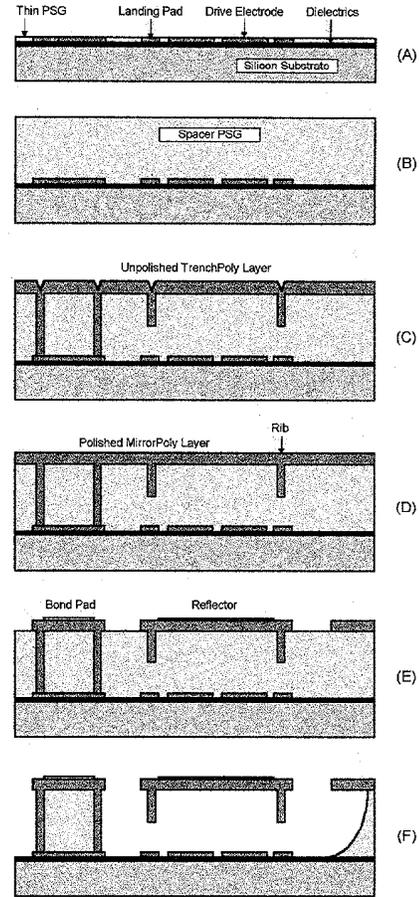
Figure 4. Fabrication sequence for mirror: (A) Thin PSG Pattern, (B) Spacer PSG Deposition, (C) TrenchPoly Deposition, (D) Polished MirrorPoly, (E) Metallizations, and (F) Post-Release Etch.
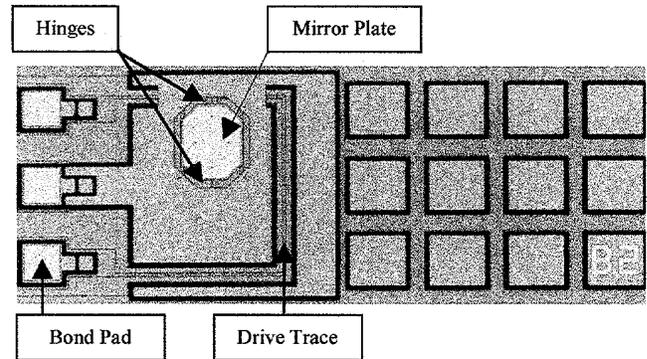
Figure 5. Silicon chip containing the micromachined torsional mirror. Square features at the right are stress relief mesas whereby the sacrificial PSG layer is removed between them.

MIRROR TESTING

The micromachined mirrors were tested using an optical approach similar to the one described in [8]. A collimated beam from a laser source is directed through a beam splitter and then

focussed onto the reflecting surface of the mirror plate. The reflected beam is collimated upon returning through the focussing lens and is then redirected by 90 degrees by the beam splitter onto a Hamamatsu S2044 position sensitive detector (PSD). Upon angular rotation of the mirror plate, the reflected beam is displaced with respect to the incoming beam by an amount proportional to the focal length of the lens and the tangent of twice the angular deflection. The PSD's raw output signal is proportional to the laser spot's position on its active area. This output is conditioned using a Hamamatsu C4674 amplifier thus yielding a signal proportional to the angular displacement of the mirror plate. A fast digital oscilloscope (Tektronix 420A) is used to display and record the output of the amplifier.

Testing of the mirror involves procedures for measuring its angular displacement as a function of actuation voltage and for determining the mirror's frequency response. For both tests the mirror is driven single-sided with the undriven electrode and the mirror plate grounded. A high-voltage high-bandwidth amplifier whose input comes from a HP 33120A arbitrary waveform synthesizer supplies the mirror drive signal. For the angular deflection test, the drive signal to the mirror is a 500 Hz voltage ramping linearly from zero to 150 Volts. The amplified PSD output is then directly proportional to the deflection angle. For the frequency response test, the drive signal is a fast 100 Volt step. A fast Fourier transform (FFT) of the differentiated PSD output (step response) then yields the frequency response.

Figure 6 shows the performance of a typical torsional mirror showing an angular range of about 2 degrees before pull-in. Pull-in voltage is about 145 volts. The mirror's frequency response is shown in Figure 7. Resonance occurs at about 36 kHz with a Q of about 5.
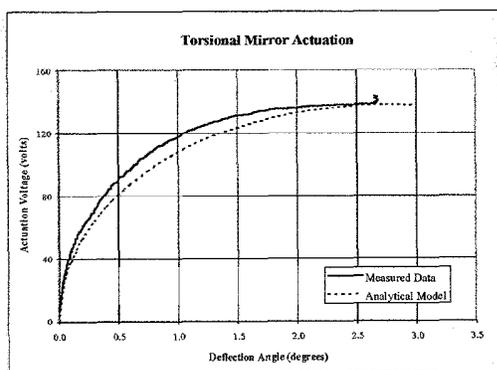
Figure 6. *Actuation characteristics for a typical mirror compared to analytical model presented above.*
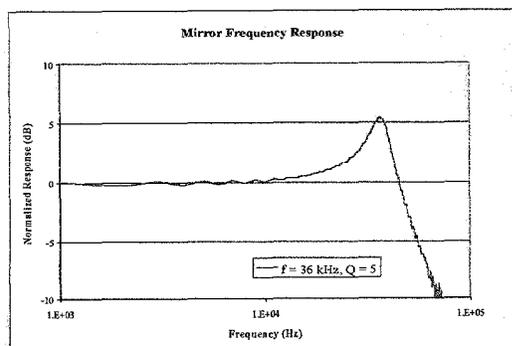
Figure 6. *Torsional mirror frequency response.*

DISK DRIVE PERFORMANCE

Operational read/write heads have been assembled and built into functioning prototypes of the OAW™ 5.25" disk drive. The high bandwidth and relatively low Q of the mirror allow for the successful implementation of the dual-stage servo system described above with a closed-loop bandwidth of 3-5 kHz. System simulations, however, suggest a maximum achievable servo bandwidth of 10 kHz. Although originally implemented for a data track density of 35 ktpi, the track misregistration error (TMR) for the system has been measured at 30 nm suggesting suitability for track densities as high as 100 ktpi.

Fine positioning by the mirror is also used to provide very fast short seeks. Prototype drives have demonstrated fully settled single-track seeks of 200 μsec. The performance for longer seeks are a function of the VCM coarse actuator; one-third stroke times are on the order of 13.5 msec. Given the small mass of the mirrors, the read/write head is quite robust to such accelerations.

CONCLUSIONS

Conventional surface micromachining techniques have been enhanced to create a polysilicon torsional mirror for second stage actuation in magneto-optical data storage. The design and fabrication process of the mirror have been optimized for optical efficiency and for electro-mechanical properties that allow for a high bandwidth servo loop within a Winchester-style disk drive. Prototypes of the disk drive have shown to be capable of supporting track densities on the order of 100 ktpi.

ACKNOWLEDGEMENTS

The authors would like to thank Kathy Jackson, Kevin Yasumura, Mark Walters, and Cronos Integrated Microsystems for help in fabrication and testing. Al Fennema was helpful in supplying data on the disk drive system-level performance.

REFERENCES

1. I. Naniwa, S. Nakamura, S. Saegusa, K. Sato, "Low Voltage Driven Piggy-Back Actuator of Hard Disk Drives", *Proceedings of the 12th Int. Conf. on MEMS (MEMS '99)*, Orlando, FL, pp. 49-52, 1999.
2. S. Kondo, S. Yoshimura, K. Yaguchi, Y. Miyata, J. Numazawa, M. Suzuki, M. Esashi, "Monolithic Piezo-Electric Actuator for Head Positioning", *Transducers '99*, Sendai, Japan, pp. 1838-1839, 1999.
3. T. Imamura, T. Koshikawa, M. Katayama, "Transverse Mode Electrostatic Microactuator for MEMS-based HDD Slider", *Proceedings of the 9th Int. Conf. on MEMS (MEMS '96)*, San Diego, CA, pp. 216-221, 1996.
4. D. Horsley, N. Wongkomet, R. Horowitz, A. Pisano, "Design and Feedback Control of Electrostatic Actuators for Magnetic Disk Drives", *1998 Solid-State Sensor and Actuator Workshop*, Hilton Head, SC, pp. 120-123, 1998.
5. M. Sekimura, M. Yonezawa, K. Uchimaru, A. Kasahara, N. Uchida, "Electrostatic Torsion Mirror", *Tech. Dig. of 16th Sensor Symposium*, pp. 167-170, 1998.
6. US Patent Nos. 5,850,375, 5,940,549, and 5,999,303.
7. J. Grade and H. Jerman, "A Large-Deflection Electrostatic Actuator for Optical Switching Applications", *2000 Solid State Sensor and Actuator Workshop*, Hilton Head, SC, 2000.
8. K. Honer, N. Maluf, E. Martinez, G. Kovacs, "A High-Resolution Laser-Based Deflection Measurement System for Characterizing Aluminum Electrostatic Actuators", *Transducers '95*, Stockholm, Sweden, pp. 308-311, 1995.

A MIXED-SIGNAL SENSOR INTERFACE MICROINSTRUMENT

Keith L. Kraver[a], Matthew R. Guthaus[a], Timothy D. Strong[a],
Peter L. Bird[a], Geun Sig Cha[b], Wolfgang Höld[c], Richard B. Brown[a]

[a]Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, MI 48109
[b]Chemical Sensor Research Group, Department of Chemistry, Kwangwoon University, Seoul, Korea
[c]Personal Systems Division, National Semiconductor Corporation, D-82256 Fürstenfeldbruck, Germany

ABSTRACT

A single-chip implementation of a microinstrumentation system is presented. The chip incorporates voltage, current, and capacitive sensor interfaces; a temperature sensor; a 10-channel, 12-bit analog-to-digital converter; and an 8-bit microcontroller with a 16-bit hardware multiplier and a 40-bit accumulator. Standard serial and parallel interfaces facilitate digital communication with a host system. Fabricated in a standard 0.35 μm digital CMOS process, the die occupies 3.8 mm x 4.1 mm, operates from a nominal supply voltage of 3 V, and draws 16 mA when fully powered (850 μA standby current). For testing purposes the chip is packaged in a 132-pin ceramic PGA.

INTRODUCTION

Sensor interfaces have been identified as critical to the development of the MEMS field [1]. Multi-chip sensor/interface-circuit solutions have been developed [2], and interface electronics have been integrated with sensors [3]. This paper presents a single-chip microinstrument: a mixed-signal microcontroller with a full analog front end--including voltage, current, and capacitive sensor interfaces; signal amplification; and analog to digital conversion circuitry. The only sensor integrated on this chip is for temperature, but other sensors and actuators could be included. For many applications, though, in which the MEMS devices cannot be made with CMOS process steps or in which sensor lifetime is limited, the most efficient system partitioning allocates all or most of the electronics on a microinstrument chip separate from the sensor.

The mixed-signal microinstrument, named the MS-8, is optimized, from instruction set definition to analog functionality, to economically support embedded sensor applications. See the block diagram in Figure 1 [4]. A minimal version of the microinstrument can be implemented with as few as eight pins. The instruction set provides excellent code density and supports the ANSI C programming language. To aid code debugging, the MS-8 includes hardware support for a single breakpoint and trace event, and a development system interface provides instructions for halting the processor, single stepping through code, and reading and writing system registers. Single cycle interrupt response and direct-memory access (DMA) capabilities facilitate usage of the MS-8 in time-critical applications, and a 16 x 16 hardware multiply, 40-bit accumulate block provides significant signal processing capability for a small microcontroller. To minimize power dissipation, data paths were limited to 8-bit widths, and the chip was designed to operate on a 3-V supply, in a 0.35-μm CMOS process. In addition, a clock manager provides programmability of the clock frequency for the digital and analog circuits, and each analog block can be individually powered down under software control.

DIGITAL CORE AND PERIPHERALS

The right side of Figure 1 depicts the digital portion of the MS-8, which includes the processor core, peripherals, and mem-
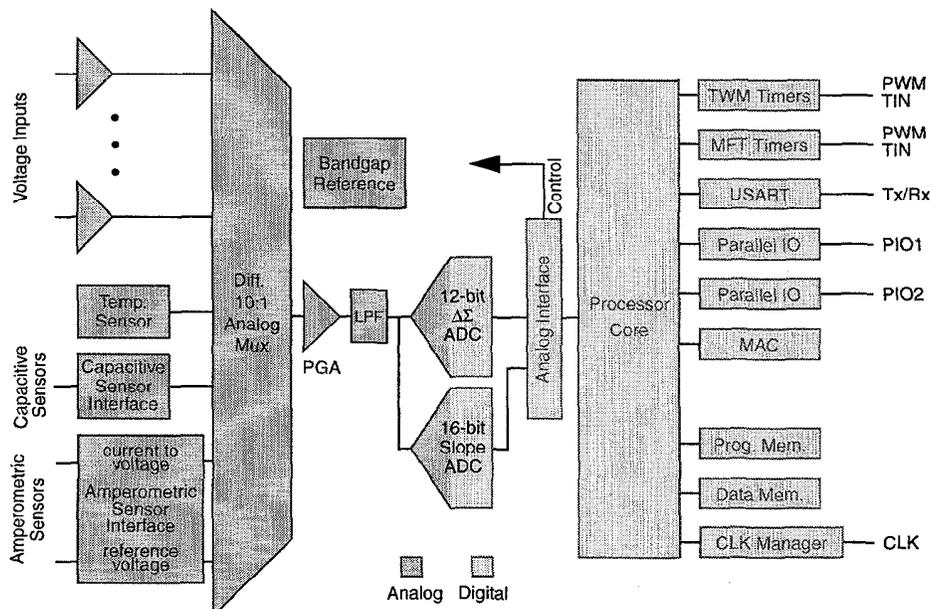
Figure 1. MS-8 block diagram. The analog (dark shade) and digital (light shade) sections interface through the analog interface block, treated as a peripheral by the processor core.

ory [5]. Communication to a host system is facilitated through two parallel input/output units (PIO) and the universal synchronous/asynchronous receiver/transmitter (USART). Standard timing functions are provided by the watchdog timer (TWM) and the multifunction timer (MFT). The multiply-accumulate unit (MAC) supports 16-bit multiplication and 40-bit accumulation, permitting on-chip signal processing. To reduce power, the clock manager can divide the system clock, select a slow clock, or selectively halt the clock to the different sections of the chip until a programmed event occurs. Program memory in this prototype version of the MS-8 consists of a 512 B boot ROM and a 4 KB RAM; a 512 B RAM provides data storage. For application versatility, both RAMs are configurable as either instruction or data storage. In addition, off-chip memories can occupy part of the 20-bit address space. Finally, the analog interface block (AIB) contains control registers for configuring the analog circuits and data registers for buffering samples from the analog to digital converters.

ANALOG CORE

The analog front end (AFE), outlined in the left half of Figure 1, includes a multiplexor which selects from an array of buffered voltage inputs (for potentiometric chemical sensors), the temperature sensor, a capacitive sensor interface (for pressure and acceleration sensors), and an amperometric chemical sensor interface. Signal conditioning provided by the programmable gain amplifier makes the interface generic and increases the dynamic range of the system. A passive low-pass filter and a 12-bit analog to digital converter (ADC) complete the AFE. The remainder of this section summarizes the analog functionality.

The buffered voltage inputs provide a high impedance interface for potentiometric sensors or any other voltage signals in the range from 0 to 3.0 V (single-ended). Potentiometric chemical sensors have output impedances on the order of 100 MΩ, which require readout circuitry with high input impedance. The CMOS rail-to-rail opamp, described later, configured as a voltage follower, provides the necessary high input impedance buffer. Buffers not in use can be powered down selectively.

The temperature sensor, derived from the bandgap reference, is designed to provide a linear conversion of temperature to voltage in a range exceeding -40 to 85 °C according to

$$V(T) = -B\frac{R_2}{R_1}\ln(A)\frac{KT}{q} + V_{ref}\left(1 + \frac{R_2}{R_{ref}}\right),$$

where B and A are temperature-independent constants, the Rs are carefully ratioed resistors, and V_{ref} is a temperature-independent voltage generated by the on-chip bandgap reference. Temperature range and sensitivity are competing characteristics set by the gain of the programmable gain amplifier, discussed later.

The capacitive interface circuit converts a capacitance to a voltage. Capacitive pressure sensors and accelerometers operate by deforming a capacitor plate from the steady state position in response to an applied force [2]. In the readout circuit shown in Figure 2, a switched-capacitor charge integrator compares the resulting sensor capacitance to a reference capacitor according to

$$V_{out} = V_A\frac{(C_s - C_{ref})}{C_f} + V_{ref},$$

where C_s and C_{ref} are the sensor and reference capacitors, respectively; V_A is the amplitude of the clock, ϕ; and V_{ref} is the on-chip reference voltage. Through program control, the user can select an
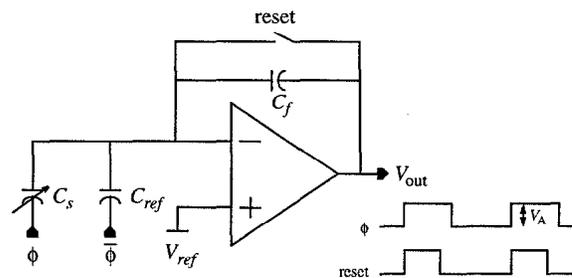
Figure 2. *Simplified circuit diagram of the capacitive readout circuit.*

on-chip programmable capacitor bank or an off-chip reference capacitor.

Amperometric sensors are frequently employed to detect dissolved gases or uncharged molecules in liquid. As shown in Figure 3, these sensors typically include three terminals: counter, reference, and working. The opamps connected to the counter and reference terminals create a potentiostat circuit whereby feedback regulates the applied voltage across the counter and working terminals independent of the current in the cell. In this version of the MS-8, a simple ramp generator, shown functionally in Figure 3, generates the applied voltage. Current sensing occurs at the working electrode with a simple current follower amplifier which outputs a voltage proportional to the cell output current. To facilitate varying sensors and current magnitude, R_f is programmable through the AIB or supplied off-chip.

The 10:1 differential analog mux consists of two 10:1 single-ended muxes. Each of the previously mentioned mux inputs feeds a channel of both single-ended muxes. With this connection scheme, the mux provides maximum versatility in selecting inputs. The AIB contains a register and logic for configuring the mux under program control.

In this implementation of the MS-8, the programmable gain amplifier (PGA) is designed after the standard three-opamp, two-stage instrumentation amplifier. (See Figure 4.) The first stage provides large input impedance, a differential input, and programmable gains of 1 to 61 in steps of 10 V/V. Stage two of the amplifier provides low output impedance, differential to signal-ended conversion, and level shifting. The level shifting maintains the DC level of the signal at mid-rail to avoid clipping. Use of the rail-to-rail opamp (described later) in both stages provides the PGA with a wide dynamic range. At the output of the PGA, a simple low-pass filter reduces amplifier distortion and aliasing (resulting from the sampling process). An AIB control register stores the gain settings.
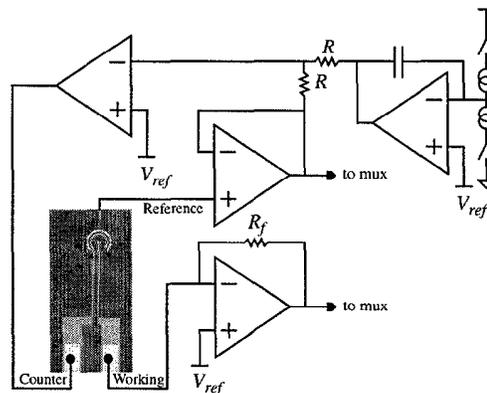
Figure 3. *Simplified diagram of the amperometric sensor interface circuit and off-chip sensor.*
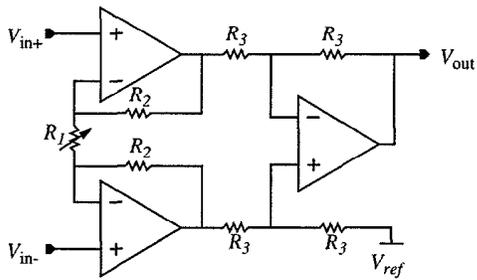
Figure 4. *Programmable instrumentation amplifier. Gain is set by an on-chip programmable resistor (R_1).*

The bandgap reference supplies bias current and a voltage reference for the analog circuitry. An opamp in a negative-feedback configuration provides voltage gain to boost the reference voltage to mid-rail and buffering to drive the reference to the $\Delta\Sigma$ ADC, programmable gain amplifier, temperature sensor, potentiostat, and off-chip. To avoid start-up transient behaviors, the bandgap circuit is powered even in low-power mode.

Two selectable data converters produce a digital representation of the analog signal. The single-slope ADC has higher resolution than the $\Delta\Sigma$ ADC but is more susceptible to substrate noise. Because of this limitation, the single-slope converter will be useful for studying noise-reduction methods in this mixed-signal environment. The ramp voltage for the converter is generated by supplying a fixed current, derived from the bandgap reference, to an off-chip capacitor.

The first-order $\Delta\Sigma$ ADC (the primary data converter) is shown in Figure 5. Under program control, the clock manager divides the system clock to generate the sampling clock and permits the user to vary the sampling frequency according to the application. In addition, the clock manager generates the four clock phases shown in the figure. This clocking scheme reduces the impact of switch-induced charge injection on the high impedance nodes.

A 16-bit counter in the AIB filters and decimates the oversampled, single-bit ADC output. The AIB stores both the counter output and the single-bit oversampled output, allowing the processor core to provide the filtering and decimation, if desired by the user.

A custom opamp, based on a topology in [6], was designed for use in the analog circuits described previously. Since in some cases the opamp drives the signal off-chip (e.g. the reference voltage and test points), the opamp is designed to handle 200 pF loads. To increase dynamic range in this low-voltage environment, the opamp processes rail-to-rail input signals and drives rail-to-rail output signals. A circuit that maintains constant transconductance, g_m, over the input common mode ranges reduces distortion.
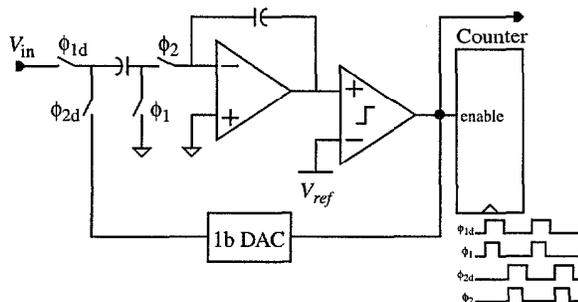
Figure 5. *Simplified diagram of first-order $\Delta\Sigma$ modulator. Filtering is provided by the counter or with software and the MAC peripheral.*
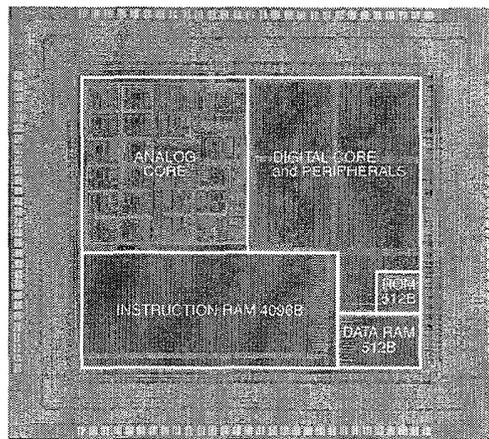
Figure 6. *MS-8 die photomicrograph.*

Finally, the transistors in the input differential pair and output stage operate in weak inversion to reduce input referred noise and offset, and to maximize g_m for a given current.

PHYSICAL DESIGN

Figure 6 is a die photomicrograph of the MS-8. The major components visible are the RAMs, ROM, analog circuitry, and the microcontroller core. The die is approximately 3.8 mm x 4.1 mm. Nearly one-quarter of the core area, 1.39 mm x 1.42 mm, is dedicated to the analog circuits. RAM and ROM memory cells occupy a little less than half of the core.

The memory drivers are placed as far from the analog circuits as possible to reduce substrate coupling effects. In addition, the analog core has ample substrate contacts surrounding the sensitive circuitry and dedicated V_{dd} and ground connections to further combat switching noise. Finally, enforcing metal-to-metal spacings greater than minimum reduces crosstalk between signal lines (at the expense of area).

EXPERIMENTAL RESULTS

The MS-8 digital core has been verified to be functional at 40 MHz, 3.0 V, and 75 °F using an HP82000 digital tester. While simulations predict static current draw around 3.3 mA (full power) and 15 μA (standby), measurements reveal 16 mA and 850 μA. Initial testing of the AFE reveals the usefulness of the device in embedded sensor applications. The USART enabled bidirectional serial communications between the MS-8 and a laptop computer running LabView (National Instruments). Specifically, programs were downloaded to the chip and digital samples from the ADC were transferred to LabView for recording and analysis. The potentiometric and capacitive sensor interfaces and the temperature sensor have been verified. In addition, the ramp generator circuit in the amperometric interface has been successfully tested, and amperometric sensors are being prepared to complete the demonstration of the MS-8 analog interfaces.

Figure 7 shows the calibration curve over seven decades of KCL concentration for a solid-state potentiometric sensor selective to potassium ions (K^+) [7]. The inset contains the time response of the sensor; each step represents an 10-fold increase in potassium concentration. A wire from the output pin of the on-chip reference generator set the potential of the solution to 1.5 V. The sensor has a linear response from around the minimum detectable limit of 10^{-5} M to about 1 M, as shown in the figure. Using the

16
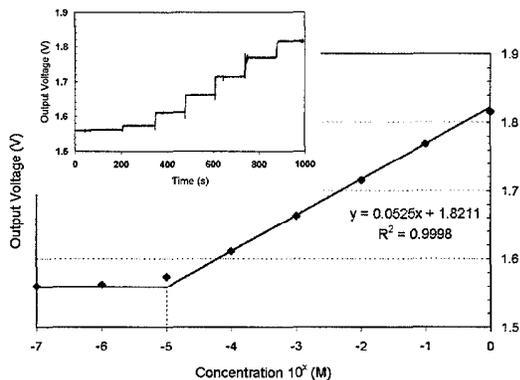
Figure 7. K^+ *calibration curve. The background solution consists of 200 ml of 0.05 M Trizma Base (Sigma, St. Louis, MO) adjusted to pH 7.4 with H_2SO_4. The PGA gain equals 1 V/V.*

microcontroller to linearize the response with a higher-order fit would extend the detection range.

The initial test of the capacitive interface circuit was performed with an 8.2 pF external reference capacitor and known capacitances to represent the sensor, see Figure 8. Actual capacitive-based pressure sensors will be interfaced to the MS-8 in the future. The figure shows a linear capacitance to voltage conversion. Pressure sensors, however, demonstrate a nonlinear relationship between pressure and output capacitance, which could be linearized with the signal processing capabilities of the MS-8.

To test the temperature sensor, the MS-8 and test-board were placed in an oven. The temperature was ramped slowly from room temperature, 22 °C, to 80 °C, as measured by a mercury thermometer. At each temperature, the chip was allowed to reach equilibrium with the oven environment before a digital voltage reading was recorded. As expected, Figure 9 shows the temperature sensor output voltage has a linear relationship to temperature.

CONCLUSIONS

The MS-8 described here is the first implementation of an on-going project that coordinates efforts among analog designers, compiler writers, and low-power digital designers to develop a low-power, low-cost, and multipurpose sensor interface and data acquisition system. This single-chip microinstrument contains a programmable analog front end capable of interfacing to a variety of sensors. The integrated microcontroller and peripherals support digital filtering and compensation of sensor outputs, timing control
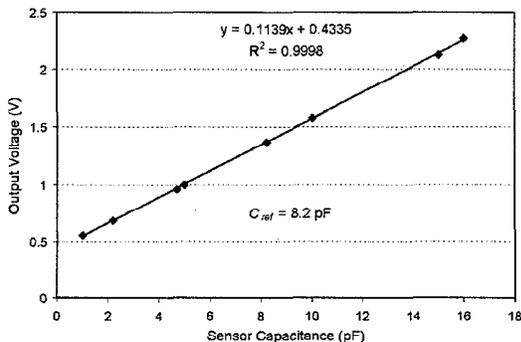
Figure 8. *Capacitive sensor readout circuit test. The PGA gain is set to 1 V/V.*
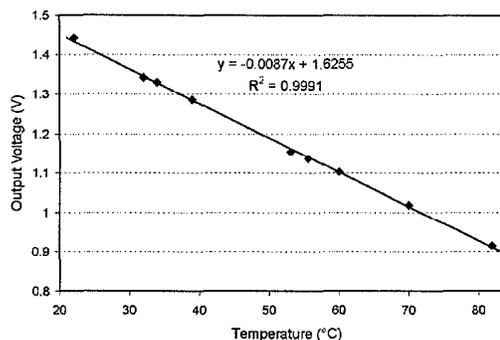
Figure 9. *Temperature sensor response from 25 to 80 °C with PGA gain set to 11 V/V.*

for sampling multiple sensors, and communication with a host system. The usefulness of the MS-8 in embedded sensor applications is evidenced by the initial results presented. More testing with pressure and chemical sensors is planned to further demonstrate the capabilities of the MS-8.

ACKNOWLEDGMENTS

REFERENCES

1.   K. D. Wise, Ed., *Proc. IEEE Special Issue on Integrated Sensors, Microactuators, and Microsystems (MEMS)*, August 1998.

2.   A. V. Chavan, A. Mason, U. Kang, and K. D. Wise, "Programmable Mixed-Voltage Sensor Readout Circuit and Bus Interface with Built-In Self-Test," *IEEE ISSCC Digest of Technical Papers*, pp. 136-137, February 1999.

3.   H. D. Goldberg, *The Batch Fabrication of Integrated Chemical Sensor Arrays*, Ph.D. Dissertation, The University of Michigan, 1993.

4.   P. L. Bird, R. B. Brown, T. N. Mudge, L. Hong, and B. Khailany, "MS-8 Design Specification," Technical Report, The University of Michigan, January 1999.

5.   R. B. Brown, P. L. Bird, M. R. Guthaus, K. L. Kraver, and B. Khailany, "MS-8 Microarchitecture," Technical Report, The University of Michigan, May 1999.

6.   R. Hogervorst and J. H. Huijsing, *Design of Low-Voltage, Low-Power Operational Amplifier Cells*. Klewer, Boston, 1996.

7.   H. J. Yoon, J. H. Shin, H. Nam, G. S. Cha, T. D. Strong, and R. B. Brown, "Solid-State Ion Sensors with a Junction-Free Solvent-Processible Polymer Membrane-Based Reference Electrode for Blood Analysis," *10th International Conference on Solid-State Sensors and Actuators, Transducers '99*, Sendai, Japan, pp. 652-655, June 7-10, 1999.

POST-CMOS MODULAR INTEGRATION OF POLY-SIGE MICROSTRUCTURES USING POLY-GE SACRIFICIAL LAYERS

A. E. Franke,[#o] Y. Jiao,[#o] M. T. Wu,[*o] T.-J. King,[#] R. T. Howe[#*o]

[#]Department of Electrical Engineering and Computer Sciences
[*]Department of Mechanical Engineering, [o]Berkeley Sensor & Actuator Center
University of California at Berkeley, Berkeley, CA 94720-1770

ABSTRACT

Polycrystalline silicon-germanium (poly-SiGe) micro-structures have been fabricated on standard CMOS wafers with aluminum metallization. Hydrogen peroxide (H_2O_2) etches the sacrificial polycrystalline germanium (poly-Ge) layer without significantly etching the p-type poly-SiGe structural layer. In contrast to HF sacrificial release etches, no special protection of the underlying CMOS layers is needed. For the first time, the fabrication of LPCVD surface microstructures directly on top of standard electronics is demonstrated, which leads to dramatic reductions in both MEMS-CMOS interconnect parasitics and device area.

Initial measurements indicate that poly-SiGe has promising material properties. Its fracture strength is 1.2 ± 0.1%, which is comparable to that of poly-Si. A folded flexure lateral resonator has a quality factor in vacuum as high as ~15,000. No stress or dopant-activation anneal of the structural layer is needed, since the *in-situ* boron-doped poly-SiGe is found to have an as-deposited stress of only -10 MPa (compressive) and a resistivity of only 1.8 mΩ-cm. From a cantilever-beam array, the strain gradient is found to be rather high (+1.4 x 10^{-4} (μm)$^{-1}$), so that a 120 μm cantilever bends away from the substrate by 1 μm at its tip. Deposition and annealing conditions have yet to be optimized to minimize strain gradient.

INTRODUCTION

The modular integration of microelectromechanical devices with driving, controlling, and signal-processing electronics is advantageous for improving performance, lowering manufacturing cost, and allowing for separate development of the micromechanics and microelectronics technology components. The microelectronics-first approach has the advantage that it is more conducive to the use of integrated-circuit (IC) foundries for low-cost electronics fabrication. However, it necessitates a low thermal budget for the microelectromechanical structures (MEMS) fabrication process, if conventional circuits employing aluminum or copper interconnects are to be used. A standard polycrystalline-silicon (poly-Si) based MEMS process typically employs deposition temperatures near 600°C and annealing temperatures at or above 900°C in order to achieve low-resistivity, low-tensile-stress structural films [1]. These temperatures are too high to be compatible with conventional IC metallization schemes. Poly-Ge has been demonstrated as a CMOS-compatible structural material, in a process which uses a conventional SiO_2 sacrificial layer [2, 3]. However, integration with CMOS requires a low-temperature, pinhole-free, HF-resistant film to protect the electronics during the lengthy HF release etch.

Low-pressure chemical vapor deposited (LPCVD) poly-SiGe has properties similar to poly-Si, but can be deposited and annealed to activate dopants at much lower temperatures than poly-Si [4, 5]. The high dopant activation found in p-type poly-SiGe [5] eliminates the need for post-deposition annealing in order to lower the resistivity. $Si_{0.7}Ge_{0.3}$ has been used recently as the structural material for surface-micromachined bolometers [6].

This paper describes a micromachining process based on p-type poly-SiGe structural films and poly-Ge sacrificial films, which enables the modular integration of conventional surface microstructures with standard CMOS circuitry. In the initial experiments, no anneal of the p-type poly-SiGe structural films was needed to reduce either resistivity or stress. Special protection of the CMOS is unnecessary, since the H_2O_2 release etchant does not damage the either the deposited SiO_2 passivation layers [7] or the exposed Al bond pads.

PROPERTIES OF P-TYPE POLY-SIGE

The p-type poly-SiGe structural film was deposited in a Tystar horizontal LPCVD furnace at 450°C and 600 mT, using 85 sccm SiH_4, 90 sccm GeH_4, and 50 sccm of 10% B_2H_6/ 90% SiH_4. Rutherford back scattering (RBS) indicated that the films were $Si_{0.36}Ge_{0.64}$. This composition is advantageous because of its combination of relatively low deposition temperature, high deposition rate of 168 Å/min, and as-deposited polycrystalline structure. Moreover, its Ge content is less than the threshold of 70%, above which significant attack by the sacrificial H_2O_2 will occur [7]. An as-deposited 3 μm-thick film had a sheet resistance of 6 Ω/square, or a resistivity of 1.8 mΩ-cm. Annealing did not lower the resistivity significantly, indicating that the boron was fully activated. This result is expected for Ge-rich p-type poly-SiGe films [5].

The residual stress of the p-type poly-$Si_{0.36}Ge_{0.64}$ was measured with the wafer curvature technique using a Flexus laser interferometer. The stress was found to be -10 MPa (compressive), which is quite low for an as-deposited film.

The stress gradient was determined by analyzing released cantilever beams (Fig. 1). The substrate is tilted 60 degrees, so the upper array of cantilevers appear to bend away from the substrate more than does the lower array. The deflection of 1 mm long beams is 70 μm, which corresponds to a strain gradient of 1.4 x 10^{-4} (μm)$^{-1}$. The deposition conditions have not been optimized to reduce strain gradient.
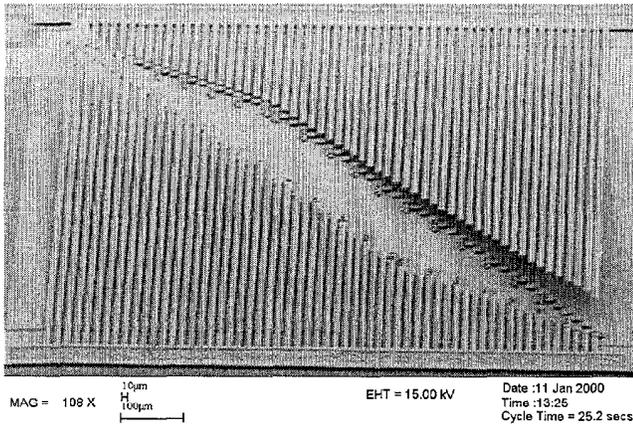
Figure 1. *Array of as-deposited poly-Si$_{0.36}$Ge$_{0.64}$ cantilever beams. The substrate is tilted by 60 degrees, which makes the strain gradient more noticeable in the upper array.*

The fracture strain of poly-Si$_{0.36}$Ge$_{0.64}$ was characterized by measuring the strain state in cantilever beams as they fractured due to in-plane bending [8]. The test structure used for these measurements is shown in Fig. 2. The fracture strain of 50 μm-long poly-Si$_{0.36}$Ge$_{0.64}$ beams was 1.2 ± 0.1% (28 beams tested), whereas the fracture strain of a MCNC poly-Si was 1.9 ± 0.2% (28 beams tested). These results indicate that the strength of poly-Si$_{0.36}$Ge$_{0.64}$ is comparable to poly-Si.
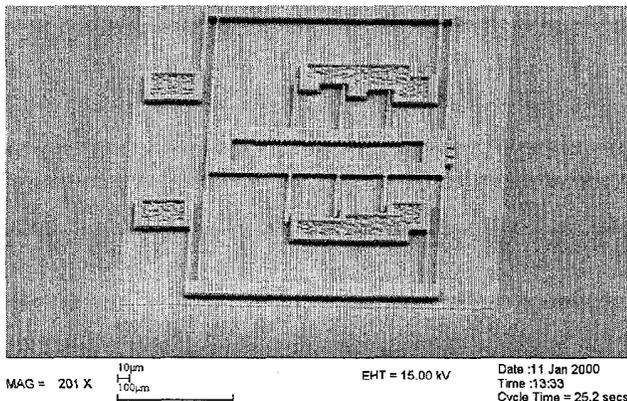
Figure 2. *Poly-Si$_{0.36}$Ge$_{0.64}$ structure to measure fracture strain.*

Poly-Si$_{0.36}$Ge$_{0.64}$ folded-flexture resonators were driven in vacuum to determine their resonant frequency and quality factor, Q. The resonators were excited with a 53 mV V$_{p-p}$ AC signal on a comb drive and 7 V DC on the shuttle and ground plane. Measurements of Q were performed in a vacuum probe station at a pressure of 290 μTorr. Current from the sense electrode was fed to an off-chip transimpedance amplifier. Quality factors as high as ~15,000 were measured at a resonant frequency of 11.783 kHz (Fig. 3). Higher quality factors may be achievable after annealing the structural film.
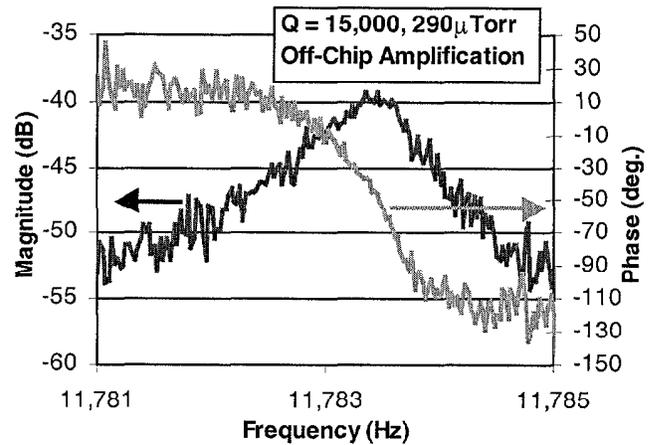
Figure 3. *Frequency (amplitude and phase) characteristics of a poly-Si$_{0.36}$Ge$_{0.64}$ resonator.*

INTEGRATED POLY-SIGE MEMS PROCESS

Transresistance amplifiers were fabricated in the UC Berkeley Microlab baseline 3 μm, p-well CMOS technology. The metallurgy for the CMOS was Al with 2% Si. It was demonstrated that this metallization can withstand 5 hours of annealing at 450°C without junction spiking problems. Electronic fabrication was completed with a 400°C sintering anneal. The CMOS circuitry was passivated with 1 μm of low temperature oxide (LTO), which was deposited at 400°C.

Folded-flexure resonators were fabricated using poly-Si$_{0.36}$Ge$_{0.64}$ as the structural material, poly-Ge as the sacrificial material and using only conventional microfabrication equipment. Contacts were made through the LTO to connect a p-type poly-Si line to the MEMS p-type ground plane poly-Si$_{0.36}$Ge$_{0.64}$. The p-type poly-Si jumper is needed since direct deposition of poly-SiGe onto aluminum is not permitted in our SiGe LPCVD furnace, which is also used for MOSFET gate research. The poly-Si jumper must be p-type to prevent diode formation. This poly-Si was doped by ion implantation along with the PMOS source and drain regions.

The poly-Si$_{0.36}$Ge$_{0.64}$ films for the ground plane and structural layers were deposited in a conventional LPCVD system at 450°C and 600 mT by pyrolysis of 85 sccm of SiH$_4$, 90 sccm of GeH$_4$ and 50 sccm of 10% B$_2$H$_6$/ 90% SiH$_4$ for 30 min. Since SiGe does not readily nucleate on SiO$_2$, a thin (<5 nm thick) seed layer of Si was deposited *in-situ* prior to the poly-SiGe. The 0.5 μm-thick poly-Si$_{0.36}$Ge$_{0.64}$ ground plane was patterned and etched in 25 s in a high-density plasma (15 mT, 300W top RF, 150W bot. RF, 5.8 cm gap, 50 sccm Cl$_2$, 150 sccm HBr). The etch was continued for an additional 35s to ensure removal of stringers.

A 200 Å-thick LTO layer was deposited on the poly-Si$_{0.36}$Ge$_{0.64}$ ground plane to act as an etch stop for the later poly-Ge anchor etch. This layer is not essential and could be eliminated, once the poly-Ge anchor etching process is optimized. The LTO etch stop reduced the likelihood of etching significantly into the poly-Si$_{0.36}$Ge$_{0.64}$ ground plane.

The poly-Ge sacrificial layer deposition consisted of two steps. A thin seeding layer of Si was deposited with 5 min, 300 mT, at 375°C, and with 200 sccm of Si_2H_6 followed by a 2 hr and 45 min deposition at 375°C, with 220 sccm of GeH_4, which resulted in a 2.8 μm thick film. The anchor windows were etched in 135 s using the same plasma chemistry used to etch the ground plane. It should be noted that standard wafer cleaning procedures are not possible whenever poly-Ge is exposed, because poly-Ge is rapidly etched in piranha and RCA, "SC1" solutions. Solvent cleans and O_2 plasma were performed instead to clean organics. Development of an alternative cleaning procedure for organic contaminants that minimizes solvent waste is desirable.

The 3 μm-thick structural poly-$Si_{0.36}Ge_{0.64}$ layer was deposited with the same conditions as the ground plane except the time was increased to 3 hours. The film was patterned with the same plasma chemistry. An endpoint signal was seen as the etching proceeded into the underlying poly-Ge after 165s, and etching proceeded for an additional 15 s to clear stringers.

Structures were released in 80°C H_2O_2 for 2 hours. Dice were rinsed in water and methanol and then dried using the CO_2 critical point drying technique. Figure 4 is a schematic cross-sectional view of the completed device. Figure 5 consists of SEM images of a resonator fabricated next to an amplifier and a resonator fabricated on top of the amplifier.
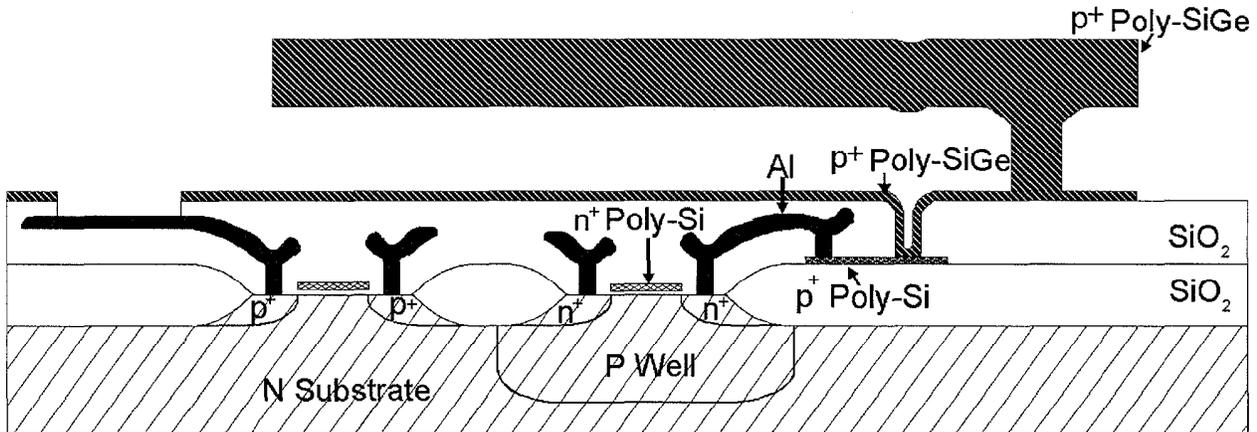
Figure 4. Schematic cross-sectional view of the fabricated CMOS and MEMS structures.
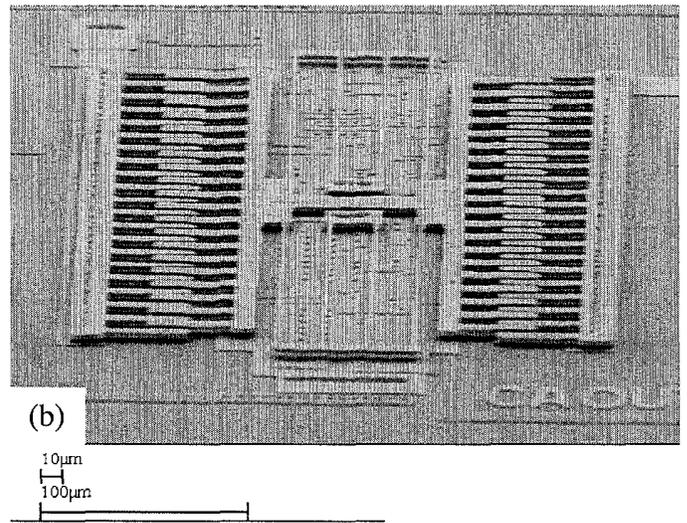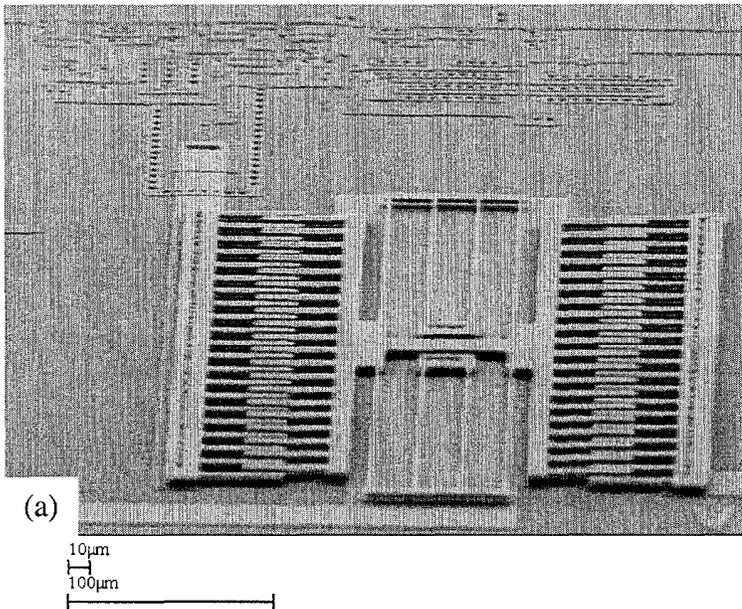
Figure 5. SEMs of fabricated poly-SiGe resonators fabricated (a) adjacent to a CMOS amplifier and (b) on top of a CMOS amplifier. Note the area reduction in fabricating MEMS structures on top of electronics.

Processing of the poly-SiGe resonators did not degrade the performance of the transistors. However, the NMOS transistors had significant leakage even before the processing of the MEMS structures. Secondary ion mass spectrometry (SIMS) of a monitor wafer indicated that the boron concentration in the "p-well" was below that of the n-type substrate dopant concentration. Therefore there was no p-well. New wafers are currently being processed.
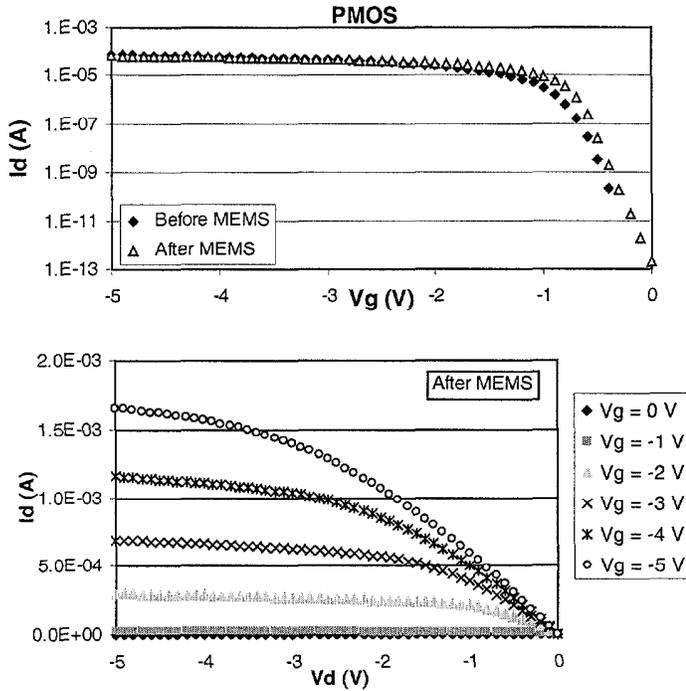
Figure 6. PMOS characteristics with W/L = 50 µm / 3 µm before and after processing of the poly- SiGe resonators.

CONCLUSIONS

A poly-SiGe structural/poly-Ge sacrificial material MEMS technology has been demonstrated which allows the integration of surface microstructures directly on top of standard electronics. No special passivation or other processing is needed to protect the CMOS layers. Stacking the MEMS on the sense amplifier drastically reduces the parasitic capacitance and resistance of the MEMS-CMOS interconnects. P-type poly-SiGe has such high dopant activation and mobility that annealing the structural film is not necessary, thus reducing the thermal budget and process complexity. No special passivation or other processing is needed to protect the CMOS layers. The etchant of the sacrificial poly-Ge, H_2O_2, does not require any special steps to protect the electronics, is safe to handle, and does not attack exposed bond pads, which further simplifies processing.

REFERENCES

1. M. Biebl, G. T. Mulhern and R. T. Howe, "Low in situ phosphorus doped polysilicon for integrated MEMS," *Technical Digest, 8th International Conference on Solid-State Sensors and Actuators - Transducers 95/Eurosensors IX* (Stockholm, Sweden), Vol. 1, pp. 198-201, 1995.

2. A.E. Franke, D. Bilic, D.T.Chang, P.T. Jones, T.-J. King, R.T. Howe, G.C. Johnson, *MEMS '99*, Orlando, CA, January 1999, pp. 630-7.

3. B. Li, B. Xiong, L. Jiang, Y. Zohar, and M. Wong, "Germanium as a Versatile Material for Low-Temperature Micromaching," *Journal of Microelectromechanical Systems*, 366, Vol. 8, No. 4, (1999).

4. T.-J. King and K. C. Saraswat, "Deposition and properties of low-pressure chemical-vapor deposited polycrystalline silicon-germanium films," *Journal of the Electrochemical Society*, Vol. 141, No. 8, pp. 2235-2241, 1994.

5. T.-J. King, J. P. McVittie, K. C. Saraswat and J. R. Pfiester, "Electrical properties of heavily doped polycrystalline silicon-germanium films," *IEEE Transactions on Electron Devices*, Vol. 41, No. 2, pp. 228-232, 1994.

6. S. Sedky, P. Fiorini, M. Caymax, A. Verbist, and C. Baert, "Thermally insulated structures for IR bolometers, made of polycrystalline silicon germanium alloys," *Technical Digest, 1997 International Conference on Solid-State Sensors and Actuators - Transducers '97* (Chicago), pp. 237-240.

7. J.M. Heck, C.G. Keller, A.E. Franke, L. Muller, T.-J. King, R. T. Howe, *Transducers '99*, Sendai, Japan, June 1999, vol. 1, pp. 328-331.

8. P.T. Jones, G.C. Johnson, and R.T. Howe, "Micromechanical structures for fracture testing of brittle thin films," *(MEMS) ASME*, DSC-Vol. 59, pp. 325-330, (1996).

SIC RESONANT DEVICES FOR HIGH Q AND HIGH TEMPERATURE APPLICATIONS

S. Roy[*†], A.K. McIlwain[*], R.G. DeAnna[‡], A.J. Fleischman[†], R.K. Burla[*], C.A. Zorman[*], and M. Mehregany[*]

[*]Department of Electrical Engineering and Computer Science
[+]Department of Materials Science and Engineering
Case Western Reserve University, Cleveland, Ohio 44106

[†]Department of Biomedical Engineering
The Cleveland Clinic Foundation
Cleveland, Ohio 44195

[‡]U.S. Army Research Laboratory
NASA Glenn Research Center
Cleveland, Ohio 44135

ABSTRACT

Polycrystalline cubic silicon carbide (polySiC) lateral resonant devices have been fabricated by surface micromachining, packaged using ceramic-based materials and nickel wirebonding procedures, and successfully operated over <10^{-5}-760 Torr and 22-950°C pressure and temperature ranges, respectively. Device Q's of >100,000 at <10^{-5} Torr and resonant frequency drifts of <18ppm/hr under continuous operation have been determined using an electron microscope-based setup. Finite element analysis revealed that changes in resonant frequency with increasing temperature depend on the interplay between the decrease in Young's modulus of polySiC and induced stresses in the suspension beams of the lateral resonant device, which occur due to mismatch in thermal expansion coefficients of the polySiC film and the underlying substrate.

INTRODUCTION

Micromachined polysilicon resonators are emerging as potential on-chip replacements for discrete oscillators and filters in high performance communication transceivers [1]. The integrated microelectromechanical polysilicon devices exhibit frequency selectivity characteristics that are superior to integrated RC active filtering techniques based upon traditional electron devices. Micromachined resonator Q's of >80,000 under vacuum conditions and resonant frequency coefficients of <10 ppm/°C have been reported [2]. In addition to high Q transceiver components, the utility of polysilicon resonators has also been demonstrated in a number of other applications including mechanical properties testing, pressure sensing, and inertial navigation systems [3-5]. However, the electrical and mechanical properties of polysilicon begin to rapidly degrade at elevated temperatures (>350°C), making it increasingly unsuitable for high temperature applications [6-8]. In contrast, SiC is well known for its mechanical characteristics, chemical inertness, high thermal conductivity, and electrical stability at temperatures well above 600°C [8, 9]. Although the material properties and microfabrication compatibility of SiC make it an attractive structural material for fabrication of high Q and/or high temperature microelectromechanical devices, the development of SiC as a structural material has been limited by a combination of fabrication, packaging, and testing challenges.

This paper presents development of cubic polycrystalline SiC (polySiC) lateral resonant devices (subsequently called resonators) that are fabricated, packaged, and successfully operated over a wide range of pressures and temperatures. First, the fabrication of polySiC resonators by a surface micromachining process using SiO_2, polysilicon, and nickel (Ni) as the isolation, sacrificial, and contact metallization layers, respectively, is presented. Details of the resonator packaging scheme using ceramic-based materials and Ni wirebonding procedures are also outlined. Next, both optical and electron microscope-based experimental setups and protocols to investigate changes in resonator Q with pressure and resonant frequency with temperature are described. Finally, resonator testing results are presented and the behavior of resonant frequency with temperature is examined using finite element analysis.

FABRICATION

The starting substrate for polySiC film growth is prepared by depositing a 3.5 μm-thick polysilicon film by LPCVD on a 1.5 μm-thick SiO_2 layer grown thermally on 100-mm diameter, (100) Si wafers. A polySiC film with a nominal thickness of 2 μm is grown at 1280°C in a cold-wall, rf-induction-heated, vertical APCVD reactor [10]. The polySiC film is then mechanically polished using a SiC slurry to reduce the surface roughness (Ra) from ~400Å on the as-grown film to <40Å. Figure 1 outlines the fabrication of
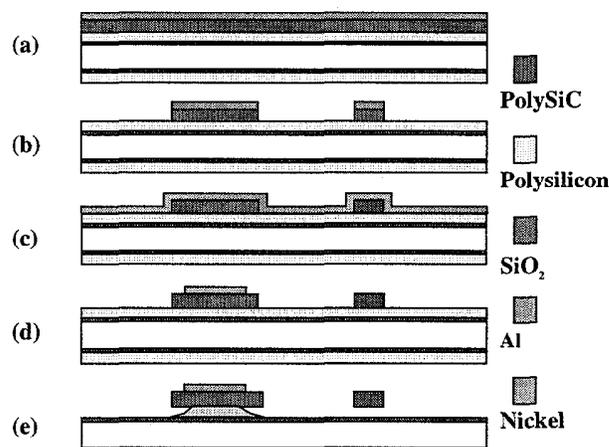
Figure 1. Schematic description of the polySiC resonator fabrication process showing cross-sections after: (a) Al deposition by sputtering; (b) definition of resonator pattern by plasma etching; (c) deposition of Ni by sputtering; (d) definition of Ni contact regions by photolithography and wet etching; and (e) release of polySiC resonators in KOH.
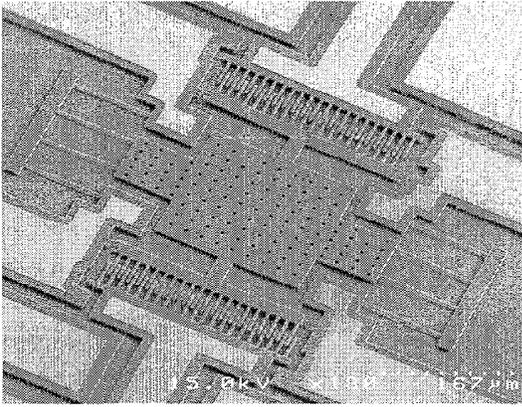
Figure 2. SEM micrograph of a released polySiC lateral resonant device. The suspension beam length and width are nominally 100 μm and 2.5 μm, respectively. Exposed polySiC shows up as dark gray, while the Ni metallization appears light gray.

Figure 3. Cross-sectional schematic view of the polySiC resonator package. External electrical connections to the resonator chip are enabled through the stainless steel screws, washers, hex nuts, gold pads, and nickel wirebonds.

polySiC resonators by surface micromachining. A 5000Å-thick aluminum (Al) film is deposited on the polySiC film by sputtering and subsequently patterned using photolithography and Al etchant to delineate the resonator geometry. Next, the resonator pattern is defined in the polySiC by dry etching in a $CHF_3/O_2/He$ plasma with Al acting as the etch mask [10]. Afterwards, the Al mask is stripped and a 7500Å-thick Ni film is sputter deposited, patterned by photolithography, and wet etched to define nickel contacts to the polySiC. Finally, the resonator is released by a timed etch of the sacrificial polysilicon in 40 wt.% KOH at 40°C and dried using a supercritical CO_2 drying process. Figure 2 presents a SEM micrograph of a released polySiC resonator.

The released polySiC resonator chip is packaged onto a ceramic plate with patterned gold pads and steel posts as shown in Fig. 3. The chip is attached to the ceramic plate using commercially available zirconia-based cement (*AREMCO Ultra-Temp 516*), which is thermally conductive, stable to 1760°C, and resistant to chemical attack. The gold pads on the ceramic plate enable electrical connection from the resonator chip to external circuitry through the stainless steel screws, nuts, and washers. Electrical connections from the Ni contact pads on the resonator chip to the gold pads on the ceramic plate are realized by 25 μm-diameter Ni wires. Attachment of Ni wires to the pads is achieved by a thermosonic wirebonding process optimized for maximum pull strength [11].

EXPERIMENTAL SETUPS

The packaged resonators are tested under different conditions of pressure and temperature. Figure 4 presents the experimental setup used to test devices at atmospheric pressure (760 Torr). The resonator package is placed on a heater capable of attaining temperatures as high as 1000°C. The heater is mounted on a thermally-insulating ceramic tile inside an aluminum chamber with openings for argon, electrical connections, and optical access. Probe tips are used to electrostatically actuate the resonator, while the microscope is used to monitor device motion. Argon is introduced at ~2 psi into the chamber during testing to minimize thermal oxidation effects on the probe tips and heater.

Figure 5 presents the experimental setup used to test devices under vacuum conditions. The resonator package is placed on a hot
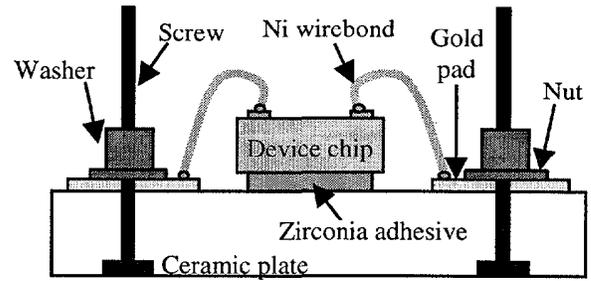
stage mounted inside an environmental SEM (ESEM) and connected to external circuitry via electrical feedthroughs. The resonator motion is observed using an attached monitor. In addition to room-temperature imaging capabilities of a standard SEM under high vacuum ($<10^{-5}$ Torr) conditions, the ESEM can also perform
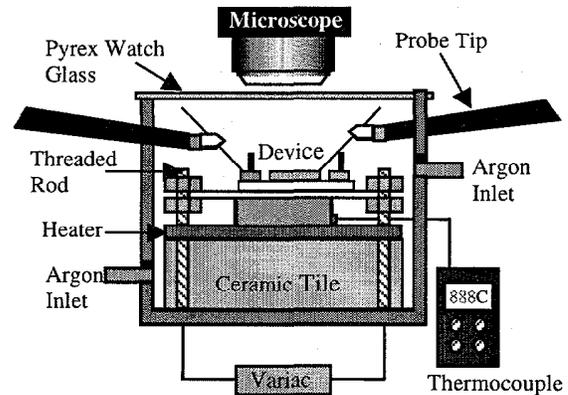
Figure 4. Test setup for atmospheric pressure (760 Torr) experiments. The optical microscope is used to observe resonator motion. Measurement uncertainty in resonant frequency is ±50 Hz.
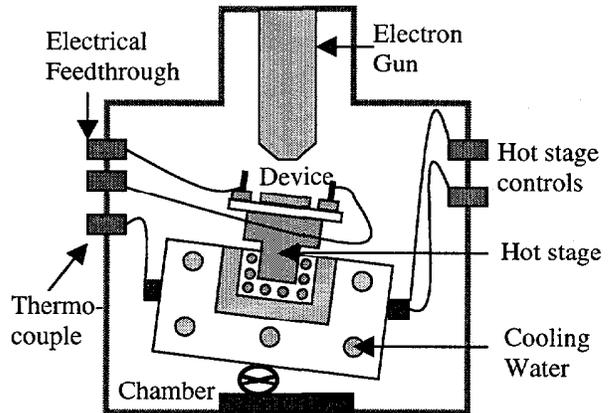
Figure 5. Schematic description of the environmental SEM setup used for experiments under vacuum conditions. Resonator motion is observed on an attached monitor. Measurement uncertainty in resonant frequency is ±0.01 Hz.

imaging under low vacuum (0.1-20 Torr) and/or high temperature (1000°C) conditions.

Experiments are performed to determine the resonant frequency and Q of the resonator under different conditions of pressure and temperature. Resonant frequency is determined visually by adjusting the frequency of the excitation voltage to the resonator until maximum resonator amplitude is observed. Resonator Q is determined from knowledge of the resonant frequency and 3dB frequencies, which are defined as those frequencies at which resonator amplitude is ~70% of the maximum [12].

RESULTS

Table I presents a summary of resonator Q's of three polySiC resonators measured at room temperature (22°C) under three different pressure regimes. The low-pressure tests are conducted inside the ESEM, while the atmospheric pressure tests are performed using the optical microscope setup. The nominal resonator frequencies are determined under high vacuum conditions. The resonator Q's vary with pressure, with maxima and minima exhibited under high vacuum and atmospheric conditions, respectively. In general, resonator Q's are below 150 for all devices at 760 Torr, between 2500 and 4000 at 0.1-0.2 Torr, and above 50000 at $<10^{-5}$ Torr. This increase in resonator Q with decreasing pressure is attributed to damping effects: at lower pressures, the damping is minimized and Q is increased. The effect of damping on resonators is dominant for pressures down to $\sim 10^{-3}$ Torr, and at extremely low pressures (high vacuum conditions), the Q is limited by the internal friction of polySiC [7, 13]. A 800 million cycle, 16 hr aging test performed at 22°C on a ~14 kHz resonator reveals that the resonant frequency drift is approximately 17.5 ppm/hr.

Figure 6 presents a graph showing the variation of resonant frequency with temperature for a polySiC resonator actuated at 150Vpp under atmospheric conditions. The corresponding resonant frequency drops from 30400 Hz at 22°C to 30000 Hz at 410°C. The resonator does not operate above 500°C due to breakdown of the electrically isolating SiO_2 layer. Figure 7 presents a graph showing the variation of resonant frequency with temperature for a comparable polySiC resonator under low vacuum (0.1-0.2 Torr) conditions. The device is actuated at 3.5Vpp and operates over a 22-950°C temperature range. The resonant frequency of the vacuum-tested device drops steadily from 28930 Hz at 22°C to 28402 Hz at 700°C, followed by an increase to 28544 Hz at 900°C, and then a slight decrease to 28500 at 950°C.

Finite element analysis is performed to investigate changes in resonant frequency with increasing temperature. A preliminary finite element model is analyzed using ANSYS software to examine interactions between the polySiC resonator and silicon substrate [14]. The bottom of the silicon substrate is fixed in the direction normal to the surface and allowed free thermal expansion

Figure 6. Variation of resonant frequency with temperature for a polySiC resonator operated at atmospheric pressure at 150Vpp actuation. Device does not operate above 500°C due to breakdown of the insulating SiO_2 layer.
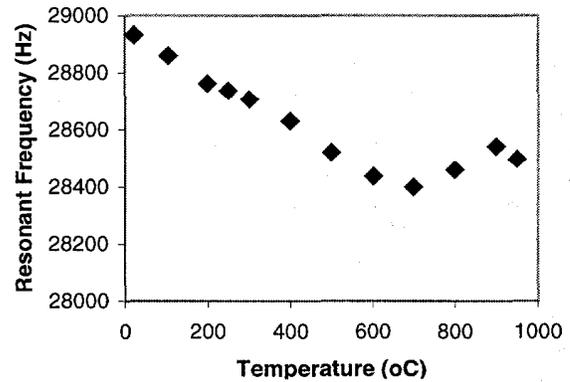
Figure 7. Variation of resonant frequency with temperature for a polySiC resonator (comparable to the one in Fig. 6) under vacuum conditions (0.1-0.2 Torr) at 3.5Vpp actuation.

in directions parallel to the surface. Figure 8 presents a graph showing the variation of thermal expansion coefficients with temperature for cubic SiC (assumed equivalent to polySiC) and silicon, while Table II lists the other material properties used in the finite element analysis [14-16].

Figures 9 and 10 present graphs of normalized resonant frequency versus temperature based on the results of the finite element analysis and the experimental data shown in Fig. 7, respectively. In both cases, there is an initial steady decrease in resonant frequency, followed by a slight increase, and another slight decrease. The differences between the finite element results and experimental data may due to variations between the actual material property parameters and those used in the model as well as the influence of intrinsic stresses in the various material layers.

Examination of the material properties used in the finite element model provides qualitative insight into the resonant frequency versus temperature behavior for the devices. The thermal expansion coefficient of polySiC is lower than that of silicon from 22°C to ~200°C, while at higher temperatures, the thermal expansion coefficient of polySiC is greater than that of silicon. Consequently, at temperatures below 200°C, the silicon substrate expands faster than the polySiC layer, which induces a resultant compressive thermal stress in the suspension beams of the

Table I. Resonator Q determined at different pressure regimes.

Device #	f_r (Hz)	Q_1 (±5%)	Q_2 (±10%)	Q_3 (±15%)
1	10060	25	2875	62888
2	11360	34	3474	107926
3	28920	148	3616	57828

f_r: resonant frequency; Q_1: 760 Torr; Q_2: 0.1-0.2 Torr; and Q_3: $<10^{-5}$ Torr.
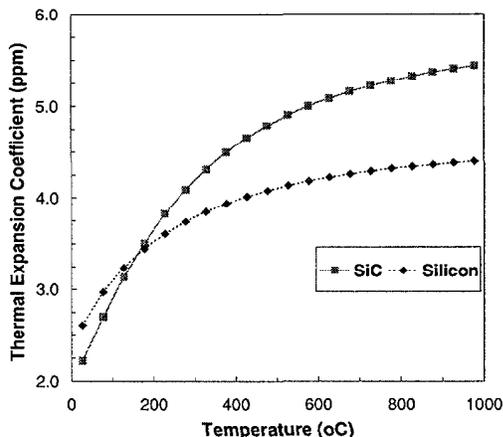
Figure 8. *Variation of thermal expansion coefficient of SiC and silicon as a function of temperature [14-16].*

Table II. *Materials properties used in finite element analysis [14].*

Material	ρ (kg/m³)	E (GPa)	α_E (ppm/°C)
SiC	3230	400	46
Silicon	2320	160	80

ρ: density; E: Young's modulus; and α_E: temperature coefficient of Young's modulus.

resonator. The net effect of the induced compressive stress and decrease in Young's modulus of polySiC leads to a decrease in the resonant frequency. As temperatures increase above 200°C, the polySiC film begins to expand faster than the silicon substrate, which induces a resultant tensile stress in the resonator suspension beams. The lowering effect on the resonant frequency due to the decrease in Young's modulus is counteracted by the induced tensile stress, which tends to increase resonant frequency. The net effect of increasing tensile stress and decreasing Young's modulus with increasing temperature leads to a resonant frequency minimum, and then, an increase. However, as the temperature increases even further, the effect of the decrease in Young's modulus dominates leading to the eventual decrease in resonant frequency.
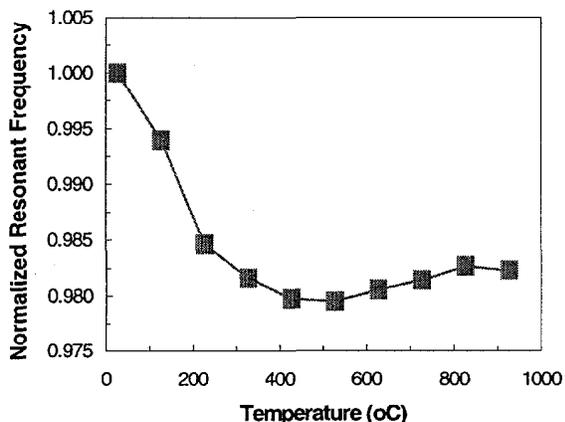
Figure 9. *Predicted resonant frequency as a function of temperature obtained using finite element analysis.*
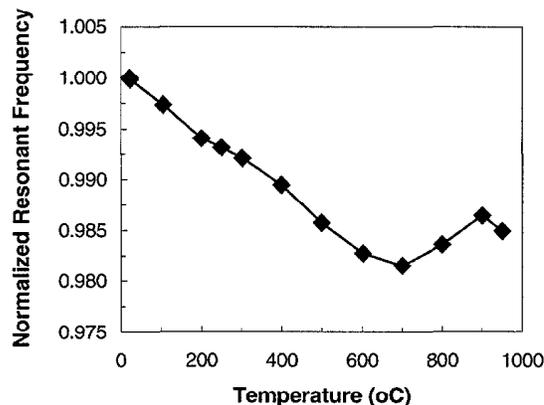
Figure 10. *Normalized experimental data of Fig. 7.*

CONCLUSIONS

PolySiC resonators have been fabricated by surface micromachining and successfully packaged for operation in vacuum and high temperature environments. Resonator Q's above 100,000 under vacuum conditions have been observed. Changes in resonant frequency with increasing temperature are attributed to a decrease in Young's modulus as well as thermal expansion mismatch with the underlying substrate.

REFERENCES

[1] C. T.-C. Nguyen, *Proc. 1997 IEEE Intl. Symp.Circuits & Syst.*, Hong Kong, (1997) pp. 2825-2828.
[2] C. T.-C. Nguyen and R. T. Howe, *Tech. Dig., 1993 IEEE IEDM*, Washington, D.C., (1993) pp. 199-202.
[3] H. Kahn, S. Stemmer, K. Nandakumar, A. H. Heuer, R. L. Mullen, R. Ballarini, and M. A. Huff, *Proc. 1996 IEEE MEMS Wkshp.*, San Diego, CA, (1996) pp. 343-348.
[4] C. J. Welham, J. Greenwood, and M. M. Bertioli, *Proc. 1998 Eurosens. XII*, (1998) pp. 298-304.
[5] K. Funk, H. Emmerich, A. Schilp, M. Offenberg, R. Neul, and F. Laermer, *Proc. 1999 IEEE MEMS Conf.*, Orlando, FL, (1999), pp. 57-60.
[6] G. Pearson, W. R. Jr., and W. Feldman, *Act. Metall.*, v5, p. 181, 1957.
[7] M. Biebl, G. Brandl, and R. T. Howe, *Proc. Intl. Conf. Solid-State Sens. & Act., & Eurosens. IX,* Stockholm, Sweden, (1995) pp. 80-83.
[8] A. J. Fleischman, S. Roy, C. A. Zorman, and M. Mehregany, *Proc. Intl. Conf. SiC, III-Nit. & Related Mat.*, Stockholm, Sweden, (1998) pp. 889-892.
[9] G. Krotz, W. Legner, C. Wagner, H. Moller, H. Sonntag, and G. Muller, *Proc. Intl. Conf. Solid-State Sens. & Act., & Eurosens. IX,* Stockholm, Sweden, (1995) p. 186.
[10] A.J. Fleischman, S. Roy, C. A. Zorman, M. Mehregany, and L. G. Matus, *Proc. IEEE MEMS Wkshp.*, San Diego, CA, (1996) pp. 234-238.
[11] R. K. Burla, S. Roy, V. M. Haria, C. A. Zorman, and M. Mehregany, *Proc. Conf. Des., Charac., & Packag. MEMS & Microelect.*, SPIE Symp. Microelect. & MEMS, Melbourne, Australia, (1999) pp. 324-333.
[12] W. Tang, *Ph.D. Dissert.*, U.C. Berkeley, 1990.
[13] C. M. Su, M. Wuttig, A. Fekade, and M. Spencer, *J. App. Phys.*, vol. 77, pp. 5611-5615, 1995.
[14] R. G. DeAnna, S. Roy, C. A. Zorman, and M. Mehregany, *Proc. Intl. Conf. Mod. & Sim. Microsyst.* San Juan, Puerto Rico, (1999) pp.644-647.
[15] R. R. Reeber and K. Wang, *Mat. Chem. & Phys.*, v46, pp. 259-264, 1996.
[16] R. R. Reeber and K. Wang, *Proc. Symp. Covalent Ceramics: Sci. & Tech. Non-Oxides*, MRS Fall Meeting, Boston, MA, (1995) pp. 211-216.

SEPARATION OF DNA IN MICROFLUIDIC SYSTEMS

H. G. Craighead, J. Han and S. W. Turner

School of Applied and Engineering Physics, Cornell University

Ithaca, NY 14853

ABSTRACT

It is possible to fabricate fluidic systems with dimensions comparable to the size of DNA molecules. By controlling the time and spatial dependence of electric fields one can drive, sort and separate molecules by a variety of physical mechanisms. We have created nanoconstrictions in fluidic channels that act as entropic barriers to DNA motion. We have used the size dependence of these entropic effects to separate DNA. In this paper we describe the design and fabrication two such types of devices and their performance.

INTRODUCTION

Gel electrophoresis is the standard method of separating DNA for genetic sequencing and analysis. The size-dependent electrophoretic mobility of DNA when driven through a polymeric restricting gel results in a spatial separation of DNA fragments by length. This separation appears as bands in an electropherogram. The efficiency of separation of DNA by gel electrophoresis decreases with increasing molecular length as the mobility differences become less pronounced for larger molecular sizes. We have addressed primarily this issue in the systems described here. We can design the molecular size region of interest by the choice of geometry and dimensions that are defined by lithography.

The use of microfabricated systems offers several additional possible improvements for DNA separation devices. Microfabricated systems could, for example, be designed to work with smaller amounts of material. The microfabrication approach also easily allows for parallelism so that many separations can take place simultaneously on a "chip". If gel-free systems can be used, it should be easier integration of different processes on a single chip and extraction of separated DNA should be more straightforward. Gel-free microfabricated approaches for the creation of artificial sieving structures have been suggested and met with some success. [1] Microfabricated devices that sort DNA molecules by diffusion rate have also been demonstrated. [2]

We have developed several fabrication approaches that can be integrated with optical detection and incorporate controlled geometrical constrictions in fluid channels. These approaches have enabled fabrication of structures smaller than the radius of gyration of large DNA molecules and could be used for rapid separation of the DNA.

ENTROPIC TRAPPING ARRAY

We fabricated a fluid channel with an array of narrow gaps. The cross sectional schematic of this is shown in Figure 1. The device is fabricated with two photolithography and etch steps, with the dimensions of the gap determined by an etch depth. The channel is sealed by anodically bonding a cap on the etched channels. DNA molecules are driven by electrophoresis through the series of narrow slits with a static electric field. [3] The narrow slit width is smaller than the radius of gyration of the DNA molecules.
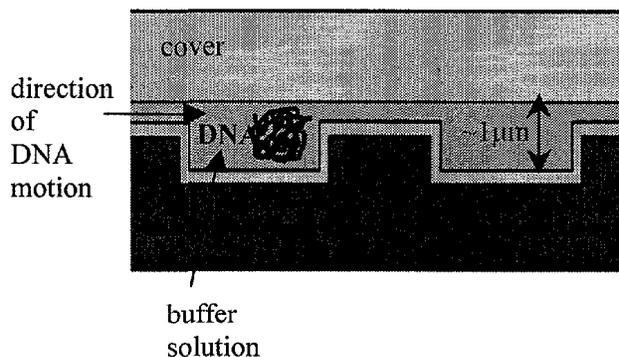
Figure 1. *Cross sectional schematic of an entropic trap array.*

In order for the DNA to enter the narrow region the molecule must undergo an entropically unfavorable change in conformation. The competition of the electrostatic forces driving the molecule into the narrow region with the entropic hindrance creates a free energy barrier. The probability that a molecule will surmount this barrier is proportional to the width of the molecule encountering the narrow gap. This width is approximately proportional to the square root of the number of base pairs in the molecule. [4]
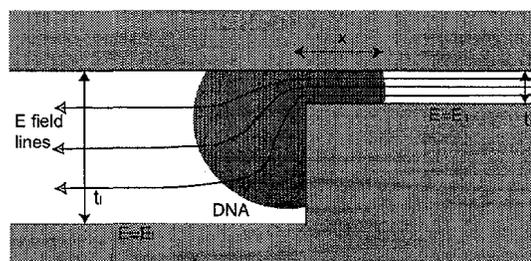
Figure 2. *Schematic of the deformation and escape of a DNA molecule from a trap.*

Figure 3 shows the results of a separation of a ladder of DNA molecules with lengths from 10 kilobases to 48 kilobases. The resolution is comparable to that from the conventional pulsed field gel electrophoresis, but can be done in substantially less time. [5]
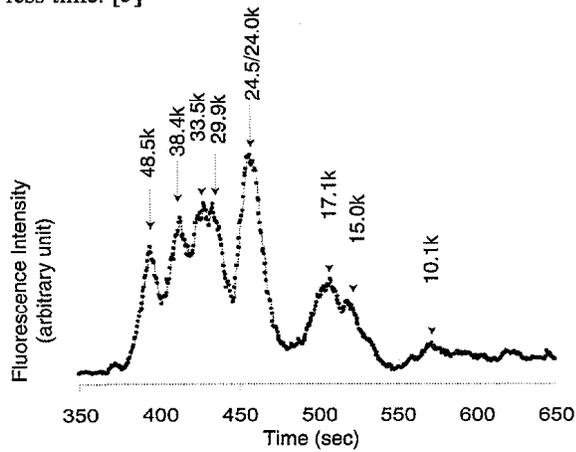
Figure 3. *Separation of a DNA sample consisting of different length fragments.*

This same approach should also work for other polymeric molecules.

ENTROPIC RECOIL DEVICE

We can exploit a related length-dependent trapping effect to separate DNA molecules by length. In this case we have fabricated a series of obstructions consisting of vertical columns in a thin fluid channel. This structure is formed by high-resolution lithography and sacrificial removal of a patterned layer to form the fluid system. [6] A micrograph illustrating the geometry is shown in Figure 4.
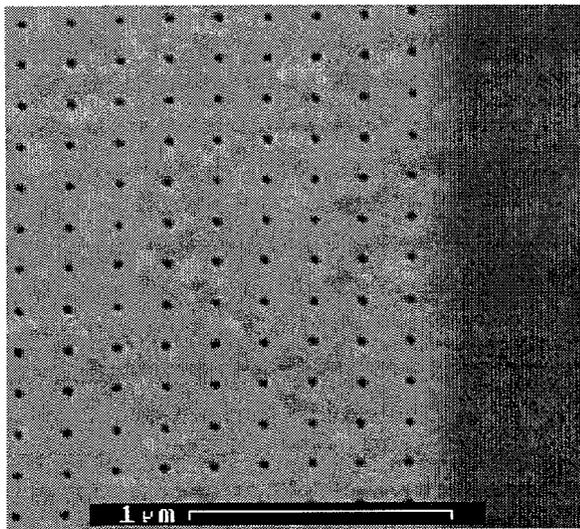
Figure 4. *Electron micrograph showing a top view of entropic recoil device with the unconstrained fluid region bounded by the obstruction filled constrained region.*

This structure creates an entropic trap similar to that described for the trapping array discussed above but with a different dimensionality. In this case we exploit the time difference for the passage of molecules into the constrained region. Figure 5 illustrates the basic principal of this approach.[7]

In this approach a pulsed electric field is used to drive the molecules. During the pulse molecules are driven from the entropic trap region. When the electric field is turned off molecules only partially into the constrained region will recoil and remain trapped. Longer pulse length results in the release of longer molecular lengths from the trapped area. Varying the pulse length, therefore, will result in sequential release of molecules of varying length.
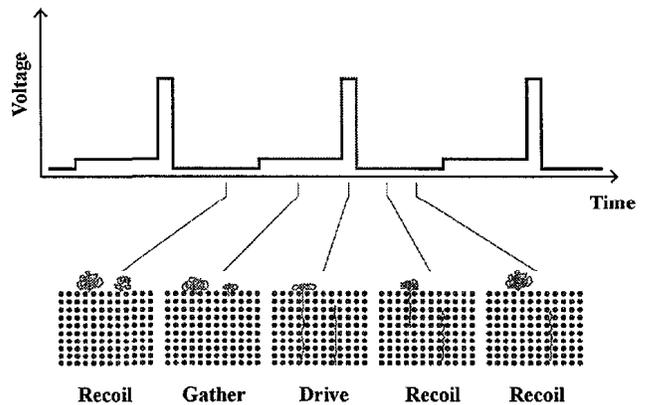
Figure 5. *Schematic of the entropic recoil device function, illustrating the behavior of a long and short DNA molecule in the time varying electric fields.*
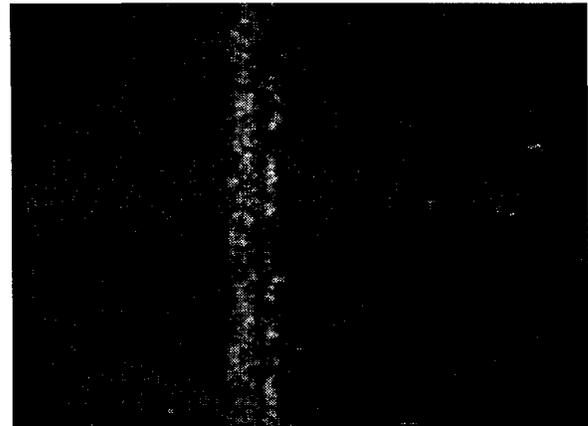
Figure 6. *Fluorescence image of a band of smaller length DNA fragments released from the trapping region while the longer molecules remain trapped.*

The result of such a pulsed separation is shown in Figure 6. This fluorescence micrograph shows a band of short fluorescently labeled DNS molecules launched from the trap and traveling in the channel. These recoil structures could be staged to enhance the separation process.

CONCLUSIONS

It is possible to rapidly sort DNA molecules by their trapping and release from nanofabricated constrictions. We have fabricated and tested several different systems for rapid separation of molecules in the kilobase size regime. The length region of operation should be scalable with choice of the constriction dimensions to similarly separate larger or smaller scale DNA molecular sizes. The fabrication approach with geometry controlling the device behavior allows this to be accomplished in a variety of material systems. It is also possible to integrate these fluid channels with other functional devices such as in-plane optical waveguides for integrated fluorescent excitation and detection.

ACKNOWLEDGEMENTS

REFERENCES

1. W. D. Volkmuth and R. H. Austin, *Nature, 358,* 600 (1992).

2. C.F. Chou et al. *Proc. Nat. Acad. Sci. 96,* 13762 (1999).

3. J. Han and H. G. Craighead, "Entropic trapping and sieving of long DNA molecules in a nanofluidic channel", *J. Vac. Sci. Technol A, 17,* 2142 (1999).

4. J. Han, S. W. Turner and H. G. Craighead, "Entropic trapping and Escape of long DNA molecules as submicron size constrictions", *Phys. Rev. Lett., 83,* 1688 (1999).

5. J. Han and H. G. Craighead, "Separation of long DNA molecules in a microfabricated entropic trap array", (to appear in *Science*).

6. S. W. Turner, A. M. Perez, A. Lopez, and H. G. Craighead, "Monolithic nanofluidic sieving structures for DNA manipulation" *J. Vac. Sci. Technol B, 16,* 3835 (1998).

7. S. W. Turner, M. Cabodi and H. G. Craighead "Entropic recoil DNA separation ", APS Meeting , March 2000.

A MICROELECTRODE ARRAY FOR REAL-TIME NEUROCHEMICAL AND NEUROELECTRICAL RECORDING *IN VITRO*

Timothy D. Strong[a,b], Hal C. Cantor[a], Richard B. Brown[b]

[a]Advanced Sensor Technologies, Inc., Farmington Hills, MI 48334
[b]Department of Electrical Engineering and Computer Science,
The University of Michigan, Ann Arbor, MI 48109

ABSTRACT

The development of a silicon neural recording device capable of measuring neurochemical and neuroelectrical activity in tissue culture is described. The device consists of a microfabricated array of platinum and chloridized silver electrodes on a silicon substrate. An affixed culture chamber on the surface of the device contained the culture media. Neuron cultures survived for over 75 days in the devices, generating a great deal of neuroelectrical and neurochemical data. The performance of the devices allowed for the study of single neuron interactions with the ability to correlate electrical signals directly with a chemical response.

INTRODUCTION

The understanding of communication between neurons is key to the advancement of neuroscience. It is important for not only purely scientific reasons, but may provide assistance in studying various neurological injuries and disorders, and possibly lead to treatments.

Miniature silicon neurosensors have been available to neuroscientists for many years [1-7]. These devices have enhanced the study of neural signaling significantly. Arrays of sites can be easily produced on such sensors, allowing the study of multiple groups of neurons concurrently. In addition, the silicon substrate enables the inclusion of active electronics, providing low noise, on-chip amplification, and multiplexing to reduce lead count. Silicon micromachining techniques allow the final shape of the devices to be tailored to the application, from fine needle-like probes to three-dimensional arrays. To date, these devices have provided neuroelectrical signals but no chemical data.

Carbon fibers affixed in pulled glass micropipettes are among the most common devices for neurochemical sensing. The pipettes are drawn under heat and broken, exposing the carbon fibers at the tip, which form the electrode [8, 9]. Electronics can be used to allow the detection of both electrical and chemical signals. Unfortunately, their manufacture makes it difficult to insure consistent performance. It is also difficult to create an array, since the sensors are manufactured individually. Amplification must be performed using discrete components, requiring long interconnects that are subject to noise.

The neural array described here combines the advantages of both the silicon neurosensors and the carbon fiber electrodes, while eliminating most of the complications. The devices are fabricated using standard silicon processing technology. This provides high yield and consistent performance due to the tight dimensional control of photolithography. Large sensors arrays facilitate the study of small neural systems in addition to individual neurons. Using these devices, real-time neurochemical *and* neuroelectrical activity of the neurons can be recorded. This provides the important capability of studying the correlation of electrical and chemical events in neuronal communication.

VOLTAMMETRY IN NEUROSENSORS

Many physiologically important neurotransmitters, including dopamine and serotonin, are electroactive. The application of a potential will cause electrons to be removed from the species of interest as it undergoes electrochemical oxidation. The observed current is related to the concentration of the species.

In voltammetry, a potential is ramped between two electrodes (usually referred to as a working and a reference electrode) while the current is recorded. As the oxidation potential of each electroactive species is reached, a corresponding current increase will be observed. Adding a third electrode, called an auxiliary electrode, increases sensitivity. With careful circuit design, the current can be made to flow between the working and auxiliary electrodes. The potential is measured between the reference and the working electrodes with little current being carried by the reference electrode. Placement of the reference electrode close to the working electrode removes the effect of differences in solution conductivity from culture to culture [10]. A single reference electrode can be shared among several sensing sites.

Cyclic voltammetry (CV) extends the voltage sweep by reversing the potential and causing the newly oxidized species to undergo reduction. The reduction cycle provides additional information for the identification of species and also mediates some of the toxicity of the chemicals produced during the oxidation scan [9, 11]. Fast sweep rates, employed in fast cyclic voltammetry (FCV), provide near real-time analysis. A baseline curve must be obtained before testing begins to allow for the charging currents of the electrodes to be recorded and removed from the final results.

A neighboring electrode can be used for recording electrical signaling, or a grounded auxiliary system can provide the capability of recording both signals from the same electrode [12-14].

DESCRIPTION AND FABRICATION

The neural arrays consist of sixteen groups of three-electrode sites. Each group incorporates platinum working and auxiliary electrodes, and a chloridized silver reference electrode. A schematic of the array is shown in Figure 1. Minimum electrode dimensions ranged from 2 to 100 μm on separate devices to allow the study of the effect of electrode size on performance. Individual site orientation was alternated. This allowed neighboring sites to be used in unison as a single larger electrode, providing additional information about the effect of electrode area on performance.
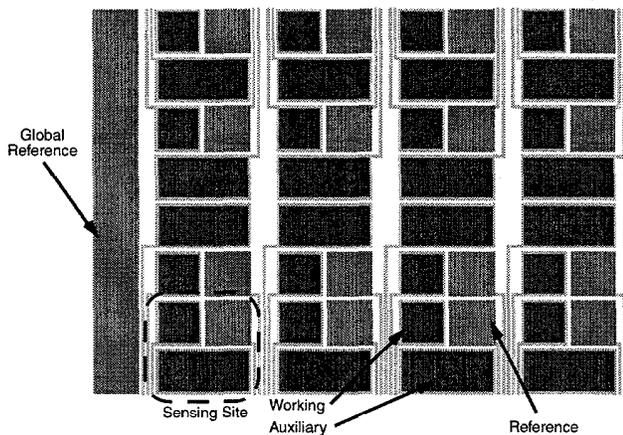
Figure 1. A schematic of the neural array. Site orientation is rotated to allow the combining of neighboring electrodes to study the effect of electrode size on performance.

Fabrication begins on standard 4" <100> silicon substrates. A lower thermal oxide of 1.2 μm provides electrical isolation. Platinum was used for interconnect due to its ability to serve as an excellent polarizable interface in solution. A modified lift-off procedure was used for patterning the interconnect. A timed etch in hydrofluoric acid "submerged" the area where the platinum was applied. This resulted in a nearly planar surface, improving the coverage of the top isolation layer. A 20 nm layer of titanium was used to help adhesion of the 100 nm platinum layer. Several sensor arrays were fabricated using a gold layer with platinum on the surface. It was hoped that the gold would provide devices with a lower resistance.

An isolation layer of 800 nm of PECVD silicon nitride was deposited over the platinum interconnect. A reactive ion etch was used to open the contacts. The resist for the silver lift-off step was then applied. However, before the silver was deposited, a second reactive ion etch was performed to etch 700 nm of the nitride layer. This left a shelf of nitride around the bottom of the silver sites as shown in Figure 2. In the past, our work with silver at these dimensions formed thin wall-like projections by coating the sidewalls of the lift-off resist. Later handling often caused these walls to collapse. The small pieces of silver left behind can move around on the device causing contamination of the platinum electrode surfaces. The etch helped to form a break line at the resist edge, preventing the formation of the walls. The etch was stopped short to prevent the formation of a seam between the silver sidewalls and the surrounding silicon nitride, since a seam might provide a route for encroachment of the test solution resulting in a dual contact.

Silver was applied 650 nm thick over a 20 nm layer of titanium as an adhesion promoter. The final lift-off was performed in
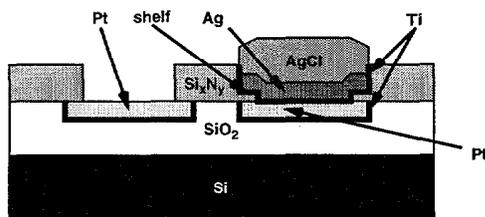
Figure 2. A cross-section of the completed passive array. Note the nitride shelf used to prevent solution from forming a dual contact to the electrode.
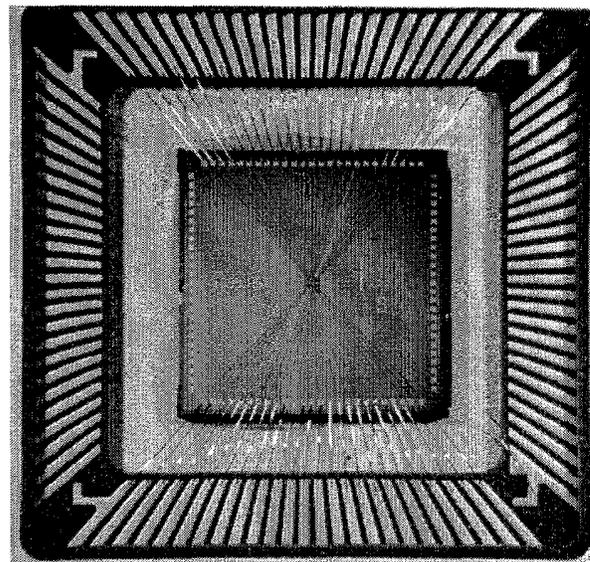
Figure 3. A packaged device wirebonded into ceramic pin grid array carrier before sealing with epoxy.

acetone. Early work using Microposit 1112A photoresist remover (Shipley, Marlboro, MA, U.S.A.) caused discoloration of the silver reference electrodes. Acetone was less reactive to the silver; however, it lacks surfactants which prevent re-deposition of the silver. Care was required to insure all the waste metallization was completely removed in the lift-off process.

Finished chips were wire-bonded into ceramic 100-pin grid array packages (#ECPG10020, Spectrum Semiconductor Materials, Inc., San Jose, CA, USA). The ceramic package provided both a high I/O count for contact to the array and a large thermal mass. Cultured neurons are sensitive to thermal disturbances, and the ceramic package formed an adequate thermal buffer to allow the neuron cultures to be moved from an incubator into a test apparatus. Epoxy was used to insulate the wire bonds from the solution. Culture cylinders were then attached to the surface to provide containment of the culture medium. The resulting package system can be autoclaved for easy sterilization. A bonded device is shown in Figure 3 before epoxy sealing.

RESULTS

The culture cylinder was infused with neuron growth media for initial testing. After initial scans to establish the charging currents for the electrode array, calibrated amounts of neurotransmitters were added to the media, and cyclic voltammograms were recorded. The temperature was carefully maintained by incubation to match neural culture conditions as closely as possible.

Three distinct peaks appeared in the dopamine calibration voltammograms. A peak at 500 mV was related to dopamine concentration. Another peak with a potential of >1 V may have been caused by the oxidation of the phenol red in the culture media. The other peak at 650 mV is currently under investigation. Figure 4 shows the response of the 500 mV dopamine peak and the 650 mV peak, to increasing dopamine concentration. While the 500 mV peak exhibits a linear dependence on dopamine concentration, the 650 mV and 1 V peaks remain constant, demonstrating that the
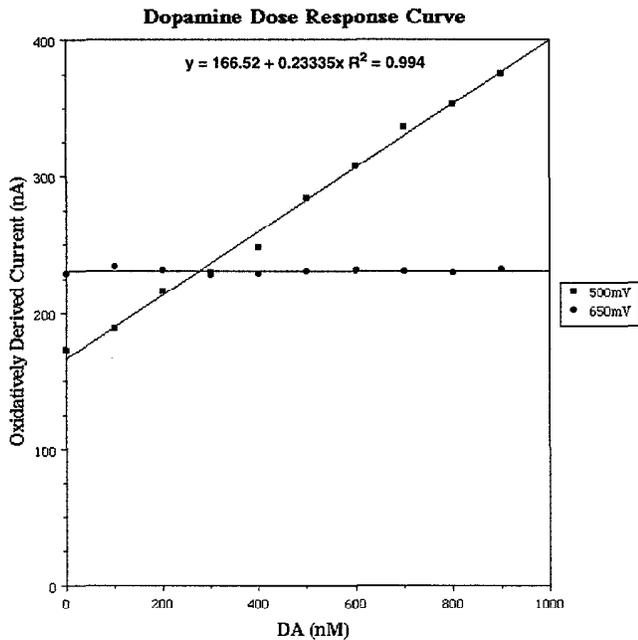
Figure 4. A dopamine calibration response curve. Notice that the neighboring 650 mV peak is unaffected by dopamine concentration and represents a different electroactive compound in the culture media.

devices can distinguish dopamine from other electrochemically active species present in the culture media.

Live cell studies were performed with neuron cultures (hNT cell line from Stratagene, La Jolla, CA, U.S.A.). This cell line was chosen because it is known to be spontaneously active and requires no external stimulation to produce coordinated neural activity. These neurons release dopamine and acetylcholine.

Twenty-eight packaged neural sensor arrays were sterilized and prepared with cultures. After six days of incubation, 24 of the cultures were electrically active. The cultures were monitored for 75 days at which point the experiment was terminated.

For these preliminary tests, the internal reference electrode was replaced with a single external electrode. This simplified instrumentation by allowing the internal electrode to be used to detect neuroelectrical signals, but required all the sites to be swept in unison while performing cyclic voltammograms.

The results demonstrate the ability to detect neurotransmitter concentrations and to differentiate between dopamine and acetylcholine. The ability to concurrently record both neurochemical and neuroelectrical activity has been demonstrated (see Figure 5). Of particular note is the effect of the second group of spikes that apparently cause a release of neurotransmitter.

The devices have sufficient resolution to show the apparent interaction of a single action potential and the resulting neurotransmitter release (see Figure 6). Correlations between action potentials across the array were observed, showing the propagation of neural signals throughout the neural network.

The devices with platinum interconnect and those with a gold conductivity layer were both found to perform well. However, the pure platinum devices performed more closely to theoretical expectations. In the gold devices, it is suspected that the gold base layer and the platinum top-layer interdiffused. This could have occurred while the sealing epoxy was being cured in an oven.
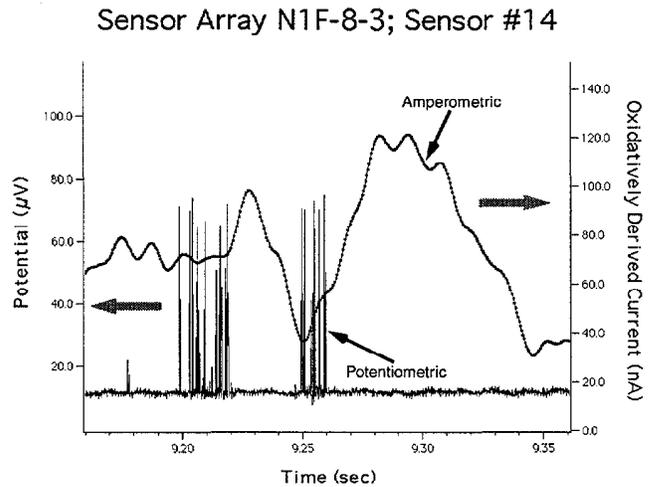
Figure 5. The combined chemical and electrical signal from a sensor site. The amperometric signal represents varying neurotransmitter concentrations; the potentiometric signal provides "spike" recording.

Diffusion between platinum and gold can occur at temperatures as low as 200 °C [15]. Further evidence of this effect was observed when the resistance of the interconnect lines was measured. A distinct change in resistance was seen from an initial value to a higher value that stabilized with time. The current being applied to measure the resistance may have been high enough to cause the metals to interdiffuse.

An additional problem occurring on some devices was corrosion of the silver reference electrodes. This was observed on devices immediately after dicing, and is believed to be an electrolytic effect caused by the water used to cool the dicing saw. An electrolytic cell is formed between the large platinum bond pads and the smaller silver reference sites. Where a manufacturing defect broke the connection, no corrosion was seen. Future devices will incorporate a polymer layer to prevent this silver corrosion. The polymer can be easily removed with a solvent after dicing.
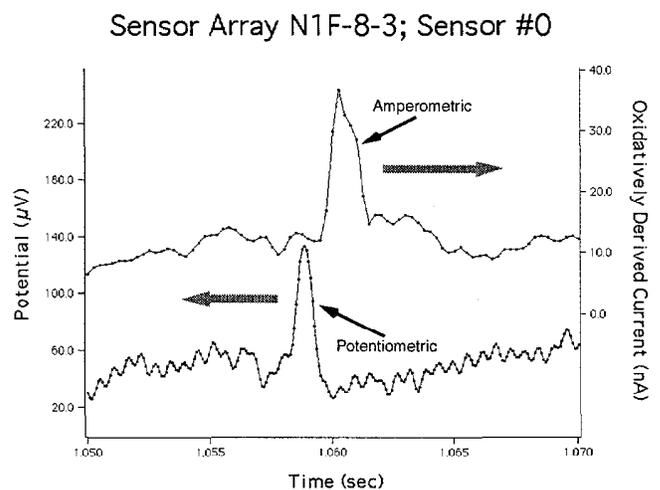
Figure 6. Superimposed potentiometric and amperometric traces. The amperometric signal appears to reveal the release of neurotransmitter in response to the action potential shown in the potentiometric recording.

31

CONCLUSIONS

These new devices provide an exciting advancement in neuroscience. Due to their fabrication using silicon processing techniques, they are relatively simple to produce in large quantities. Their performance is more consistent from device to device than the carbon fiber electrodes that have been used in previous studies. Unlike previous silicon neural devices, these arrays provide access to both neuroelectrical and neurochemical activity concurrently, providing the ability to study electrical and chemical interactions in neuronal communication. In addition, they provide the capability to study a system of neurons, allowing the study of cause-effect relationships. The use of cell culture provides the ability to study neural systems with no trauma to the tissue, such as that incurred by *in vivo* studies, and provides a long window of study, over 75 days.

The use of spontaneously active neurons eliminated the need for external stimulation of the culture. This reduces potential external influence on the neurons and improves the validity of the results. The silicon substrate allows for the incorporation of active electronics for amplification, noise reduction, and expansion of the array size while reducing lead counts.

The current, patent-pending, passive neural arrays hold a great deal of promise in future work, and research and development on them continues. Improvements in metallization are intended to remove remaining corrosion problems and more advanced lithography techniques will be applied to reduce the electrode size below 2 μm.

This work will allow neuroscientists to study the communication between neurons in much greater detail than was previously possible. Such research will enhance our understanding of the nervous system, and may provide valuable insights into a wide variety of neural disorders.

ACKNOWLEDGEMENTS

REFERENCES

1.   K. D. Wise, J. B. Angell, A. Starr, "An Integrated-Circuit Approach to Extracellular Microelectrodes," *IEEE Transactions on Biomedical Engineering*, v. BME-17, n. 3, pp. 238-247, 1970

2.   M. Kuperstien, D. A. Whittington, "A Practical 24 Channel Microelectrode for Neural Recording *in Vivo*," *IEEE Transactions on Biomedical Engineering*, v. BME-28, n. 3, pp. 288-293, 1981

3.   N. A. Blum, B. G. Carkhuff, H. K. Charles, Jr., R. L. Edwards, R. A. Meyer, "Multisite Microprobes for Neural Recordings", *IEEE Transactions on Biomedical Engineering*, v. 38, n. 1, pp. 68-74, 1991.

4.   P. Connolly, P. Clark, A. S. G. Curtis, J. A. T. Dow, C. D. W. Wilkinson, "An Extracellular Microelectrode Array for Monitoring Electrogenic Cells in Culture," *Biosensors & Bioelectronics*, v. 5, pp. 223-234, 1990.

5.   E. Valderrama, P. Garrido, E. Cabruja, P. Heiduschka, A. Harsch, W. Göpel, "Microfabrication and Characterization of Microelectrode Arrays for *in Vivo* Nerve Signal Recording," *Proceedings of the 8th International Conference on Solid-State Sensors and Actuators, and Eurosensors IX*, pp. 63-66, June 25-29, 1995.

6.   S. L. BeMent, K. D. Wise, D. J. Anderson, K. Najafi, K. L. Drake, "Solid-State Electrodes for Multichannel Multiplexed Intracortical Neuronal Recording," *IEEE Transactions on Biomedical Engineering*, v. BME-33, n. 2, pp. 230-241 (1986).

7.   C. Kim, K. D. Wise, "A 64-Site Multishank CMOS Low-Profile Neural Stimulating Probe," *IEEE Journal of Solid-State Circuits*, v. 31, n. 9, pp. 1230-1238 (1996).

8.   J. Krüger, "Multiple Microelectrodes," *Methods in Neurosciences: Electrophysilogy and Microinjection*, v. 4, New York:Academic Press, pp. 249-265 (1991).

9.   J. A. Stamford, *Monitoring Neuronal Activity: A Practical Approach*, IRL Press, Oxford, pp. 113-145, 1992.

10. A. J. Bard, L. R. Faulkner, *Electrochemical Methods*, New York:John Wiley & Sons, pp. 22-26, 1980.

11. J. A. Stamford, "Fast cyclic voltammetry: measuring transmitter release in 'real-time,'" *Journal of Neuroscience Methods*, v. 34 pp. 67-72, 1990.

12. J. A. Stamford, "In Vivo Voltammetry," *Methods in Neurosciences: Electrophysiology and Microinjection*, v. 4, New York:Academic Press, pp. 127-141, 1991.

13. J. Millar, T. G. Barnett, "Basic instrumentation for fast cyclic voltammetry," *Journal of Neuroscience Methods*, v. 25, pp. 91-95, 1998.

14. J. A. Stamford, P. Palij, C. Davidson, C. M. Jorm, J. Millar, "Simultaneous 'real-time' electrochemical and electrode physiological recording in brain slices with a single carbon-fibre microelectrode," *Journal of Neuroscience Methods*, v. 50, pp. 279-290, 1993.

15. R. Belser, "Alloying Behavior of Thin Bi-metal Films, Simultaneously or Successively Deposited," *Journal of Applied Physics*, v. 31, n. 3, pp. 562-570, 1960.

ADHESION FORCE MEASUREMENTS ON SINGLE GECKO SETAE

Yiching A. Liang[1], Kellar Autumn[2], S. Tonia Hsieh[3], Wolfgang Zesch[4],
Wai Pang Chan[5], Ronald S. Fearing[6], Robert J. Full[5], Thomas W. Kenny[1]

[1]Department of Mechanical Engineering, Stanford University, Stanford, CA 94305
[2]Department of Biology, Lewis & Clark College, Portland, OR 97129
[3]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138
[4]Department of MTS, Zuehlke Engineering AG, CH-8952 Schlieren-Zurich
[5]Department of Integrative Biology, University of California at Berkeley, Berkeley, CA 94720
[6]Department of Electrical Engineering and Computer Science, U. of California at Berkeley, Berkeley, CA 94720

ABSTRACT

Geckos are known for their exceptional ability to climb up vertical walls. The adhesion mechanism has been the subject of debate for nearly a century. It is commonly assumed in the biology community that this remarkable wall-climbing ability in geckos is attributed to van der Waals forces [1]. In this theory, the tiny hair-like structures on the bottom of gecko feet, known as setae, form van der Waals bonds with the surface to facilitate adhesion. These forces are on the order of µN per seta. Various attempts have been made to experimentally quantify the adhesive forces on whole animals, but none had been able to measure the forces on single setae. Recent development in MEMS technology has produced numerous highly sensitive devices for small force measurements, including a piezoresistive cantilever capable of simultaneous force detections in two directions [2]. We have successfully performed series of adhesion measurements on single setae using this device to support the van der Waals adhesion hypothesis.

This paper describes the setup and procedures used to obtain adhesive force measurements on single gecko setae. Experimental results showed that adhesion increases with increasing contact preload and lateral pulling speed. Our data also revealed some interesting characteristics of setal adhesion. The magnitude of the adhesive forces measured supports the hypothesis that individual setae operate by van der Waals forces and sum to generate sufficient forces to explain macroscopic behavior such as clinging or vertical running.

INTRODUCTION

Tokay geckos (*Gekko gecko*) are one of the largest Gekkonid lizards (snout-vent length of approximately 37cm) and are found primarily in Southeast Asia. They have large toe pads and are remarkable in their ability to climb up vertical surfaces. Studies on whole animals have found that they are capable of sustaining over 40 times their own body weights on a nearly vertical (85° to the horizontal) acetate surface [3].

Figure 1 shows the morphology of a gecko foot. The toe pads are covered with leaf-like arrays termed lamellae. The ventral surfaces of the lamellae are densely packed with arrays of keratinous hairs called setae. On average, a gecko has approximately one million setae, each 30 - 130 µm in length and approximately 5 µm in diameter. The shorter setae are generally proximal and the longer ones distal. Each seta divides into branches before terminating in a hundred to a thousand projection tips called spatulae. Spatulae are approximately 0.2 µm in diameter [4, 5].
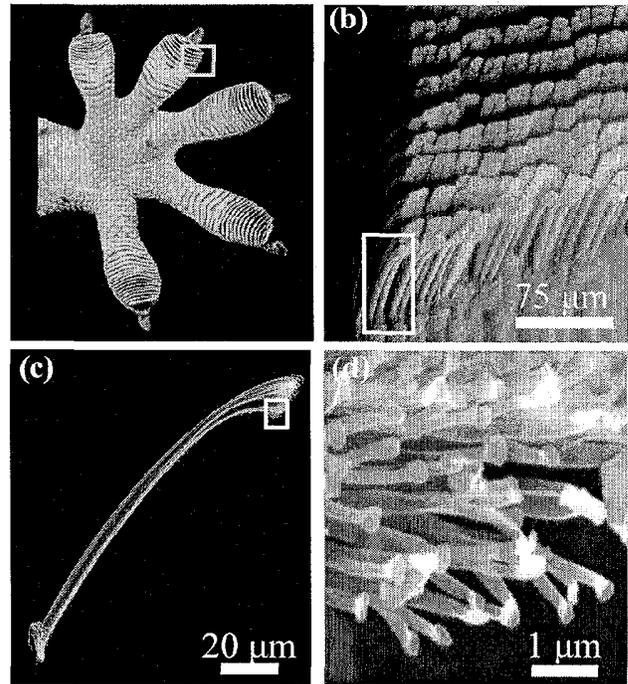
Figure 1. SEM images of (a) the foot of a Tokay gecko (photo courtesy of Mark Moffett); (b) an array of lamellae; (c) an isolated Seta; (d) spatula-shaped tips [1].

Although much attention has been given to the structure of setae, the mechanism of setal adhesion is still controversial, and has been debated for nearly a century. Table 1 lists the various theories suggested. All but the last one has been rejected based on experimental results.

Table 1. Theories suggested for the adhesion mechanism.

Theory	Reasons for rejection
Friction [6]	Adhesion is much higher than friction [1].
Suction [7, 8]	Setae can adhere in vacuum [9].
Capillary wet adhesion	Skin glands are not present on gecko feet [6, 9, 10].
Micro-interlocking [6]	Surface irregularities on the scale of the spatulae not necessary for adhesion [11].
Electrostatics [12]	X-ray bombardment does not affect adhesion; setae can function in ionized air [9]
van der Waals forces [5, 13]	(None so far.)

The adhesion by van der Waals forces hypothesis is supported by an observation that adhesive force of a whole gecko increases with increasing surface energy of the substrate [13].

Irschick [3] performed experiments to measure the macro-scale wall climbing abilities of whole animals. A live gecko with 227.1 mm^2 of pad area could sustain an average clinging force of 43.4 N. Assuming the force is divided evenly among all setae, calculated using the typical setal density of 5300 setae/mm^2 [5], this is equivalent to approximately 17 μN of force per seta. Although in reality this value can be several times higher since it is unlikely that all setae are used at the same time. Experiments of this sort provide rough measurements of the magnitude of whole-animal adhesion. They are however not sufficient for concluding what mechanism is actually responsible for the adhesion.

Theoretical calculations of the complete van der Waals force between two complex macro-scale objects are simply too complicated to perform [14], and therefore adhesion force measurements on smaller sub-units, such as individual setae, are necessary for verifying the van der Waals mechanism.

THEORY

van der Waals forces are attractive forces between two non-polar objects. They are caused by fluctuations in the instantaneous dipole moments of these two objects due to the uneven distribution of electrons in their electron clouds. Two adjacent particles tend to synchronize their dipole moment fluctuations to minimize the total potential energy; therefore van der Waals forces are usually attractive in nature. van der Waals forces are the smallest among all intermolecular forces, but they become significant when a large number of particles are involved. We can calculate the van der Waals attraction between two solids by integrating the London dispersion energy over all particles in both volumes, and differentiating with respect to the separation distance between them [15].

A single spatula adhering to a surface can be roughly modeled as the interaction between two parallel surfaces.
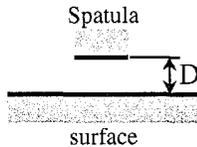
Spatula

Figure 2. Spatula and surface approximated as two parallel surfaces in contact.

van der Waals force for the configuration in Figure 2 is

$$F_{vdW} = \frac{A}{6\pi D^3} \qquad \text{per unit contact area,} \qquad (1)$$

where A is the Hamaker constant that depends on the materials of the two surfaces. This constant is typically on the order of 10^{-19} J between two solids, and does not vary significantly for different materials. A separation distance of $D = 3$Å is typically used for two solids in real contact. Since F_{vdW} is inversely proportional to the third power of D, the attractive force is negligible when the separation distance is more than tens of Angstroms. Although the exact Hamaker constant between the sensor and setae is unknown, the nominal value of 10^{-19} J can be used to obtain a reasonable

approximation. The force is calculated to be approximately 200 μN/μm^2. This corresponds to 200N of attractive force for 1 mm^2 of real contact area.

Note that there is a significant difference between real and apparent contact areas. Solid surfaces are rarely ideally planar. Therefore the real contact area is merely the total of the areas between the few opposing asperities actually in position to touch each other [16].
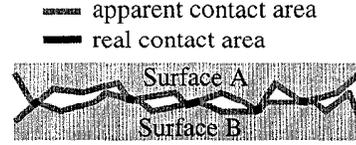
Figure 3. Real vs. apparent contact areas between two surfaces.

Since only the real contact area is considered in the calculation, van der Waals forces between most solid surfaces are negligible. But if one of the surfaces is sufficiently compliant to allow deformation under modest loads, the real contact area can be substantial. A gecko seta in contact with a hard surface is a good example of this kind of interface. The compliant setal structure can conform to the topography of a surface to form large area of very intimate contact, which increases the magnitude of van der Waals forces.

Since the tips of the spatulae are cup-shaped, only the perimeters are in real contact with the surface. For a seta with one thousand spatulae at ϕ0.2 μm, a total van der Waals force of over 200 μN results even if only 2.5% of the total tip area is in real contact with the surface. Note that this value is a very rough estimate. The actual contact area is difficult to quantify due to the irregular shapes and unknown numbers of the spatulae on each seta. It varies greatly between setae and depends also on the roughness of the surface. More accurate values can be calculated with better shape and size estimates on the specific seta sample. Clearly, larger setae would be able to provide larger sticking forces due to the larger contact area. The specific Hamaker constant that corresponds to the silicon-keratin interface can also be used to obtain more accurate force calculations.

DEVICE

All adhesive force measurements in this paper are obtained using a micromachined dual-axis piezoresistive force sensor [2] shown here with a seta in Figure 4. The device is fabricated on SOI (Silicon-On-Insulator) wafers formed using the SIMOX process.
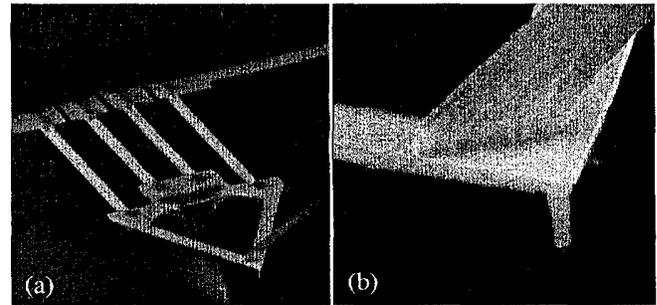
Figure 4. (a) SEM image of the force sensor with a seta. (b) Close-up at the sensor tip shows detailed structure of the spatulae.

The entire device is made of single crystalline silicon. It has two independent force sensors, each with one predominant direction of compliance. The vertical sensor consists of a thin triangular probe, and the lateral sensor consists of four tall slender ribs. A special 45° oblique ion implantation allows piezoresistive and conductive regions to be implanted on both horizontal and vertical surfaces simultaneously. Forces applied to the tip of the sensor are resolved into these two orthogonal directions, and can be measured by the changes in resistance of the piezoresistors. Since this device is originally designed for data storage applications, each cantilever contains a sharp conical tip at the apex of its triangular probe. Therefore the tipless backside of this device is used to provide a flat surface for setal adhesion. The backside of the sensor has a surface RMS roughness of less than 2.5nm, 2 orders of magnitude smaller than the size of the spatula tips on setae. Therefore it is safe to assume that micro-interlocking does not play a major role in the adhesion.

Typical noise spectra for this device are shown in Figure 5. The dotted line in each graph corresponds to the Johnson noise floor for a resistor value equal to the nominal resistance of the piezoresistor. The integrated noise level in each piezoresistor is equal to a minimum detectable force of ~5nN in a 10 kHz bandwidth, more than sufficient for measuring the μN-scale adhesion of individual setae.
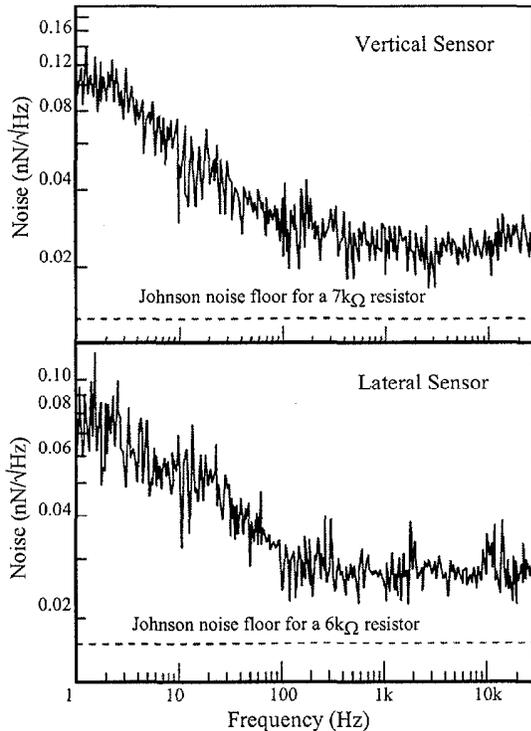
Figure 5. Noise spectra of the dual-axis force sensor.

EXPERIMENTAL SETUP

A single seta, peeled off of the toe of a ~100-g Tokay gecko, is carefully glued to the tip of a #2 insect pin using 5-minute epoxy. The pin is mounted on a computer-controlled piezoelectric manipulator (Melles Griot, NanoFlex[TM]) and orientated so that the active surface of the seta is facing down. The sensor is mounted on a movable microscope stage with the desired surface facing up. The seta is brought into contact with the sensor from

above and a small downward preload is applied to initiate adhesion. The seta is then pulled away.
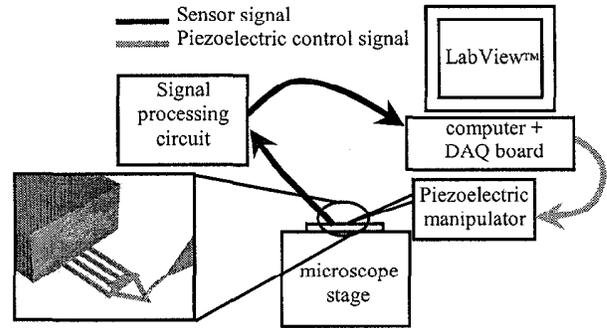
Figure 6. Experimental setup.

Sensor signals in both directions are amplified and filtered through a pair of 300-Hz low-pass filters, and digitized at 100 Hz using LabView[TM] on a PC with a 16-bit data acquisition board (National Instrument PCI-MIO-16XE-50). The collected data are then converted to sensor deflections through calibration constants, and multiplied by cantilever stiffness to obtain force data.

RESULTS

Figure 7 shows the top view of the first configuration used to study the dynamics of this adhesion-breakaway event. The seta is brought into contact with the tip area of the sensor for simultaneous vertical and lateral force measurements.
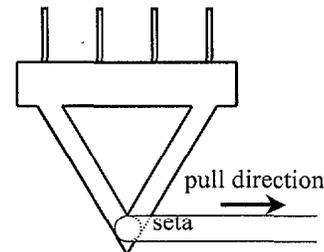
Figure 7. Experimental configuration #1.

After a small vertical preload is applied, the seta is pulled away laterally at a rate of 5 μm/sec. This pulling action is equivalent to a linearly increasing applied lateral force. Dellit [9] suggested that tension on the setae in this direction is necessary to initiate traction.

Once the pulling begins, the lateral sensor deflects with the pull until the pulling force overcomes adhesion. At this point the seta loses adhesion and breaks away from the sensor. The cantilever springs back to its initial position and the sensor signal returns to zero. The force measured at the point of breakaway is the adhesion capacity of the seta on this surface. A typical force curve in this configuration is shown in Figure 8. The slight buckling of the seta due to the preload pushes the sensor to the left and causes an initial negative lateral force of ~30 μN. In this figure, a maximum lateral force of ~45 μN (point C) is achieved with a ~5 μN preload.

Figure 8 also shows discrete sections of varying slopes on the lateral force curve. From these different sections, we can hypothesize the dynamics of the adhesion-breakaway events as

follows: Point A denotes the start of the pulling motion. As the pulling continues, lateral force builds up linearly, as evident from the somewhat constant positive slope between A and B. At B, there is a significant drop in the slope of the lateral force curve, indicating slippage between the seta and sensor. Since the active surface of the seta is comparable in size to the surface area at the sensor tip, the seta begins to slide off the edge of the sensor almost as soon as it begins to slip, therefore reducing the contact area. However, despite the reduction in contact area, lateral force continues to build up as pulling continues between B and C. This suggests that the adhesive force per area must be increasing, since the total adhesive force is the product of force per area and the contact area. At point C, most of the seta surface has lost contact with the sensor. The remaining part can no longer support the sensor deflection. At this point the seta breaks loose from the sensor. The sensor then recoils, occasionally hindered by parts of the seta still touching its top surface, before returning to its unloaded position at D.

the vertical sensor is not accessible, but the setae are able to slide on the sensor surface and build up much larger adhesion forces.

Figure 9. Experimental configuration #2.

A typical force curve obtained in this configuration, shown in Figure 10, shows a maximum adhesion force of nearly 180µN.

Figure 10. Force curve obtained using configuration #2.

Note that in Figure 10, a constant maximum force is maintained until detachment, when the seta slides off the edge of the sensor. The maximum force presumably is reached when nearly all of the spatulae are in contact with the sensor surface. Even though the vertical preload is unknown, the maximum force in Figure 10 is several times larger than can be obtained when the seta is placed at the tip of the sensor with any preload. This demonstrates that setal adhesion does build up significantly through sliding.

Preload

To study the effect of preload, the detachment force, or the adhesive forces at first slippage (point B in Figure 8), are recorded and plotted against various preload values in Figure 11. A clear positive correlation between adhesion and preload can be seen. This experiment shows that preloading improves the initial contact between the seta and the surface. To rule out friction as the dominating adhesion mechanism, we also measured the friction coefficient between the inactive non-spatular backside of the seta and the same surface on the sensor. The coefficient of friction between this silicon/keratin interface is measured to be approximately ≈ 0.2. The magnitude of frictional force is drawn as the dashed line near the bottom of the graph, and clearly indicates that friction plays a minimal role in setal adhesion.

Preload of above 15 µN causes the seta to buckle. This introduces an unnecessary complexity in the experiment. Therefore only data with preload values less than 15 µN are

Figure 8. Force curve obtained using configuration #1 and schematics showing detachment events.

If this hypothesis were true, we have just revealed an interesting possible mechanism by which geckos possess such remarkable wall-climbing ability: a fraction of the spatulae at the setal tip will adhere to the surface at initial contact. If the initial adhesion is not sufficient to sustain a gecko's body weight, the animal begins to slide down the wall, equivalent to the sliding of a single seta on the sensor surface in our experiment. It is possible that initially attached spatulae will force their unattached neighboring spatulae to become engaged in the adhesion through sliding. Thus more and more spatulae will adhere to the surface as the pulling motion continues, building up adhesion, until the force is great enough to support the animal's weight. Note that the sliding distance necessary for building up to maximum adhesion is very small, on the order of a few microns.

A second set of experiments was performed using configuration #2, shown in Figure 9, to support this hypothesis. In order to provide a larger surface for sliding, the base of the triangular probe region on the sensor is used. In this configuration

plotted. Figure 11 consists of 42 data points taken using the same seta in one sitting, therefore minimizing variations in all other parameters, and demonstrating an apparent positive relationship between preload and adhesive forces. The number of data points that can be reliably taken using one seta is limited because setae tend to dry out and lose adhesion when exposed to the strong light on the microscope for an extended period of time.
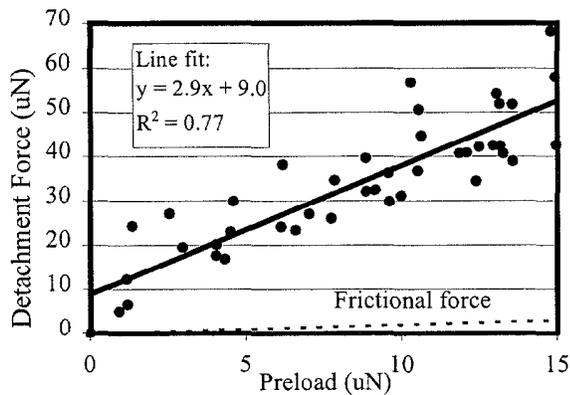
Figure 11. Detachment force as a function of preloads. (N=42)

This dependency is presumably due to the larger number of spatulae in initial contact with the surface with larger preload. The effect of preloading on setal adhesion supports the hypothesis that a small perpendicular force in concert with a rearward displacement may be necessary to engage adhesion [4]. Experiments were repeated on several other setae which all showed a positive correlation between preload and adhesion.

Pulling Speed

We also performed a series of experiments using the second configuration (Figure 9) to study the effect of pulling speed on the force buildup through sliding. The magnitude of the preload is unknown, but kept constant throughout the experiment.
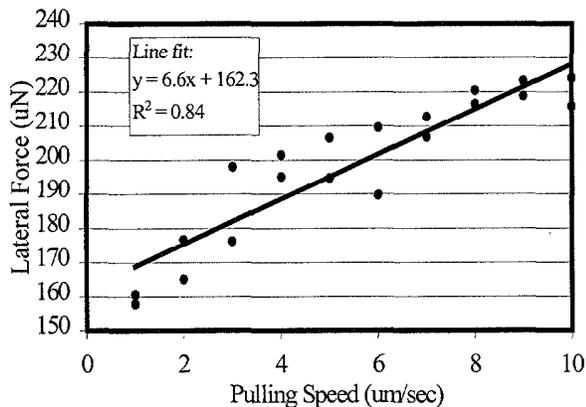
Figure 12. Maximum lateral force with sliding build-up. (N=18)

Figure 12 shows the maximum adhesive forces (dotted line in Figure 10) obtained at pulling speeds varying from 1 to 10 μm/sec. Each point represents a single data point. All 18 points are taken using the same seta in one sitting. Maximum pulling speed is limited by the response time of the piezoelectric manipulator used to pull the setae. Again a clear positive correlation can be seen. As discussed previously in the results section, initially unattached spatulae can be pulled into contact with the surface by their

neighbors through sliding. It is possible that somehow a higher pulling rate increases this dynamic effect. More experiments are necessary to support this hypothesis.

DISCUSSION

The magnitude of the lateral adhesive force measured for each seta depends greatly on the size and orientation of its active surface. In general larger forces are measured when a larger area of the setal active surface is in contact with the sensor surface. Experimental data showed that adhesion increases with increasing preload and pull speed. Most importantly, adhesion increases significantly if the seta is dragged a short distance across the surface. All these can be attributed to the increase in number of spatulae that come in contact with the surface, which supports the hypothesis that the wall-climbing ability of geckos is likely due to van der Waals forces. Furthermore, the measured contact forces are consistent with the predicted levels of van der Waals forces between these two surfaces.

Our data also showed that geckos are overbuilt in their adhesive capacity. If all setae on the gecko can provide up to ~200 μN of force, an entire animal can sustain 200 N of force with one million setae, orders of magnitude larger than the animal's total body weight. The elaborate pad structure with microscopic setae is responsible for this remarkable clinging ability. We have also found that setae can still stick to surfaces very well more than one month after they have been peeled off of the animals, as long as they are kept moist for the duration.

CONCLUSION

All of our experimental results support the van der Waals assumption, but further experiments are necessary to confirm this theory. We are in the process of fabricating an improved version of cantilevers designed specifically for this application. These new cantilevers have been redesigned and optimized for the desired force range [17], and they will also have wider tip areas for studying the effect of sliding while simultaneously measuring vertical and lateral forces. Possible future directions also include measuring vertical detachment forces, correlating the pulling angle to the maximum adhesion, as well as varying surface material and roughness by depositing various coatings on the sensor surface.

We have demonstrated the use of a novel dual-axis MEMS cantilever for high-quality measurement of dry adhesion on gecko setae. Several interesting characteristics of the setae are revealed based on data obtained using this MEMS device. In addition to addressing an important and long-standing controversy in the animal biomechanics community, this work illustrates an important application area for MEMS devices - the measurement of relatively small forces in biomechanical structures. Experiments using the newly fabricated cantilevers will allow us to perform more extensive study on the dynamics of adhesion. We also seek further opportunities for MEMS measurement capabilities to address biomechanical measurement problems.

REFERENCE

1. K. Autumn, *et al.*, "Adhesive Force of a Single Gecko Foot-Hair", accepted by *Nature* for publication (2000).

2. B. W. Chui, *et al.*, "Independent detection of vertical and lateral forces with a sidewall-implanted dual-axis piezoresistive cantilever", *Applied Physics Letters*, 72(11): 1388-1390 (1998).

3. D. J. Irschick, *et al.*, "A comparative analysis of clinging ability among pad-bearing lizards", *Biological Journal of the Linnean Society*, 59(1): 21-35 (1996).

4. R. Ruibal and V. Ernst, "The Structure of Digital Setae of Lizards", *Journal of Morphology*, 117: 271-294 (1965).

5. A. P. Russel, "A Contribution to the Functional Analysis of the Foot of the Tokay, *Gekko-Gecko* (Reptilia, Gekkonidae)", *Journal of Zoology London*, 176(4): 437-476 (1975).

6. B. C. Madhendra, "Contribution to the Bionomics, Anatomy, Reproduction, and Development of the Indian House-Gecko *Hemidactylus flaviviridis* Ruppell. Part II. The problem of locomotion", *Proc. Indian Acad. Sci.*, 3: 288-306 (1941).

7. H. Gadow, *The Cambridge Natural History Vol. 8: Amphibia and Reptiles*, London, McMillan and Co. (1901).

8. J. F. Gennaro, "The Gecko Grip", *Natural History*, 78: 36-43 (1969).

9. W. F. Dellit, "Zur Anatomie und Physiologie der Geckozehe", *Jena. Z. Naturw*, 68: 613-656 (1934).

10. A. Bellairs, *The Life of Reptiles*, New York, Universe Books (1970).

11. P. F. A. Maderson, "Keratinized Epidermal Derivatives as an Aid to Climbing in Gekkonid Lizards", *Nature*, 203: 780-781 (1964).

12. H. R. Schmidt, "Zur Anatomie und Physiologie der Geckopfote", *Jena Z. Naturw*, 39: 551 (1904).

13. U. Hiller, "Untersuchung zum Feinbau und zur Funktion der Haftborsten von Reptilien", *Zeitschrift fur Morphologie der Tiere*, 62(307-362): (1969).

14. J. T. Feddema, *et al.*, "Micro-Assembly Planning with van der Waals Force", *Proceedings of the IEEE International Symposium on Assembly and Task Planning*, Porto, Portugal (1999), pp. 32-38.

15. J. N. Israelachvili, *Intermolecular And Surface Forces, 2nd Edition*, London, Academic Press (1992).

16. R. E. Baier, *et al.*, "Adhesion: Mechanisms That Assist or Impede It", *Science*, *162*: 1360-1368 (1968).

17. Y. Liang and T. Kenny, "Mechanical Optimization of a Dual-Axis Piezoresistive Force Sensor", *Proceedings of the ASME International Congress and Exposition MEMS Symposium*, Anaheim (1998).

REMOVAL OF CUMULUS CELLS FROM MAMMALIAN OOCYTES IN A MICROFLUIDIC SYSTEM

H. C. Zeringue and D. J. Beebe
Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801
Department of Biomedical Engineering, University of Wisconsin-Madison, Madison, WI 53706
M. B. Wheeler
Department of Animal Science, University of Illinois at Urbana-Champaign, Urbana, IL 61801

ABSTRACT

We have designed, fabricated, and tested a micro fluidic system for the removal of cumulus cells from unfertilized mammalian oocytes. Building devices in poly-dimethylsiloxane (PDMS) with channel dimensions on the same order as that of the oocyte-cumulus complex (~400 μm) allows for the tracking and fluidic manipulation of individual complexes. The system uses pressure driven flows to achieve positioning and movement of the oocyte complex leading to the complete separation and removal of the cumulus cells. A series of narrowed regions are used to "condition" the cumulus. The complex is then brought to ports in the channel that mechanically remove the cumulus. This technique could be combined with other assisted reproductive processes that are amenable to automation in microfluidic systems thereby reducing labor costs and potentially improving *in vitro* production efficiencies.

INTRODUCTION

The use of *in vitro* embryo manipulation is expanding as new applications emerge which include: increasing genetic improvement of livestock breeding herds; producing transgenic animals that generate pharmaceutical products in their milk; producing chimeric animals for cancer and genetic disease research; and enabling infertile couples to become pregnant. One of the key steps in many of these processes is *in vitro* fertilization (IVF). The oocyte (unfertilized egg) is surrounded by the granulosa cell mass, called the cumulus oophorus. The cumulus mass has a consistency similar to honey and is approximately spherical with a diameter between 200 and 500 μm for most mammalian species (including mice, swine, cattle, and humans). The cumulus cells are a barrier to sperm penetration. For this reason, the cumulus mass is often removed prior to IVF and embryo maturation [1]. A typical procedure for cumulus removal includes high-speed vortexing of embryos in a centrifuge tube. This causes high stresses on the oocyte complex and is done on large numbers of oocytes simultaneously. Oocyte to oocyte contact facilitates the cumulus removal [2, 3].

The ability to handle and track individual embryos is very important. Eggs removed from the donors are typically handled in batches to save costs. However, it is advantageous in some cases to keep track of each egg individually[4] from the time it is removed from the donor, through the transfer to the recipient, until the birth of a healthy offspring.

Microfluidic devices have been used for cell manipulation [5-7]. Glasgow et al [8] have shown precise manipulation and retention of single embryos in microfluidic systems. The size scale of the channel environment and laminar flow characteristics of microfluidic systems make them well suited to the automation of embryological procedures. Being able to position and retain an embryo during testing or culture allows the individual identity of each embryo to be maintained. These advantages are employed in the present research to achieve microfluidic cumulus removal.

DEVICE DESIGN, FABRICATION, AND METHODS

The microfluidic device for removing the cumulus mass (Figure 1) has the following key features: 1) non-toxicity; 2) properly sized cross-sectional geometry; 3) simple loading scheme; 4) "conditioning regions" and ports for mechanical stripping of the cumulus; and 5) connections that allow precise control of the fluid.

The device reported here (Figure 1) is fabricated with glass and PDMS using recently developed techniques[9]. Chan et al [10] have shown that many typical microfabrication materials,
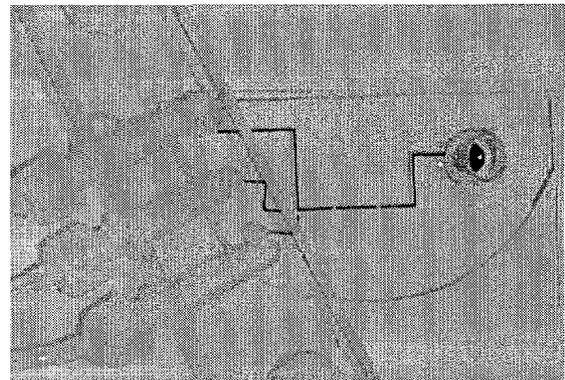
Figure 1: Picture of a device with the inlet well (right) and outlet connections (left).

including PDMS, are not toxic to embryos. The main channels in the devices (Figures 1 and 2) are sized to allow a single oocyte-cumulus complex to travel freely. These channels are 400μm deep by 500μm wide.

A polypropylene well (Figure 1) is bonded to the loading port to provide a larger fluid reservoir at the inlet. An acrylic syringe connection module (Figure 3) is clamped over the exit ports so that standard syringes or other fittings can be connected to the device. Syringes enable manual pressure control, or a syringe

pump can serve as a precise flow controller. Similar to previous glass/silicon devices[8], the PDMS devices allow for embryo positioning throughout a channel network and parking of the oocyte at desired locations during testing. In addition, the PDMS devices provide complete optical access (important for embryo analysis), rapid prototyping, and easy integration with future analysis sensors.

Loading oocytes in the device is simplified through the use of a funnel-shaped inlet well (Figure 2b). The funnel shape is molded at the entrance to the channel with the tip of the funnel connected to the head of the channel. This configuration allows an oocyte complex to be easily inserted
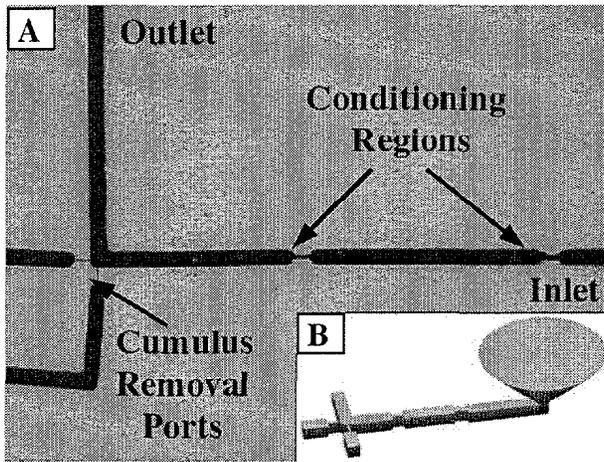
Figure 2: (a) Picture of an actual device with parts labeled. (b) An isometric view of the channel design showing the funnel shaped inlet well.

in the wide funnel opening using standard embryo handling methods. Oocytes will typically sink to the funnel bottom. The sloped walls guide the complex into the channel entrance at the bottom of the funnel. This method of loading simplifies the handling procedures because it does not require precise lateral positioning .

To manipulate the cumulus cells into a configuration that allows for complete cumulus removal, the complex is passed through two narrow channel regions, termed "conditioning regions" (Figure 2a). These narrowed regions
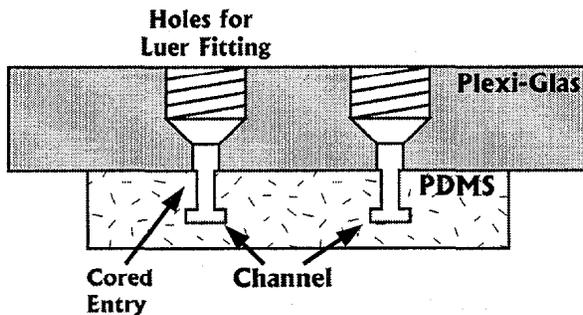
Figure 3: Cross-sectional schematic view of the connection module on the PDMS device.

force the cumulus into two main clumps at the front and back of the oocyte (Figure 4a). The two conditioning regions (Figure 2a) are 200 and 150 μm wide, respectively.

From the conditioning regions, the oocyte complex flows to the removal ports (Figure 4b). The removal ports are two thin
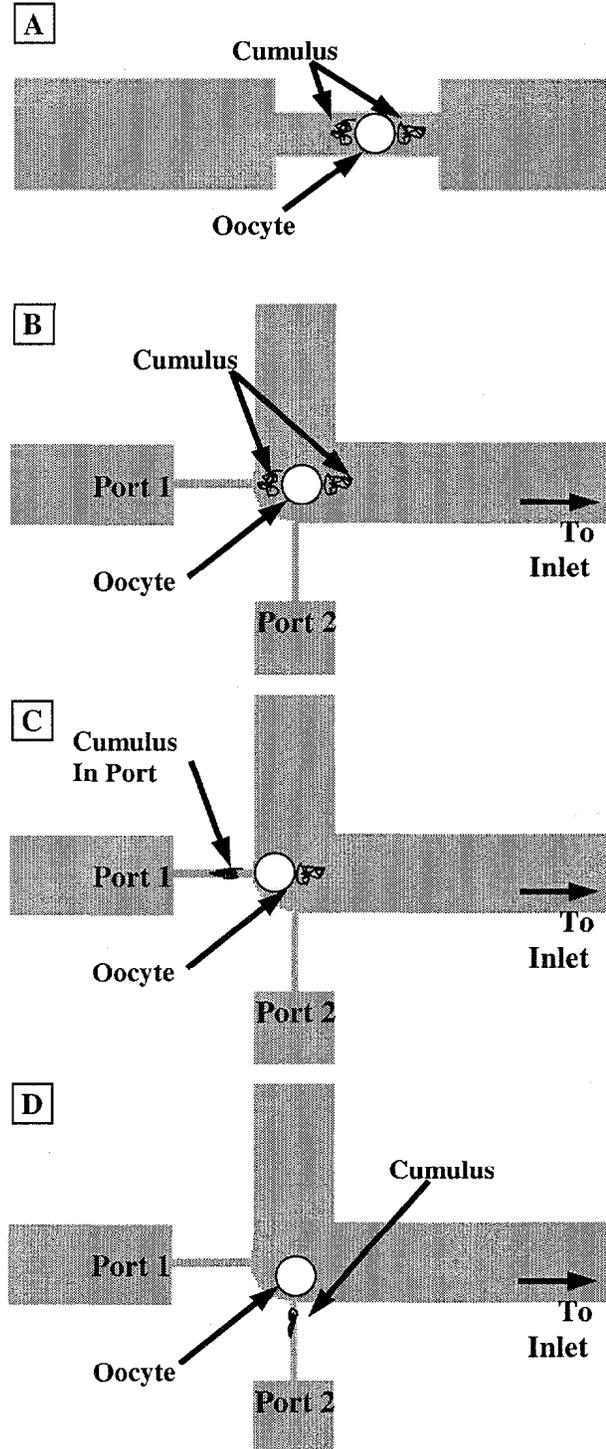
Figure 4: Diagram of the cumulus removal process. The oocyte-cumulus complex is brought to port 1 after passing through the conditioning regions (a). The first bundle of cumulus is removed at port 1 (b). The complex is moved to port 2 where the second bundle of cumulus is removed leaving the clean oocyte (c).
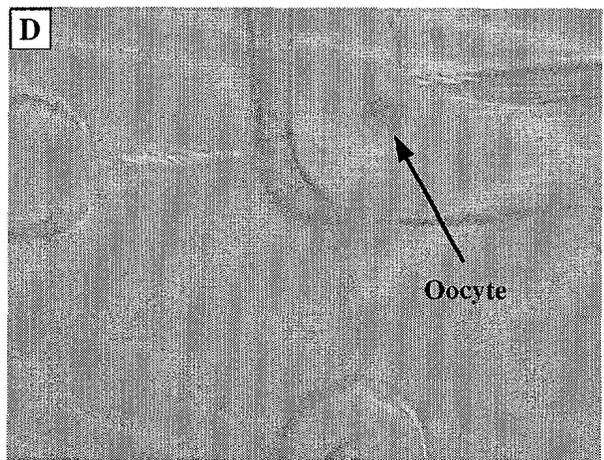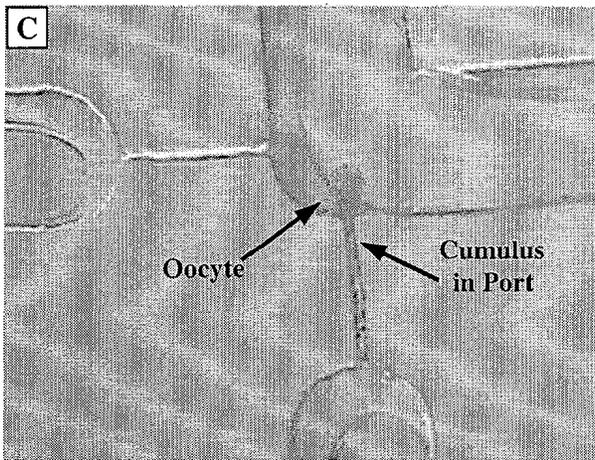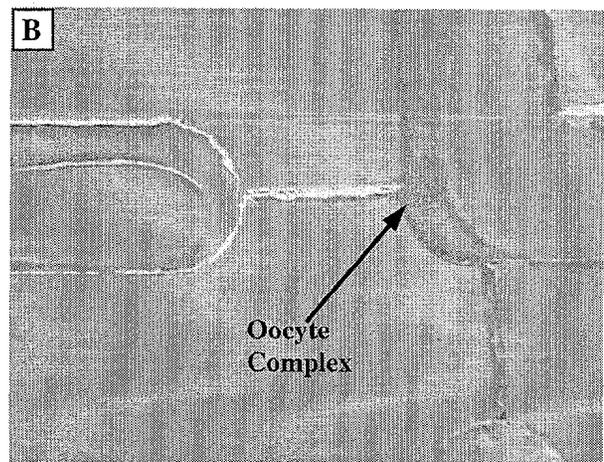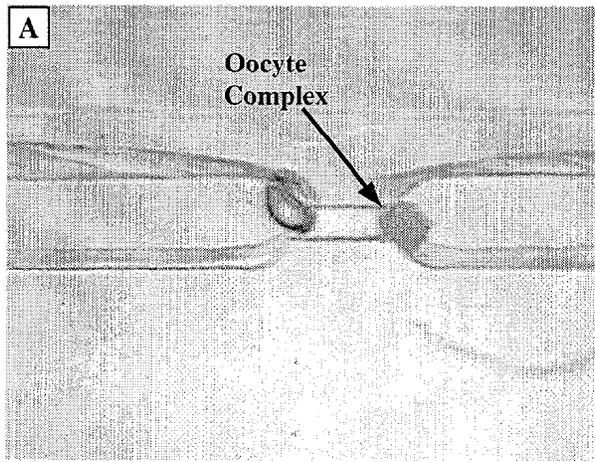
Figure 5: The oocyte-cumulus complex entering the first conditioning region (a). The same complex at the cumulus removal port 1 (b) and again at port 2 (c). The clean oocyte in the elbow region (d). The variations of the channel walls are due to undercutting of the epon master during device fabrication. This undercut was transferred to the PDMS during the molding process.

channels placed 90 degrees from one another at a bend in the channel (Figures 2 and 4). The ports allow the cumulus to enter while being too small for the oocyte (Figures 4c and 4d). Using fluid flow, the cumulus can be stripped from the oocyte and carried out of the embryo environment via the port (Figure 4c). The ports are placed at 90 degrees to allow controlled removal of the main cumulus masses. After the first mass has been removed, the oocyte is pulled fluidically to the second port, where it aligns to facilitate removal of the second mass (Figure 4d).

RESULTS AND DISCUSSION

Cumulus removal from bovine oocytes has been achieved in a microfluidic system. The cumulus removal process was performed under a stereomicroscope. Initial testing was done using presumptive zygotes from matured bovine oocytes with the cumulus cells intact (oocyte-cumulus complex). Using syringes connected by tubing to the outlet ports (Figure 1), the network was flushed with M16 (Eagel's modified media) then flushed with fresh M16 to remove bubbles from the channels. Using fine-bore glass pipettes (a standard embryology tool) an oocyte-cumulus complex was placed in the entrance funnel and allowed to sink to the bottom (20-30 sec). The fluid was then drawn into the channel using the syringes connected to the outlet ports. This caused the complex to move down the channel through the two conditioning regions (Figure 5a). When the complex cleared the second conditioning region, the syringes were adjusted to direct flow to outlet port 1. The fluid carried the oocyte such that the cumulus entered removal port 1 ahead of the oocyte (Figure 5b). By drawing on the syringe, flow was increased until the cumulus in the port was removed. The exit flow was then switched to exit from outlet port 2. The complex aligned itself such that the remaining mass of cumulus entered the port (Figure 5c). Again, flow was increased until the cumulus was removed. This left a stripped oocyte at the elbow region (Figure 5d). By reversing the flow in the device, the oocyte was transported back through the channel and back into the funnel entrance region. It could then be easily removed via pipetting techniques.

The cumulus mass normally surrounds the oocyte in an arbitrary manner. Initial testing revealed that forcing the complex through a narrow channel actively grouped the cumulus mass in a non-uniform pattern. The conditioning regions act to reconfigure the cumulus cells into two main bundles. As the oocyte complex passes through the conditioning regions that are the same size as the oocyte, the cumulus cells are forced to the front and rear. The stiffness of the oocyte allows it to hold it's shape while passing through these regions while the cumulus will conform to the

surrounding spatial restrictions (Figure 4a). The bundles of cumulus cells can then be easily removed via fluid flow. When the complex is brought to a region where the channel is too small for the oocyte to fit, the cumulus will move into this region and can then be sheared from the oocyte using fluid flow (Figures 4c and 4d). The device presented here utilizes this phenomenon to achieve controlled removal of the cumulus mass.

CONCLUSION

In this paper, we present a microfluidic platform for cumulus removal from mammalian oocytes. Using pressure driven flow in micro channels allows for the removal of the cumulus from single oocytes. Precise control of oocyte placement and flow rates allow for controlled pressures yielding predictable cumulus removal. This work illustrates the potential impact micro technologies will have on the embryology industry. The ability to precisely control embryo position and mechanically manipulate the local environment are also important enablers of basic developmental biology studies. Future work will include fertilization. Viability tests to ensure that embryo health is not adversely affected by these procedures are planned.

ACKNOWLEDGEMENTS

REFERENCES

1. Bavister, B.D., "Analysis of Culture Media for In Vitro Fertilization and Criteria for Success", in *Fertilization and Embryonic Development In Vitro*, J. L. Mastroianni and J.D. Biggers, Editors. 1981, Plenum Press: New York. p. 41-60.

2. Rutledge, J. and R. Munson, Personal Communication, Department of Animal Sciences, University of Wisconsin-Madison. February 2000.

3. Wang, L., and David J. Beebe, Allen R. Williams, Kim D. Easley, "Electrothermal Branding for Embryo Labeling",. *IEEE Transactions on Biomedical Engineering*, 1997.

4. Li, P.C.H. and D.J. Harrison, "Transport, Manipulation, and Reaction of Biological Cells On-Chip Using Electrokinetic Effects", *Analytical Chemistry*, 1997. **69**(8): p. 1564-1568.

5. Austin, R.H., R.H. Carlson, S.S. Chan, J. Knight, S. Turner, and H. Craighead. "Micro and Nanopore Structures for Biological Applications", in *1997 International Conference on Solid-State Sensors and Actuators*. 1997. Chicago, IL.

6. Glasgow, I., H.C. Zeringue, D.J. Beebe, S.-J. Choi, J.T. Lyman, and M.B. Wheeler. "Individual Embryo Transport and Retention on a Chip" in *μTAS '98*. 1998. Banff, Canada.

7. Duffy, D., J.C. McDoanld, O.J.A. Schueller, and G. Whitesides, "Rapid Prototyping of Microfluidic Systems in Poly (dimethyl siloxane)", *Angewandte Chemie. International edition in English*, 1998. **37**: p. 550-575.

8. Chan, N.G., J.T. Lyman, S.-J. Choi, H.C. Zeringue, I.K. Glasgow, D.J. Beebe, and M.B. Wheeler, "Development of an Embryo Transport and Analysis System: Material Biocompatability". *Theriogenology*, 1999. **51**(1): p. 234.

DEMONSTRATION OF A MICROFABRICATED HIGH-SPEED TURBINE SUPPORTED ON GAS BEARINGS

Luc G. Fréchette, Stuart A. Jacobson, Kenneth S. Breuer[1], Fredric F. Ehrich, Reza Ghodssi[2],
Ravi Khanna, Chee Wei Wong, Xin Zhang, Martin A. Schmidt and Alan H. Epstein

Gas Turbine Laboratory and Microsystems Technology Laboratories
Massachusetts Institute of Technology
Cambridge, MA 02139

ABSTRACT

A single-crystal silicon air turbine supported on gas lubricated bearings has been operated in a controlled and sustained manner at rotational speeds greater then 1 million rpm and power levels approaching 5 W. The device is a second-generation version of the microbearing rig first reported by Lin et al. [1], and is the first micromachine to operate at circumferential tip speeds of hundreds of meters per second, comparable to conventional scale turbomachinery. To achieve this level of peripheral speed, microfabricated rotors must withstand large induced stresses, need a sufficient power source to drive them, and require stable, low friction bearings for support. This paper focuses on process improvements in the microbearing device fabrication and on the advances in operating micro-gas bearings and micro-turbomachinery. The successful operation of this device motivates the use of this technology for high-power density MEMS.

INTRODUCTION

To achieve high power and efficiency from a rotating device, high circumferential tip speed is a necessity. Conventional scale turbomachinery typically run with tip speeds of order 500 m/s, enabling high-power density applications such as gas turbines for aircraft propulsion and power generation. In order to achieve high levels of power density, microfabricated rotors will need to run at comparable tip speeds. Typical rotating micromachines, such as gears and micromotors, are formed either by surface micromachining or LIGA, supported by solid contact on a pin bearing, and entrained by electrical or contact forces acting on the edges of the rotor. These micro-rotors have reached of order 2 m/s tip speed [2], which is two orders of magnitude lower than desired for Power MEMS applications.

An effort was undertaken at MIT to develop high-speed rotating devices to enable high-power density MEMS [3]. In the current approach, the rotor is etched from the bulk substrate, taking advantage of the high strength-to-density ratio of single crystal silicon. Fluid film lubrication supports the rotor, reducing the resistance to rotation and preventing wear due to solid contact. Micro-turbomachinery is used as a drive source for high-power density fluid-to-mechanical energy conversion. In order to demonstrate this approach and develop the requisite technology, a microturbine-driven bearing rig was developed. This paper presents the design, fabrication, and operational achievements since the previously reported results of Lin et al. [1].

DEVICE DESIGN AND DESCRIPTION

The microbearing device, shown in Figures 1 and 2, consists of a 4.2 mm diameter rotor enclosed in a fusion-bonded stack of five aligned, through-etched wafers. The rotor is a planar disk with radial turbine blades on its front side. Pressurized air enters the device near the outer edge and flows radially inward, first through a set of stationary (stator) vanes and then through the rotor blades, exhausting axially near the center of the device. A gas bearing on the periphery of the rotor supports radial motion, while a pair of gas thrust bearings on the device centerline supports axial motion. The stator vanes, shown in Figure 3, are curved, turning the flow away from purely radial, imparting angular momentum to the flow as it accelerates. The turbine rotor blades turn the flow back toward radial, extracting angular momentum in the process. The angular momentum change across the rotor produces a torque. The power transferred to the rotor is the product of this torque and the rotor's angular rotation rate. For this application, the turbine was designed to provide sufficient power to overcome the viscous drag in the bearings and on the back side of the rotor, which goes like the square of the rotation rate. The total drag is approximately 13 W at a circumferential tip speed of 500 m/s. While viscous drag is relatively large in microsystems due to the small length scale [4], it is still quite small compared to the capabilities of high-speed turbomachinery. The turbine for this device had to be intentionally designed to match the relatively low power requirements of the viscous drag. Alternative turbine designs, compatible with the current process and geometric constraints, that produce tens of watts of power (beyond the drag requirements) have been designed for Power MEMS applications.
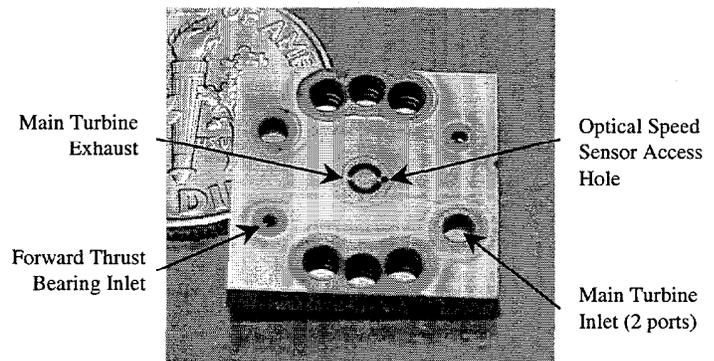
Figure 1. *Microturbine-driven bearing rig die, consisting of a diced five-wafer bonded stack, which encloses a 4.2 mm rotor.*
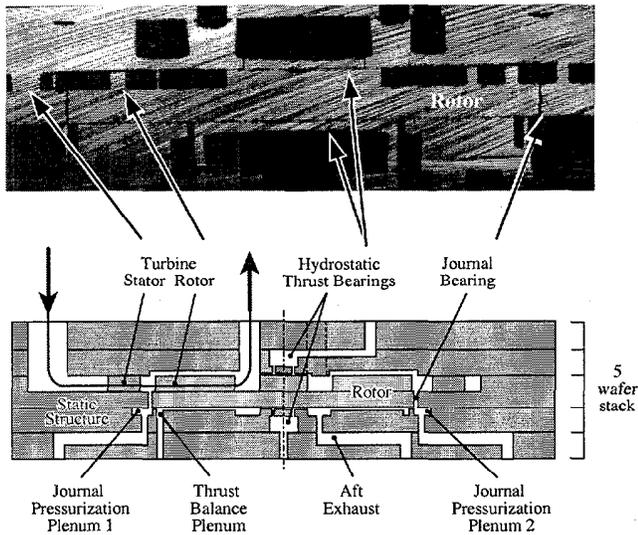
Figure 2. *Optical photograph (top) and schematic (bottom) of a cross-section of the microturbine-driven bearing rig.*

The turbomachinery was designed using MISES, a two-dimensional computational fluid dynamics code developed at MIT [5]. MISES is an Euler solver coupled to an integral boundary layer formulation. The actual flow through the turbine is expected to be three-dimensional, mainly due to boundary layer growth on the end-walls. However, because the turbine flow is accelerating, the three-dimensional effects are not expected to be substantial [6]. The particular challenge in the design of microturbomachinery for MEMS results from the microfabrication constraint that limits us to constant blade height. A more traditional design for centrifugal turbomachinery would allow the blade height to vary inversely with radius to compensate for the increase in circumference with radius. The resulting blade designs reflect the different strategies that a designer must use to adjust for the differences in fabrication capability.

The hydrostatic gas thrust bearings were demonstrated in the first generation of this device [1]. The hydrostatic gas journal bearing, which supports radial motion, is new to this device. The journal bearing, formed in the gap between the rotor and the stator at the periphery of the rotor, acts as a radial spring. The journal bearing is 300 µm deep and has an average width of 15 µm with the rotor centered. A pressure differential is maintained across the journal bearing, generating an axial through flow. On the front side of the rotor, the journal bearing boundary pressure is set by the stator exit/rotor inlet pressure, called the interrow pressure. On the back side of the rotor, the journal bearing opens into a large plenum whose pressure can be set externally. The current implementation splits the back side plenum into two symmetric journal pressurization plenums, as shown in Figure 2, although for current testing these two plenums are typically set to the same pressure.

The pressure supplied to the journal pressurization plenums is maintained higher than the interrow pressure so that there is an axial flow across the journal bearing from the back side to the front side. As the flow goes through the journal bearing gap, there is a pressure drop associated with entering the bearing, with the development of the boundary layer in the gap, with flow through the bearing from viscous effects, and finally with exiting the bearing. Each of these losses can be viewed as a resistance, and several of these resistances are functions of the local journal bearing width. When the rotor moves off center so that the journal
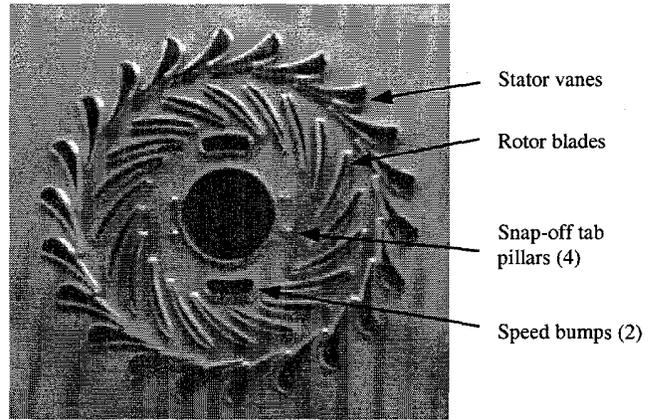
Figure 3. *SEM of the 4.2 mm diameter microturbine showing the 150 µm tall stator and rotor blades, two symmetric speed bumps, and four pillars for the snap-off tabs. In this picture, the journal bearing gap remains to be etched.*

bearing width is smaller on one side than the other, the resistances on each side of the rotor result in axial pressure profiles through the journal bearing with differing distributions. When the pressure is integrated across and around the journal bearing, the resultant force on the rotor is a restoring one, in the direction to re-center the rotor. This hydrostatic bearing scheme was first demonstrated by Orr on a scaled-up bearing rig of similar geometry [7].

In addition to the two journal pressurization plenums for the journal bearing, there is a circumferential plenum on the back side of the device, referred to in Figure 2 as the thrust balance plenum. While there is no stiffness associated with this plenum, it is used to help balance axial thrust on the rotor. The thrust balance plenum can be coupled to pressures elsewhere in the system, allowing the operator to compensate for axial loads on the rotor that develop as the pressure across the turbine is increased. In this way, the thrust balance plenum complements the thrust bearings.

Speed is measured with a fiber optic sensor that detects the passage of two diametrically opposite features on the silicon rotor. These two features, referred to as speed bumps in Figure 3, are at the same level as the blades and the thrust bearing hub, and are located on the rotor in the exhaust flow from the turbine. A fiber optic displacement measurement system is located above the speed bumps. With the rotor spinning, the output from this sensor approximates a low duty cycle square wave with frequency twice that of the rotor rotation rate.

FABRICATION

The fabrication process is based on deep reactive ion etching (DRIE) and aligned fusion-bonding of five silicon wafers to create fluidic interconnects and enclose the rotor. An exploded view of the five-wafer stack device is shown in Figure 4. The process flow was first presented by Lin *et al.* [1] and remains the foundation for this second generation device. This section summarizes the fabrication sequence, describes photolithography development necessary for precision high-aspect ratio DRIE, and presents a new rotor release strategy.

The end plates (wafers 2 and 4) first receive shallow silicon etches on their side facing the rotor. All the wafers are then deep etched halfway through their thickness, except for wafer 1, which

Main Air
Exhaust Inlet

#1
Forward
Foundation
Plate

#2
Forward
End
Plate

#3
Rotor
Plate

#4
Aft End
Plate

#5
Aft
Foundation
Plate

Top View   Bottom View
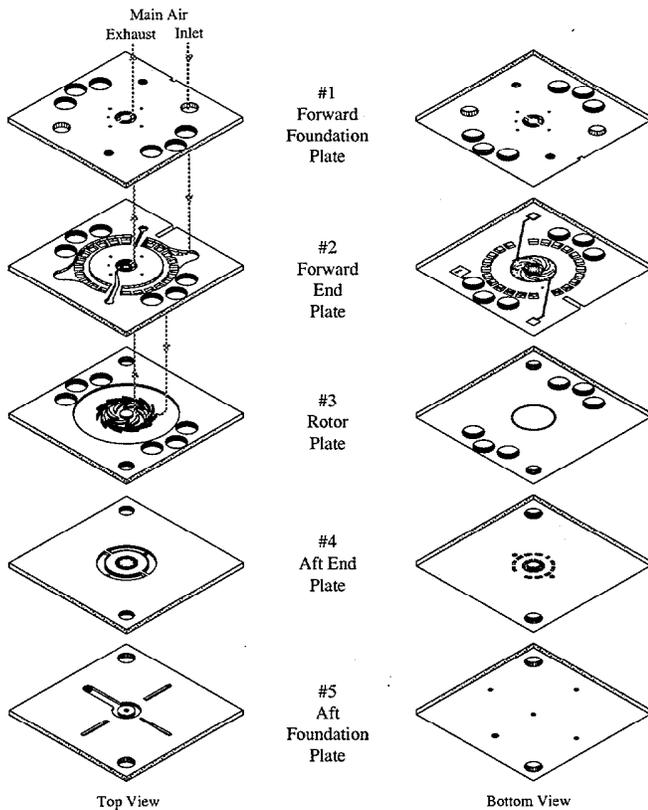
Figure 4. *Top and bottom view of the five-stack device. The die is 15 mm per side and 2.3 mm thick.*

Turbine Exhaust

Tab Beam

Forward End Plate
Wafer 2

(a)

Bonding Interface

Tab Pillar

Rotor Plate - Wafer 3

(b)

Figure 5. *(a) Schematic of snap-off tab configuration: (b) SEM of bonded snap-off tab structures before and after fracture.*

is etched through from its bottom side. Wafers 2, 4 and 5 are then deep etched on the opposite side until through-flow channels are created. Wafer 3 is then bonded to wafer 2 and etched from the bottom side to define the journal bearing gap. Finally, wafers 1, 4, and 5 are fusion-bonded to wafers 2 and 3 in a single step.

During through etching, the wafers are reversibly mounted to a quartz substrate to prevent leakage of He coolant when the etch breaks through. The bonding surface on the back side of wafers are protected to prevent deposition of hydrocarbon residue from the DRIE passivation step [8] and damage from ion back-scattering off of the quartz handle wafer. For the die shown in Figure 1, the back of wafer 1 was not protected during its through etch from the bottom side, and the resulting surface contamination can be seen around the holes.

The required level of accuracy for the micromachined geometries is driven by their functionality. The two outermost wafers provide external fluidic and optical ports to the die, resulting in dimensions of the order of hundreds of microns without stringent precision requirements. Moving inwards toward the rotor, the end plates (wafers 2 and 4) feature pressurized plenums (on their outer side) and 10 μm diameter by 100 μm long orifices for the thrust bearings (on their side facing the rotor). Accurate dimensions are critical since the axial stiffness of the thrust bearing depends on the pressure drop across the orifice, which is a strong function of its diameter. Orifices with a diameter of 10-11 μm were reproducibly achieved by exposing a 7 μm diameter circle through a 6 μm thick resist (AZ4620). The optimized process resulted in an 8-9 μm diameter opening in the resist, and 1 μm of blowout in the silicon after 90 minutes of DRIE. The rotor is cut from the back of
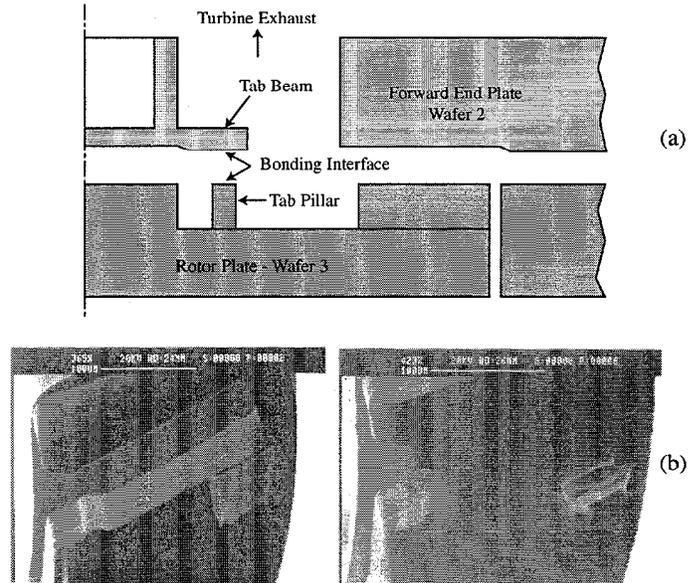
the center wafer by a high-aspect ratio circumferential trench, which simultaneously creates the journal bearing geometry. From an initial 4 μm line width in the mask, a final DRIE trench tapering from 17 to 9 μm wide and 300 μm deep was achieved using a 10 μm thick photoresist. In both processes, the minimum feature size consistently achievable was limited by the trade-off between the photoresist thickness necessary to withstand the etch duration, the increase in etch duration for smaller features, and the lower selectivity of DRIE recipes with less lateral etch.

An on-going challenge in the fabrication of micro-rotors is a reliable process flow to create free parts within enclosed structures, without tedious manual assembly and risk of damage. The approach adopted consists of bonding wafers 2 and 3 before the journal is etched, creating silicon tabs which connects the rotor to the adjacent wafer, similar to the approach presented by Lin et al. [1]. As shown in Figure 5(a), a silicon pillar extending from the rotor bonds to a horizontal beam on the forward end plate. The journal bearing is then etched and the remaining three wafers are aligned fusion bonded to complete the five-stack. The silicon links are then mechanically fractured to release the rotor before testing, acting as *snap-off tabs*. This differs from Lin et al., who used a laser assisted etch to remove the tabs. The tab geometry was designed to provide a solid support against axial and in-plane motion of the rotor, while allowing a controlled failure mode upon lateral force. An array of configurations were designed using simple beam theory and stress concentration relations, and tested aiming for a clean fracture at the roots of the pillar and beam. Micromachined silicon needles were specially designed to reach down the main exhaust port and apply the necessary lateral force to snap the tabs. Figure 5(b) shows such a snap-off tab structure before and after being fractured. The fractured surfaces are raised from the disk surface, which is important so as not to introduce stress concentrations. These snap-off tabs have been successfully used in two five-stack builds. Particles generated during snap-off were removed by blowing nitrogen through the main turbine flow path, out the main exhaust.

45

EXPERIMENTAL DETAILS

The microturbine-driven bearing rig die mounts in an acrylic package using o-rings to create sealed fluidic connections to the die. Plastic tubing connects the package to a gas handling system, which includes a high-pressure nitrogen source, valves to distribute pressures to the various ports, pressure sensors, mass flow meters, a PC-based data acquisition system, and a fiber optic speed sensor.

During operation, the signal from the speed sensor is monitored on a spectrum analyzer. The spectrum analyzer shows a strong peak at twice the rotation rate (two speed bumps per rotation), at the rotation rate due to asymmetry, and also at higher harmonics. Under certain conditions, extra peaks also appear, as shown in Figure 6. These peaks are evidence of excitation of the bearing natural frequency. As mentioned earlier, the hydrostatic journal bearing acts like a spring, and as such has a natural frequency associated with it. Whirling of the rotor at the natural frequency of the bearing results in precession of the speed bumps, which appears as a pulse width modulation of the signal at the whirl frequency. The measured interaction between the rotation rate and the bearing natural frequency also results in peaks at their combinations, as identified in Figure 6. Operation has shown that appearance of these extra peaks is associated with the approach of a
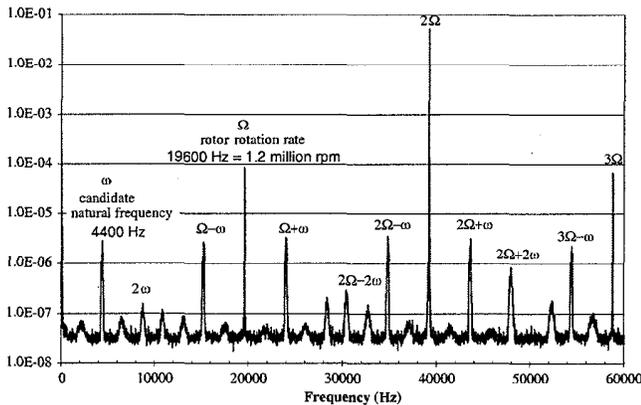
Figure 6. *Power spectrum of the optical speed sensor signal showing the speed bump frequency (2Ω) and the rotor rotation rate (Ω), as well as other peaks corresponding to the natural frequency (ω) and its combinations with the rotation frequency.*
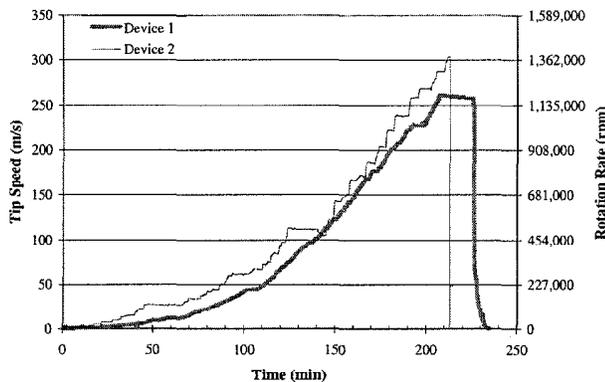
stability boundary. The differential pressure across the bearing can usually be adjusted to stabilize the rotor and make these peaks vanish. The ability of the speed sensor to pick up this additional information has been an important diagnostic tool.

Figure 7 shows the rotation rate of two devices as they are slowly accelerated in a stable manner. As the turbine supply pressure increases, the pressures applied to the hydrostatic journal bearing are also increased according to a predetermined schedule. Device 1 was brought up to a tip speed of 260 m/s (1.2 million rpm) and was held there for 20 minutes (a pressure leak resulted in a slight deceleration). The rotor was then rapidly decelerated to stop. Device 2 reached a tip speed of 303 m/s (1.4 million rpm) prior to going unstable and crashing. Both devices achieved of order 10^8 revolutions. The operating schedule is one suspect in the cause of the crash of Device 2. More dies are currently being tested to help map out safe areas of operation. Four other dies from this build have already been tested, but they have only been able to reach tens of thousands of rpm prior to instability. The main problem limiting high-speed operation is believed to be rotor imbalance. Balance is limited by our ability to align masks on the front and back side of the rotor wafer, and by the uniformity of the blade etches across a rotor. When the bearing operates at rotation rates higher then the journal bearing natural frequency (supercritically), as it does through most of its operating range, the rotor spins around its center of mass rather than its geometric center. This offset of the rotor produces circumferential pumping of fluid, causing the development of hydrodynamic forces in the journal bearing gap. These hydrodynamic forces are relatively small if the imbalance is small, but they increase with the square of speed. The hydrodynamic forces tend to be destabilizing [9].

Figure 8 shows the turbine operating line for Devices 1 and 2. These operating lines are comparable in shape to those observed for macroscale turbines. Also shown in this figure are three data points calculated using MISES with the corresponding power delivered by the microturbine. At 1.4 million rpm, the microturbine is predicted to deliver nearly 5 W of mechanical power with a torque of over 30 μN-m. The computational results are dependent on the axial position of the rotor, which affects the amount of drag. The uncertainty bars on the MISES data points are associated with an expected +/- 0.5 μm uncertainty in the axial position of the rotor. The MISES results do not compensate for the decrease in mass flow due to blockage from boundary layer growth on the end-walls

Figure 7. *Speed evolution over time. Uncertainty in the tip speed measurement is ±0.5% of the measured speed.*
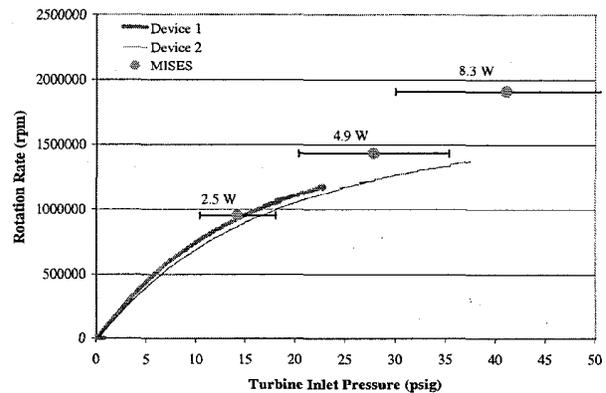
Figure 8. *- Rotation rate as a function of the air pressure supplied to the turbine. There is a ±0.8 psi uncertainty in the pressure.*
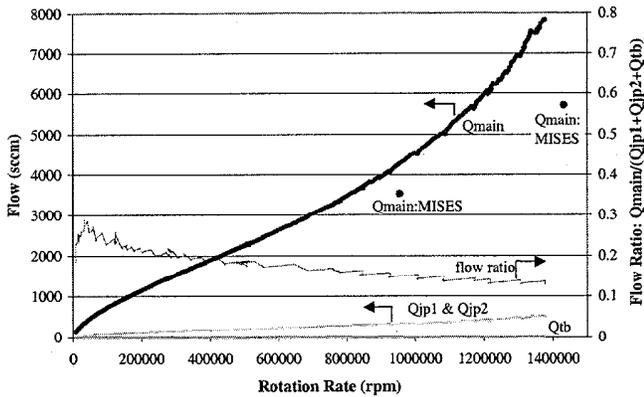
Figure 9. Turbine and bearing flow rates over the range of operating speeds for Device 2. Uncertainty in the main flow is ±200 sccm, in the journal pressurization plenums (jp1 & jp2) is ±50 sccm, and in the thrust balance plenum (tb) is ±0.5 sccm.

(three-dimensional effect), which likely explains why the actual device required a somewhat higher turbine pressure for a given speed to generate sufficient torque. At 1.4 million rpm, the differential pressure across the journal bearing was about 5 psi.

The mass flow rates supplied to Device 2 are shown in Figure 9. Here the experimental turbine flow rate (Qmain) is somewhat higher than predicted in MISES, due to the increased pressure required to operate the turbine. The flow rates supplied to each journal pressurization plenum (Qjp1 and Qjp2) are essentially equal, lying on top of each other in Figure 9. The thrust balance plenum was sealed with negligible leakage. Also shown in Figure 9 is an aggregate of the back plenum flows normalized by the main turbine flow. This ratio peaks at a value of 0.3 at low speed, but decreases to 0.13 at the highest speed and is still trending down. Thus, for high-speed applications, the flow needed to operate the hydrostatic journal bearing is small compared to the main turbine flow.

The thrust bearings proved sufficiently stiff so as not to require much manipulation during a run. For Device 2, the back side thrust bearing ran with a constant supply pressure of 35 psig while the front side thrust bearing was varied from about 60 – 85 psig. Each thrust bearing required about 10 sccm of nitrogen, which is negligible in the scale of the other flow rates (Figure 9).

CONCLUSIONS

A 4.2 mm diameter silicon rotor has been spun in a sustained manner at high speed using gas lubricated bearings. These microturbines have reached tip speeds of 300 m/s, corresponding to rotation rates of 1.4 million rpm. At this speed, the turbine provides nearly 5 W of power and achieves a power density greater than 4,000 MW/m³ (based on turbomachinery volume), which is more than twice that achieved by modern aircraft engine turbine technology. This high level of power density is a direct benefit of the reduced length scale [4]. We are currently trying to understand more fully the stability characteristics of micro-gas bearings, with the goal of achieving self-sustained operation. Our fabrication effort is focusing on improving alignment precision and etch uniformity, to decrease the rotor imbalance. We believe this technology will serve as a building block for a wide array of high-power density MEMS devices, including pumps, compressors, heat engines and coolers.

REFERENCES

1. C.-C. Lin, R. Ghodssi, A. A. Ayon, D. Z. Chen, S. A. Jacobson, K. S. Breuer, A. H. Epstein, and M. A. Schmidt, "Fabrication and Characterization of a Micro Turbine / Bearing Rig," *in Proc. 13th IEEE Workshop on Micro Electro Mechanical Systems, MEMS'99,* Orlando, FL, Jan. 1999.

2. J. J. Sniegowski, S. L. Miller, G. F. LaVigne, M. S. Rodgers, and P. J. McWhorter, "Monolithic Geared-Mechanisms Driven by a Polysilicon Surface-Micromachined On-Chip Electrostatic Microengine", *Proc. Solid-State Sensors and Actuators Workshop,* Hilton Head Is., SC, June 2-6, 1996, pp. 178-182.

3. A. H. Epstein *et al.*, "Micro-Heat Engine, Gas Turbine, and Rocket Engines – The MIT Microengine Project", AIAA 97-1773, *28th AIAA Fluid Dynamics Conference,* Snowmass Village, CO, June 1997.

4. S. A. Jacobson, "Aerothermal Challenges in the Design of a Microfabricated Gas Turbine Engine," AIAA 98-2445, 29th Fluid Dynamics Conference, Albuquerque, NM, June 1998.

5. M. Drela and H. Youngren, "MISES 2.1," MIT Computational Aerospace Sciences Laboratory, June 1995.

6. A. Mehra, "Computational Investigation and Design of Low Reynolds Number Micro-Turbomachinery," SM Thesis, Massachusetts Institute of Technology, 1997.

7. D. J. Orr, "Macro-scale Investigation of High Speed Gas Bearings for MEMS devices," Ph.D. Thesis, Massachusetts Institute of Technology, 1999.

8. A. A. Ayón *et al.*, "Characterization of a Time Multiplexed Inductively Coupled Plasma Etcher," *J. Electrochemical Society,* Vol. 146, No. 1, Jan 1999, pp. 339-349.

9. E. S. Piekos, D. J. Orr, S. A. Jacobson, F. F. Ehrich, and K. S. Breuer, "Design and Analysis of Microfabricated High Speed Gas Journal Bearings", *28th AIAA Fluid Dynamics Conference,* 1997, AIAA paper 97-1966.

Electrolysis-Bubble Actuated Gate Valve

Alexandros P. Papavasiliou, Dorian Liepmann, Albert P. Pisano
Berkeley Sensor and Actuator Center, University of California Berkeley
Berkeley, California 94720-1774

ABSTRACT

The extreme efficiency of producing bubbles via electrolysis of water has been put to use in a new design for an integrated MEMS microvalve. These valves are designed to be part of larger systems in which pumps, mixers and chemical sensors are fabricated simultaneously into one integrated device. The power consumption, flow characteristics and movement of the valve have been observed. The valve design used in this research requires only 4.3 μW for actuation. This is more than 4 orders of magnitude less than the power required by similar thermally-actuated valves demonstrated in previous research. With the current design a ratio of 4.7 has been achieved between open and closed flow resistance. In addition, the partially closed valve has demonstrated an intermediate flow resistance, demonstrating the possibility of proportional control.

INTRODUCTION

The microfluidics group at Berkeley is currently developing a system to deliver drugs such as insulin. The system must have integrated fluid control including valves, pumps and mixers. These later two devices also need valves in order to operate. The valves must be fabricated in the same process as the rest of the device and minimize the number of bonding steps in order to make the complete system cost effective. The valves must use as little power as possible so that the entire system can be powered by a small battery. Since the rate of insulin injection needed after meals is about one twentieth the rate needed at other times the goal is to produce a valve with a ratio of open to closed flow resistance of 1 to 20.

Electrochemically generated bubbles have been used previously as low power micro actuators. Neagu [1] used a copper sulfate electrolyte to produce oxygen gas and deflect a membrane. By passing a current through an aqueous solution, Böhm [2],[3] and Jackel [4] have evolved oxygen and hydrogen at the electrodes. Since these gas bubbles are the same temperature as their surroundings, they do not lose any energy through heat loss to the surroundings. As a result, they use quite a bit less energy than thermal bubbles. Böhm [3] and Jakel [4] have also shown bubbles shrinking when in contact with a catalyst. The electrolysis products have a thermodynamic proclivity to revert to water. They remain as gasses only as long as kinetics prevents their reaction. The presence of platinum catalyzes the reaction and allows the gasses to revert to water. This process allows one to control both the creation and destruction of bubbles making them an effective actuator. The addition of energy to the system in the form of heat, or a spark could increase the reaction rate and decrease the cycle time for the actuator.

Previous work [5] demonstrated a micro valve in which vapor bubbles move a free floating silicon gate across a channel. This thermally-actuated microvalve requires as much as 100 mW for actuation. The vast majority of the power is consumed by heat loss
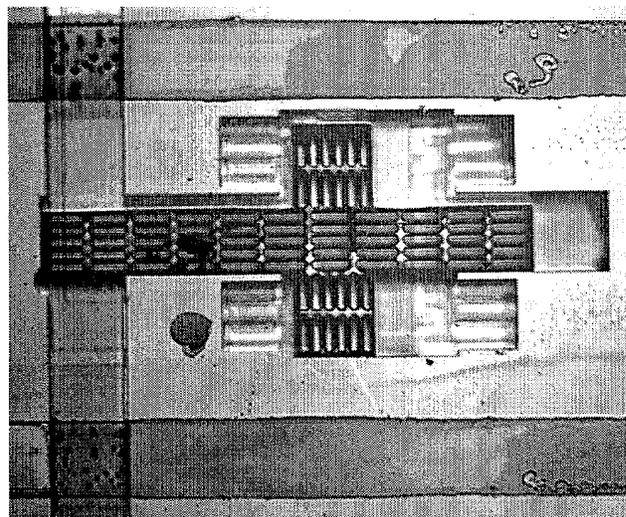
Figure 1. Top view of the gate valve.

to the surroundings. Reducing this power consumption would make the device able to be powered by a small battery for extended periods of time.

Another problem with the valve presented previously was the excessive leakage past the valve. While the gate did move into and block most of the channel and was shown to effect flow, the lack of a proper seal between the valve housing and the cover plate allowed excessive fluid flow past the valve. Providing a better seal would greatly increase the value of the valve.

THEORY

Conventional MEMS electrostatic or thermal actuators are impractical in a fluid environment. Electrostatic actuators will not work in a fluid with mobile ions and thermal actuators require more power in a conductive liquid than they do in air or a vacuum. However, as Lin [6] demonstrated, bubbles can act as large displacement actuators with substantial force. The mixture of hydrogen and oxygen gasses in a 100 μm diameter bubble is 617 times less dense than the water from which they are produced. This allows for very large displacements. The pressure inside a bubble is proportional to surface tension σ divided by radius of curvature r.

$$\Delta p = \frac{2\sigma}{r} \tag{1}$$

Thus, at small scales, surface tension forces become much more significant. The force on the moving gate is given by:

$$F = \frac{2\sigma}{r} \times A \tag{2}$$

where A is the cross sectional area of the part of the gate in contact with the actuating bubble.

The power required to create an electrolysis bubble is consumed by three phenomena. First, 4 electrons must be stripped from 2 hydroxide ions and those 4 electrons must be added to 4 hydrogen ions to produce one molecule of oxygen gas and two molecules of hydrogen gas. This requires 1.23 volts or 4.92 eV. Energy is also consumed by the need to move the hydrogen and hydroxide ions to the electrodes where they are being reacted. As the current density is increased it the rate at which reactants must be replaced increases and the this portion of the energy consumption increases. In addition, transmitting current through the electrolyte consumes energy. This energy consumption is a function of the resistivity of the electrolytic solution and the geometry of the system.

EXPERIMENTAL DESIGN AND FABRICATION

As in the previous work, each valve used in this experiment consists of a piece of single crystal silicon which can be moved across a channel to create a microscopic gate or pin valve. The basic design of the valve is shown in Figure 1. Bubbles are created at the electrodes on either side of the cross piece which expand until they are able to push the gate. The gate is made of a lattice work of silicon to allow etchants to release it from the substrate.

The electrodes are designed as interdigitated fingers to spread the gasses fairly evenly promoting bubble coalescing and mixing. Increasing the complexity of the array should increase the uniformity of the mixing, however, it also increases the proportion of the bubble generators occupied by the spaces between electrodes. This in turn decreases the total electrode area and as a result, increases current density. As discussed above, increasing current density decreases the efficiency of the electrolysis reaction. A compromise was struck between electrode complexity and current density with four interdigitated electrodes designed to create a well mixed bubble while not compromising total electrode area.

As in the previous work, these valves were fabricated in the Silicon on Epoxy or SOE process. The moving parts of the valve are defined in a thin silicon "device" wafer which is sandwiched between a conventional silicon "handle" wafer and a quartz wafer. The electrodes and leads are fabricated on the handle wafer. The handle wafer is oxidized to insulate the platinum electrodes from the silicon substrate. The platinum is sputtered on with a chrome adhesion layer, over a photoresist lift-off layer. The photoresist is
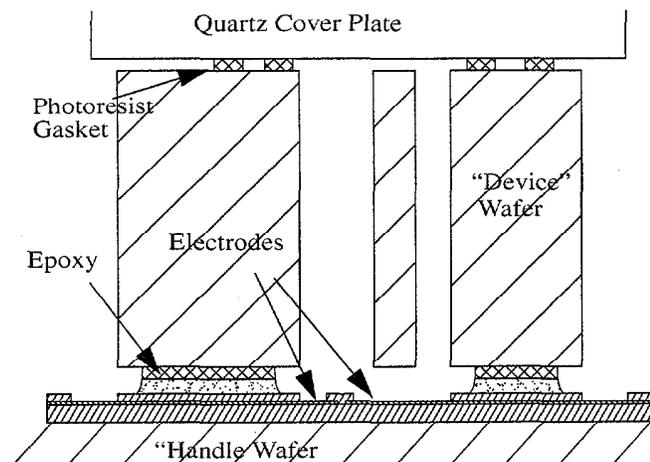
removed in acetone to pattern the platinum. The platinum electrodes are covered with a layer of LPCVD silicon dioxide which is patterned to expose only the desired electrode areas. The device wafer is bonded to the handle wafer with Epo-Tek 301 epoxy and the moving pieces are defined by DRIE etching. A sulfuric acid etch releases the valve elements from the epoxy and exposes the electrodes.

Unlike the devices in Papavasiliou [5] the quartz plate over the top of the valves in this paper are actually bonded on with a photoresist gasket to improve sealing. A layer of Olin OCG 825 is spun on a quartz wafer and patterned as a gasket. The quartz and silicon dies are aligned and pressed together. The sandwich is then heated to 115 °C.

Assuming that the sealing of the gasket is perfect, the major source of leakage in the valve will be through the gap between the top of the valve and the cover plate. The thickness of this gap should be the sum of the thickness of the epoxy/photoresist layers removed from underneath the gate and the thickness of the gasket layer. After fabrication, this gap was measured optically to be 12μm. Leakage was estimated by modeling the gap as pousille flow through a rectangular duct. The difference in flow resistance between the open and closed valve was estimated as $4.04 \times 10^{12} \, Pa \cdot s/m^3$

EXPERIMENTAL RESULTS

Electrolysis bubbles were first created in test channels to characterize their performance. Large polysilicon heaters and platinum electrodes for creating sparks were included in addition to platinum electrode arrays at the bottom of the channels. Bubbles were evolved over the electrodes by passing a current through the water in the channels. The power required to create these bubbles was a function of both the electrode area and generation rate but ranged from around 1 to 10 μwatts.

Figure 2. Representative cross section of the SOE process.
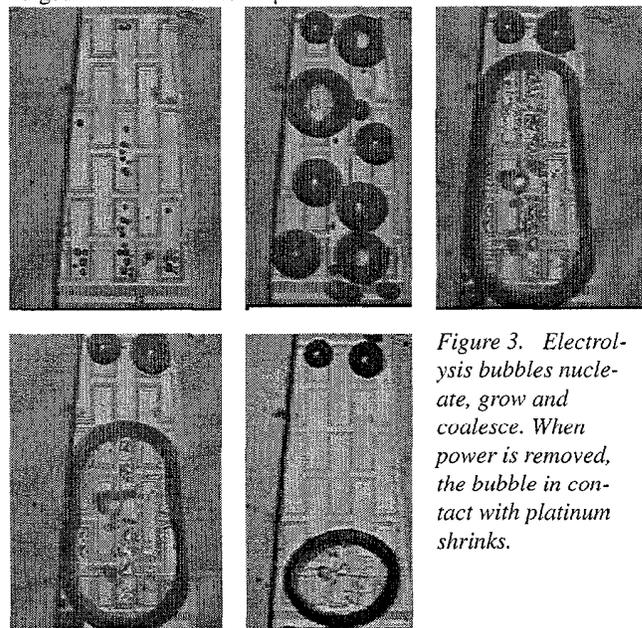
Figure 3. Electrolysis bubbles nucleate, grow and coalesce. When power is removed, the bubble in contact with platinum shrinks.

Three methods were attempted to remove the bubbles. The first method used was merely waiting for the gasses to react and the bubble to shrink. This method although slow, seems effective. Figure 3. shows bubbles generated above electrodes in the test

49
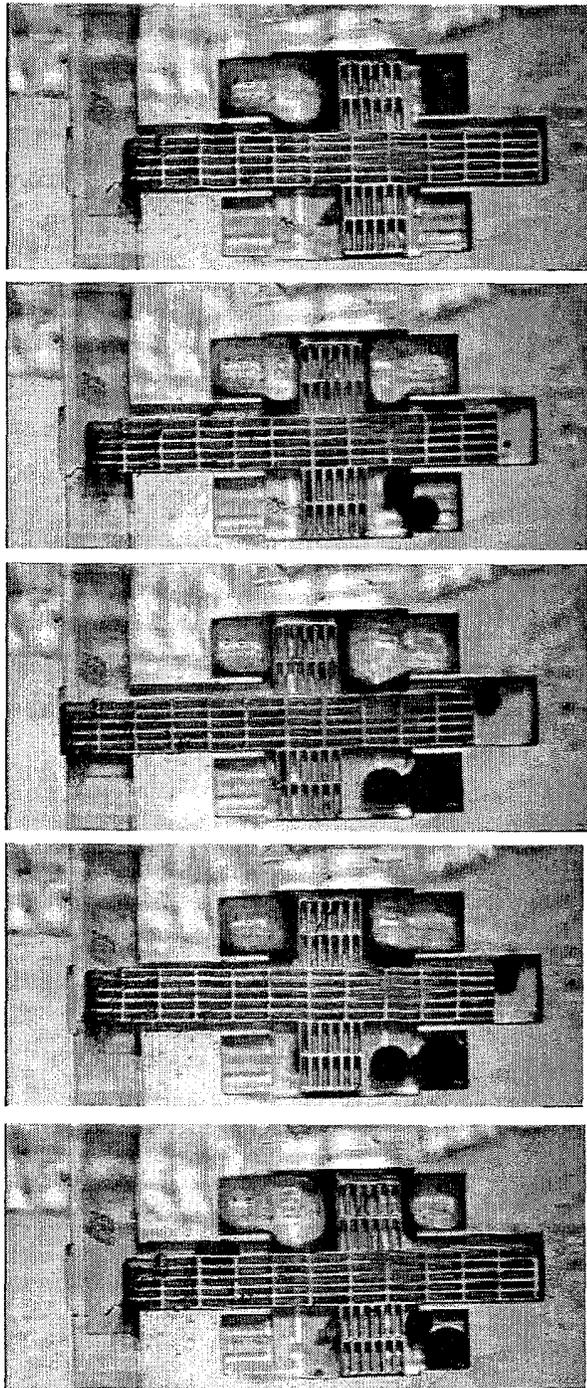
Figure 4. The gate valve in motion. Starting in the open position, bubbles push it closed, then open again

structure. The bubbles grow and coalesce into one larger bubble. After power is removed the bubbles shrink. Notably the small bubbles in the top of the frame shrink much more slowly than the larger bubble. These bubbles are small enough that they do not actually touch the platinum at the bottom of the channels. As a result their shrinkage is due only to dissolution into the fluid. The fact that these bubbles shrink so much more slowly than the bubbles in contact with the platinum indicates that a catalytic reaction is indeed taking place.

Removing the bubbles should proceed more quickly with the addition of energy from either the sparking electrodes or the polysilicon heater. However, the experiments showed that the power input made these methods impractical. In repeated experiments, the bubble shrinks nearly instantaneously when a spark is put across the sparking electrodes, however the bubble invariably does not completely go away. Apparently, the combustion of the bubble which was initiated by the spark is quenched by the rate of heat transfer out of the bubble. Adding heat with the resistor had similarly disappointing results. In addition to increasing the reaction rate, heating the area decreases the solubility of gasses. As a result, heating the area actually swells the bubble with gasses coming out of solution. It was found that cyclic heating can reduce the bubble size but not dramatically faster than simply allowing the reaction to proceed at room temperature.

Creating the actuation bubbles in the test channels requires from 10-50 µJ. In contrast, creating a spark requires around 30 mJ and heating required around 5-20 J. Since both methods require orders of magnitude more power than the initial actuation and do not dramatically improve performance, they were abandoned. It seems that the best way to improve the speed of such actuators may be a method of removing a bubble from the actuation chamber akin to the work done by Evans [8].

Electrolysis bubbles were then incorporated into the microvalve. The movement of the valve is demonstrated in Figure 4. It starts in the fully closed position, moves to the fully open position, and then moves to the fully closed position again. Once the valve has been moved in one direction, the first actuation bubble must be removed in order to move the valve in the opposite direction. Destruction of the bubble through catalytic reaction is a relatively slow process and as a result, one entire cycle, from fully open to fully closed an back to fully open, takes on average around 120 seconds. Carefully regulating the voltage supplied to the two sides of the valve, will allow a controlled bubble on each side of the cross piece to hold the gate in any position allowing proportional control of the flow resistance. Power consumption was measured by regulating voltage and measuring current drawn through the device. Maximum power consumption, while expanding the bubble to push the valve, was measured as 4.3 µW (3.3 V and 1.3 µA). As soon as power is removed, the bubbles begin to shrink due to catalysis of the gasses. Power must be added to maintain the bubbles and hold the valves in place. Electrical input of 2.5 Volts at 0.1 µA has been shown to be sufficient to hold the bubbles in place. The Energizer™ 364, a typical watch battery, contains 106 J [7], approximately enough power to actuate this valve continuously for 285 days or hold the gate in position for 11.2 years.

The flow resistance of the valve was measured by imposing a constant flow rate with a syringe pump and measuring the pressure drop across the die. This measurement was done at several flow rates to create a plot of pressure drop as a function of flow rate. These measurements were performed with the valve open, closed and partially closed. The results of these experiments are shown in Figure 5. The flow resistance of the valve is still several times lower than the expected value, however, there is more than a factor of 5 difference between the open and closed valves. The flow resistance of the closed valve has been improved over the valves shown in previous efforts [5] by a factor of 23. In addition, the flow resistance was measured when the valve was partially closed demonstrating the possibility of proportional control.

The initial actuation bubble must be removed to allow the

Delta P vs Flow Rate
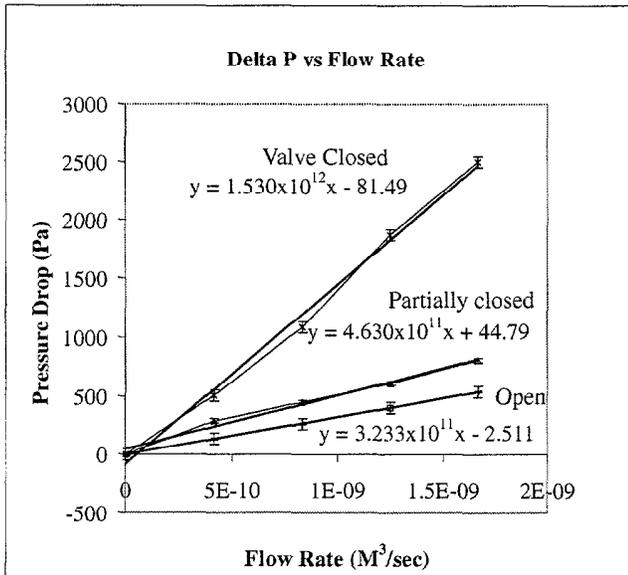
Figure 5. Experimental measurements of flow resistance for the valve fully closed, partially closed, and fully open.

gate to move in the opposite direction. This requires waiting for the relatively slow catalytic reaction process. One operational cycle (fully closed to fully open and back) requires approximately 120 seconds.

The gasket provides this valve with substantially improved sealing. Flow resistance was measured for the entire device including interconnects and channels. The resistance of the device with the valve closed was 4.7 times the resistance with the valve opened. The difference between opened and closed resistance was

Table 1: Valve Data

parameter	value	units
Flow resistance of open valve	1.53×10^{12}	$Pa \cdot s/m^3$
Flow resistance of closed valve	3.23×10^{11}	$Pa \cdot s/m^3$
differential resistance	1.30×10^{12}	$Pa \cdot s/m^3$
predicted differential resistance	4.04×10^{12}	$Pa \cdot s/m^3$
previous differential resistance	4.14×10^{10}	$Pa \cdot s/m^3$
Acutation power consumption	4.3	μW
Actuation Voltage	3.3	Volts
Actuation Current	1.3	μA
Holding power consumption	0.3	μW
Holding voltage	2.5	Volts
Holding current	0.1	μA
Minimum actuation period	120	seconds

37 times as high as that of the previous valve, however, it is still more than a factor of 4 away from the requirement for insulin regulation. However, simply putting 4 or 5 such valves in series would presumably provide sufficient control. In addition to the open and closed positions, the gate was held partially closed and an intermediate flow resistance was measured demonstrating the possibility of variable flow control.

CONCLUSIONS

An extremely low power electrolysis bubble actuated gate valve has been demonstrated. The actuation scheme requires only 4.3 μW for actuation and 0.3 μW to hold the valve in place more than 4 orders of magnitude less than similar thermal bubble actuators. The speed of the valve is limited by the need to find a way to remove the bubble quickly. The flow characteristics of the valve have been measured in the opened, closed and partially closed positions, demonstrating the possibility of proportional control. A new photoresist gasket has been shown to greatly reduce leakage around the valve. However, work still needs to be done on decreasing the leakage rate in order to make it commercially viable.

ACKNOWLEDGMENTS

REFERENCES

1. C. Neagu, J.G.E. Gardeniers, M. Elewnspoek, and J. J. Kelly, "An electrochemical microactuator: principle and first results," *J. Micromechanical Syst.* 5 (1), pp 2-9, (1996)

2. S. Bohm, W. Olthuis, and P. Bergveld, "An Electrochemically Actuated Micro Dosing System with Improved Dosing Control," *Proceedings of the ASME MEMS Division, 1999 IMECE*, vol.1 pp.391-395 (1999)

3. S. Böhm, W. Olthuis, P. Bergveld, "A Bi-Directional Electrochemically Driven Micro Liquid Dosing System With Integrated Sensor/Actuator Electrodes," *Proceedings 10th International Conference on Micro Electro Mechanical Systems*, pp92-95 (2000)

4. J. L. Jakel, J. J, Johnson, and W. J. Tomlinson, "Bistable optical switching using electrochemically generated bubbles," *Optics Letters*, Vol. 15, No. 24 pp 1470-1472 (1990)

5. A. P. Papavasiliou, D. Liepmann, A. P. Pisano, "Fabrication of a Free Floating Silicon Gate Valve," *Proceedings of the ASME MEMS Division, 1999 IMECE*, vol.1 pp.435-440 (1999)

6. L. Lin "Selective Encapsulations of MEMS: Micro Channels, Needles, Resonators and Electromechanical Filters," *Ph.D. Thesis, University of California at Berkeley, Berkeley, CA.*

7. Energizer Product Sheet EPS-4322A

8. J. Evans, and D. Liepmann, "The Bubble Spring and Channel (BSAC) Valve: An actuated, Bi-stable Mechanical Valve for In-plane Fluid Control," *Digest of Technical Papers, 10th International Conference on Solid-State Sensors and Actuators, Transducers '99.* pp 1122-1125 (1999)

AN ACTIVE MICRO MIXER USING ELECTROHYDRODYNAMIC (EHD) CONVECTION

Jin-Woo Choi and Chong H. Ahn
Center for Microelectronic Sensors and MEMS (CMSM)
Department of Electrical and Computer Engineering & Computer Science
University of Cincinnati
Cincinnati, OH 45221-0030

ABSTRACT

This paper presents a new active micro mixer using electrohydrodynamic (EHD) convection for the application of microfluidic-based biochemical analysis systems and biochips. To understand the EHD convection mixing, analytical analyses on the micro mixer have been performed for two different liquid samples with different electric conductivities. Using the analytical results, a new active micro mixer has been designed, fabricated, and also characterized by flowing two liquid samples through the microchannel. DI water and saltine water solution, which have less than the volume of 10 pl, have been fully mixed in the mixing zone at 5 ~ 20 V of the applied voltage. Since the realized micro mixer has simple structure and no mechanical moving parts, it shows very reliable mixing performance. The active micro mixing device studied in this work also shows feasible mixing capabilities in nano- or pico-liter range of liquid volumes by applying a low voltage of 5 V across the microchannel.

INTRODUCTION

In microfluidic-based biochemical analysis systems, mixing of the liquid samples is considered as one of the most challenging tasks to achieve an appropriate reaction in a short period of time. There are, however, some difficulties in realizing reliable micro mixing devices, because the fluid in microchannels behaves as a laminar flow in most case due to low Reynolds number in microchannels [1-2]. Mixing of the liquid samples in macro-scale systems is usually achieved by stirring or agitating of the liquid samples, but these methods are almost not applicable to micro-scale devices due to its small size and fabrication compatibility. For these reasons, several micro mixing devices have been recently developed and reported [3-7]. Most of them are passive micro mixers [3-6], but few semi-active micro mixers with enough mixing capabilities have been reported [7]. Passive mixing devices can also be useful in micro total analysis systems (μ-TAS) and biochip applications, but they have limitation when the system requires precise control of mixing performances concerning mixing volume and time.

In this paper, we propose a new active micro mixer using electrohydrodynamic (EHD) convection phenomena [8] for the application of microfluidic-based biochemical analysis systems (Figure 1).

The electrohydrodynamic (EHD) and magnetohydrodynamic (MHD) phenomena have been explored since early 1960's and there have been studies to realize the EHD and MHD micropumps [9-13]. Both EHD and MHD phenomena are attractive as scales go down, specifically, for microfluidic control because of its simple structural advantage in micro- and nano-scale fluidic control. In addition, since these EHD and MHD devices include no mechanically moving parts, they provide more reliable mixing.

Figures 2 and 3 show schematic illustration and simple working principle of the proposed active micro mixer using EHD convection. From the electromagnetic theory, surface charges are induced and accumulated on the boundary of dielectric materials, which are the liquid samples in this case. When an external electric field is applied over the surface charges, the charges will be moved with liquids due to a shear stress generated at the
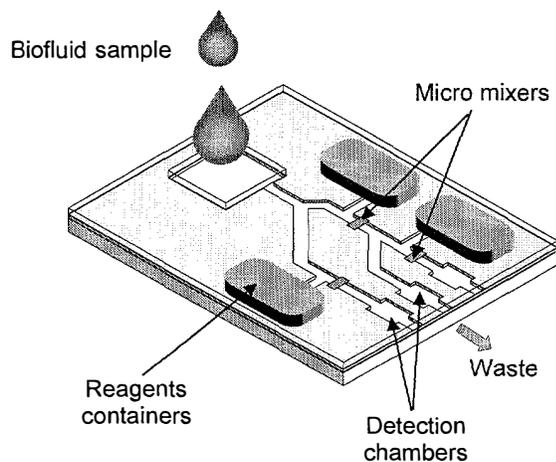
Figure 1. *Schematic illustration of an on-chip microfluidic biochemical analysis system*
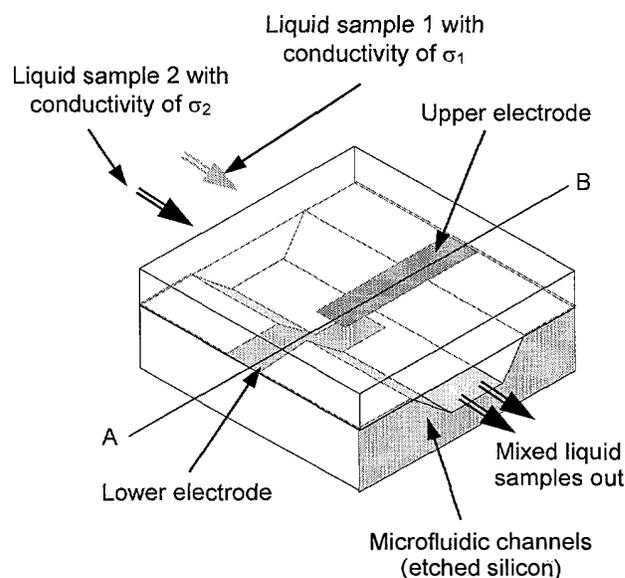
Figure 2. *Schematic illustration of the active micro mixer*
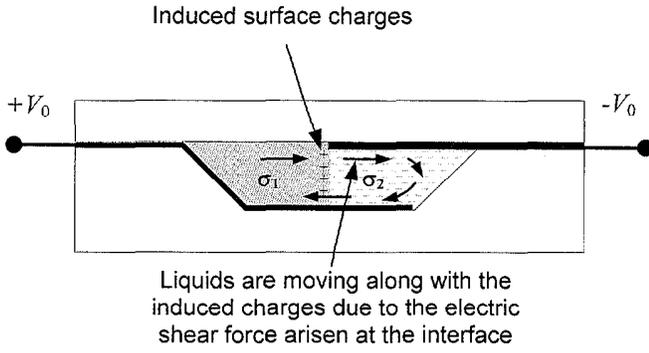
Figure 3. *Cross sectional view of the active micro mixer (along the line A-B in Figure 2)*

interface layer between the liquids to be mixed. These phenomena can continuously occurs and thus the convection of the liquid samples will continue until the liquid samples get fully mixed to eliminate the interfacial shear stress. The mixing speed is governed by the parameters of applied electric fields, electric properties like conductivity, and geometry of the electrodes. All these procedures can be understood analytically for the ideal case.

THEORY

The parameters and geometry of the device are defined in Figure 4. Each region indicates two different liquid samples which have different conductivities. σ_I and σ_{II} denote the conductivities of liquid I and liquid II, respectively.

In the region I, we assume that y-directional electric fields can be described as

$$E_y^I = \frac{V_0}{l}\hat{y},\tag{1}$$

where V_0 is the applied voltage.

From continuity condition of the tangential component of electric fields at the boundary of the dielectric materials, y-directional electric fields near the interface in the region II can be written as

$$E_y^{II} = \frac{V_0}{l}\hat{y}.\tag{2}$$
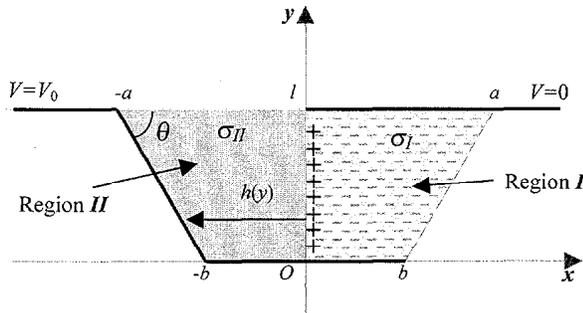
Figure 4. *Model and parameters for analytical calculation*

The interface assumes a distribution in electrical potential that varies from V_0 at the upper electrode to zero at the lower electrode. Because the lower electrode has zero potential, with a spacing $h(y)$

that varies essentially linearly with y, there is a surface charge induced on the interface. Hence, the distribution of potential on the interface is

$$\Phi\big|_{x=0} = \frac{V_0}{l}(l-y),\tag{3}$$

and $h(y)$, the distance from electrode B to the interface, is

$$\begin{aligned}h(y) &= \cot\theta \cdot y + b \\ &= a - \cot\theta \cdot (l-y).\end{aligned}\tag{4}$$

In the region II, therefore, x-directional electric fields are

$$E_x^{II} = -\nabla\Phi = \frac{V_0}{l}\cdot\frac{y}{a-\cot\theta\cdot(l-y)}\hat{x}.\tag{5}$$

To obtain x-directional electric fields near the interface in the region I, the continuity condition of the normal current density at the interface

$$\hat{n}\cdot(\sigma_{II}E^{II} - \sigma_I E^I) = 0\tag{6}$$

is used. Then, the E_x^I is described as

$$E_x^I = \frac{\sigma_{II}}{\sigma_I}E_x^{II} = \frac{\sigma_{II}V_0 y}{\sigma_I l[a-(l-y)\cot\theta]}\hat{x}.\tag{7}$$

Maxwell stress tensor at the interface is given as

$$T_{ij} = \varepsilon\left(E_i E_j - \frac{1}{2}\delta_{ij}E_k E_k\right),\tag{8}$$

where δ_{ij} is Kronecker delta.

With these equations, we can derive x-directional shear stress at the interface.

$$\begin{aligned}T_{xx} &= \frac{\varepsilon}{2}\left(\frac{\sigma_{II}V_0^2 y^2}{\sigma_I l^2[a-(l-y)\cot\theta]^2} - \frac{V_0^2}{l^2}\right) \\ &= \frac{\varepsilon V_0^2}{2l^2}\left(\frac{\sigma_{II}y^2}{\sigma_I[a-(l-y)\cot\theta]^2} - 1\right)\end{aligned}\tag{9}$$

The electric force on the interface can be obtained by surface integral of the shear stress. Although we only calculated that the x-directional shear stress, y-directional forces also exist along the interface of the liquids.

From Eq. (9), the force on the interface is determined by applied voltage (V_0), depth of the channel (l), width of the channel (a), and the ratio of the conductivity of the liquid samples (σ_{II}/σ_I). Looking into the term within the parenthesis in Eq. (9), the major parameter which effects on the force is the ratio of the conductivity. Figure 5 shows the plotted electric force profile on
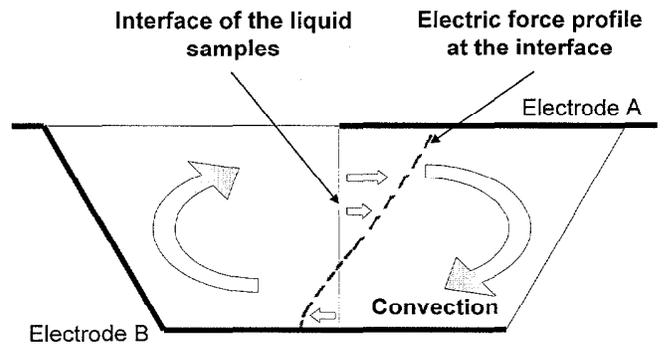
Figure 5. *The plotted electric force profile on the interface*
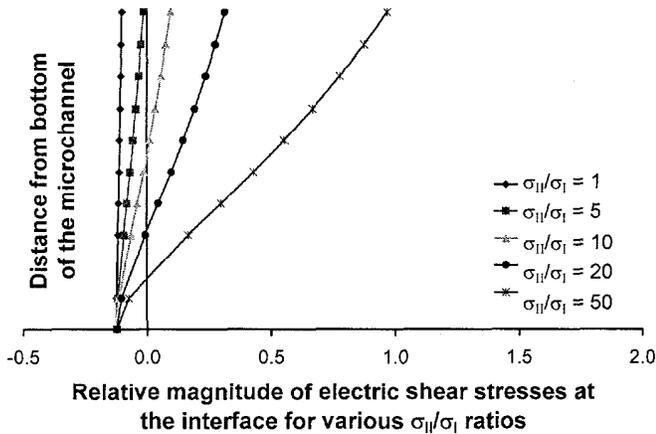
Figure 6. *Relative magnitude of electric shear stresses at the interface for various conductivity ratios*

the interface for the case that the ratio of the conductivity is larger than 1. The electric force profile varies along the interface and the liquids in the microchannel are assumed incompressible, so the unbalance between top and bottom of the channel along the interface causes clockwise convection in the channel and the two liquid samples will be mixed. Figure 6 shows the relative magnitude of the electric shear stress profiles at the interface for the various conductivity ratios. For the higher conductivity ratio, the electric shear stress has much larger unbalance along the interface and causes much faster convection than for the lower conductivity ratio. If the two liquids have same conductivity, then convection is hardly occurred. Further simulations using CFD are now under development.

FABRICATION AND EXPERIMENTS

Based on the analytical study, an active micro mixer has been designed fabricated using simple micromachining techniques. The (100) silicon wafer was patterned and unisotropically etched in potassium hydroxide solution for the microchannel of the device. The width and depth of the microchannel are 200 μm and 60 μm, respectively. After etching process, silicon wafer was oxidized and the electrodes (Cr/Au) were deposited and patterned on both
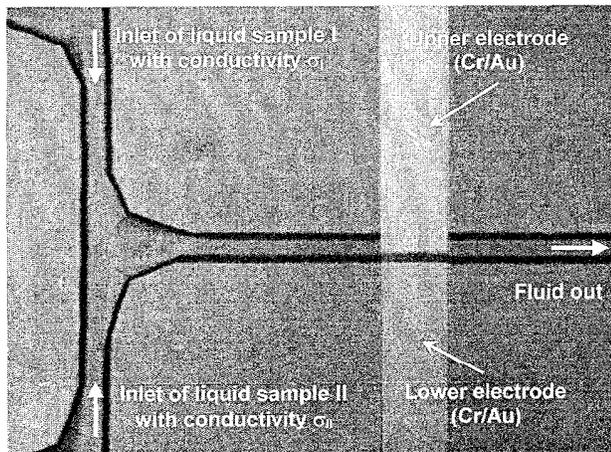
Figure 7. *Microphotograph of the fabricated active micro mixer (upper electrode are shown from back side through the glass wafer)*

silicon and glass wafer. Finally, those two wafers have been properly bonded using fluoropolymer [14]. The fabricated device is shown in Figure 7.

To demonstrate the proposed mixing concepts, two different liquid samples have been chosen: one is water (low conductivity) and the other is saltine water (high conductivity) which was dyed for the optical monitoring.

Two liquid samples have been injected through the fabricated device as shown in Figure 8 (a). With no applied electric fields, the two injected liquid samples were not mixed in the microfluidic channel as clearly showing two separate liquid streams along the microchannel. By applying electric fields to the electrodes, however, the flowing liquid samples were fully mixed after passing the electrodes due to the electric shear force generated on the interface between the liquid samples. Figure 8 (b) obviously shows the function of the proposed active micro mixer, demonstrating two separate liquid streams before reaching the electrodes and one liquid stream after passing the mixing zone.
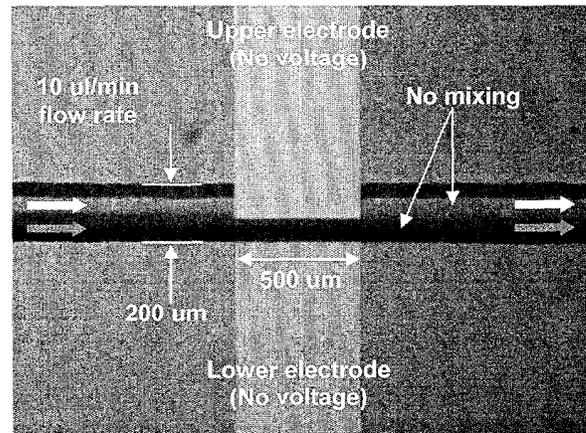
(a) No potential applied (no mixing)
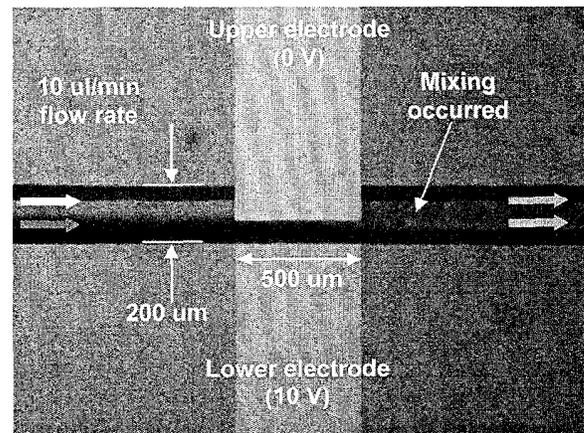
(b) 10 V applied with the flow rate of 10 μl/min (Liquid samples are mixed passing electric fields)

Figure 8. *Microphotograph for the mixing capability tests*

RESULTS AND DISCUSSION

At a given geometry of the device and with the selected liquid samples, the mixing speed and capability of the mixer depends on the strength of the applied electric fields and the flow rate of the

samples. Figure 9 shows the characteristics of the micro mixer by measuring the voltage at which the flowing liquid samples get mixed. At a flow rate of 10 μl/min, two fluids are immediately mixed with a low mixing voltage of 7 V.
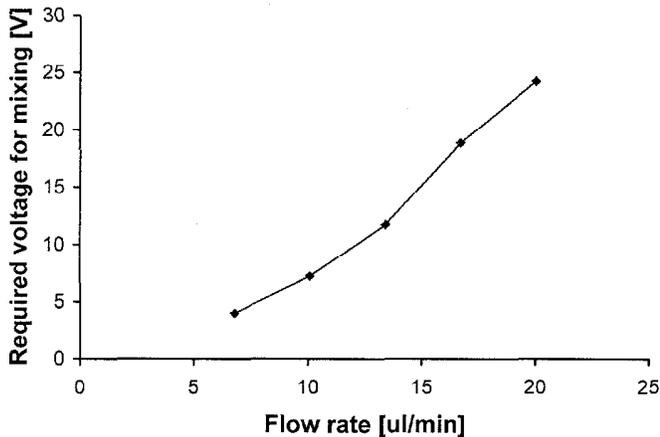
Figure 9. Minimum voltage required for mixing of the flowing DI water and saltine water

The excellent mixing performance of the proposed active micro mixer has been demonstrated using saltine water solution and DI water. However, the mixing of the solutions, which have very similar low conductivities, was not successfully achieved with such as DI water–methanol and DI water–isopropyl alcohol sets. When we use very high conductive aqueous solutions or operate the mixing device at very low flow rate, we could see the electrodes to be occasionally electrolyzed because the electrodes are exposed to the relatively high electric current once mixing begins. So further research on the dynamic characteristics related to conductivity of the solutions and voltage levels is now undergoing in addition to the structural optimization of the device.

CONCLUSION

A new active micro mixer using electrohydrodynamic convection has been proposed, fabricated, and characterized in this work. Basic analytical analyses have been studied to describe mixing mechanism. The surface charges are induced at the interface of the liquid samples which have different conductivities, and these surface charges react with the applied electric fields to generate electric shear forces. Mixing capability has been clearly shown with DI water and saltine water solution, which have less than 10 pl of the volume. By applying electric fields, the separate flow streams get mixed passing the electrodes. The micro mixing device realized in this work has simple structure and no mechanical moving part, which can provides a reliable mixing function on biochips.

With the demonstrated device, active mixing control of electrolytes can be envisaged toward microfluidic-based biochemical analysis systems and biochip applications.

ACKNOWLEDGMENT

REFERENCES

1. J. Pfahler, J. Harley, H. Bau, and J. Zemel, "Liquid Transport in Micron and Submicron Channels," *Sensors and Actuators, A21-23*, pp. 431-434, (1990).
2. G. M. Mala and D. Li, "Flow Characteristics of Water in Microtubes," *International Journal of Heat and Fluid Flow, 20*, pp. 142-148, (1999).
3. J. Evans, D. Liepmann, and A. P. Pisano, "Planar Laminar Mixer," *Proceedings of the 10th International Workshop on Micro Electro Mechanical Systems (MEMS)*, Nagoya, Japan, 1997, pp. 96-101.
4. Z. Yang, H. Goto, M. Matsumoto, and T. Yada, "Micro Mixer Incorporated with Piezoelectrically Driven Valveless Micropump," *Proceedings of the Micro Total Analysis Systems '98 (μ-TAS '98) Workshop*, Banff, Canada, 1998, pp. 177-180.
5. U. D. Larsen, W. Rang, and P. Telleman, "Design of Rapid Micromixers Using CFD," *Technical Digest of the 10th International Conference on Solid-State Sensors and Actuators*, Sendai, Japan, 1999, pp. 200-203.
6. T. Fujii, K. Hosokawa, S. Shoji, A. Yotsumoto, T. Nojima, and I. Endo, "Develpoment of a Microfabricated Biochemical Workbench – Improving the Mixing Efficiency," *Proceedings of the Micro Total Analysis Systems '98 (μ-TAS '98) Workshop*, Banff, Canada, 1998, pp. 173-176.
7. R. U. Seidel, D. Y. Sim, W. Menz, and M. Esashi, "Capillary Force Mixing Device as Sampling Module for Chemical analysis," *Technical Digest of the 10th International Conference on Solid-State Sensors and Actuators*, Sendai, Japan, 1999, pp. 438-441.
8. H. A. Haus and J. R. Melcher, *Electromagnetic Fields and Energy*, Prentice Hall, 1989.
9. A. V. Lemoff, A. P. Lee, R. R. Miles, and C. F. McConaghy, "An AC Magnetohydrodynamic Micropump: Towards a True Integrated Microfluidic System," *Technical Digest of the 10th International Conference on Solid-State Sensors and Actuators*, Sendai, Japan, 1999, pp. 1126-1129.
10. J. Jang and S. Lee, "MHD (Magnetohydrodynamic) Micropump Using Lorentz Force," *MEMS Symposium, ASME International Mechanical Engineering Congress and Exposition*, Anaheim, CA, 1998, pp. 439-443.
11. G. Fuhr, R. Hagedorn, T. Muller, W. Benecke, and B. Wagner, "Microfabricated Electrohydrodynamic (EHD) Pumps for Liquids of Higher Conductivity," *IEEE/ASME Journal of Microelectromechanical Systems (JMEMS)*, Vol. 1, No. 3, pp. 95-98, 1992.
12. A. Richter and H. Sandmaier, "An Electrohydrodynamic Micropump," *Proc. of IEEE MEMS*, 1990, pp. 99-104.
13. J.-W. Choi and Y.-K. Kim, "Micro Electrohydrodynamic Pump Driven by Traveling Electric Fields," *Proceedings of 1995 IEEE Industry Application Society Annual Meeting*, Orlando, FL, 1995, pp. 1480-1484.
14. A. Han, K. W. Oh, S. Bhansali, H. T. Henderson, and C. H. Ahn, "A Low Temperature Biochemically Compatible Bonding Technique Using Fluoropolymers for Biochemical Microfluidic Systems", *Proceedings of the 13th International Conference on Micro Electro Mechanical Systems (MEMS)*, Miyazaki, Japan, 2000, pp. 414-418.

A SILICON HEAT EXCHANGER WITH INTEGRATED INTRINSIC-POINT HEATER DEMONSTRATED IN A MICROPROPULSION APPLICATION

Robert L. Bayt
United Technologies Research Center
E. Hartford, Connecticut 06108

Kenneth S. Breuer
Division of Engineering, Brown University
Providence, Rhode Island 02912

ABSTRACT

Efficient heating of a fluid is demonstrated using a novel heat exchanger in which bulk silicon forms both the heater structure and the resistive heating elements. Current passed through the heater raises the temperature of the heater fins and this energy is transferred to a fluid flowing between adjacent fins. By exploiting the change in sign of the temperature coefficient of resistivity of the heavily doped silicon, the temperature of the system is stably maintained at the intrinsic point. A heat exchanger of this nature is integrated with a nozzle, resulting in a microthruster with elevated chamber temperature, which greatly improves the specific impulse, or thrust per unit weight flow of propellant. A numerical model is presented to optimize the heater design. Benchtop tests demonstrate the inherent stability of the intrinsic point heater design while thrust tests demonstrate the improved fuel economy of the micropropulsion system.

INTRODUCTION

A wide variety of MEMS applications employ an integrated heater, including microreactors (to catalyze reactions) and valves (to actuate the poppet). In general, these heaters typically use a thin-film metal or polysilicon as a resistor to generate heat. These systems work well, but are typically hampered by the relatively low contact area between the fluid and the heating element [1]. The heater effectiveness can be characterized as the ratio of the observed temperature rise of the fluid to the maximum possible temperature rise that could occur for a given wall temperature:

$$\varepsilon = \frac{T_{outlet,\,fluid} - T_{inlet,\,fluid}}{T_{outlet,\,wall} - T_{inlet,\,fluid}} . \qquad [1]$$

Thus, a heater with very good thermal contact between the fluid and heating element would bring the fluid up to the wall temperature and achieve an effectiveness of unity. Since the dominant heat transfer mechanism is forced convection, heat transfer to the fluid is governed by Newton's Law of cooling,

$$q = hA(T_{fluid} - T_{wall}) , \qquad [2]$$

where the area of fluid-heater interaction (A) is often the practical limitation for achieving high effectiveness.

A second issue of common concern in gas flow heat exchangers is heater stability. This arises because the gas viscosity increases with temperature. Thus, if a local hotspot develops, the fluid viscosity will increase, increasing the pressure drop across the heat exchanger passage and thus reducing the mass flow through the channel. This, in-turn, reduces the cooling and further raises the temperature in the channel. In this manner a thermal runaway can ensue leading to device failure.

This paper presents a microfabricated solution for achieving a highly effective heater in which the heater elements are formed

from the structural material of the fluid system - single crystal silicon. This architecture has two principal advantages. First, the electrical and mechanical functions are combined. Second, the thermo-electric properties of silicon allows for an inherently stable operation at high temperature without risk of thermal runaway.

In this paper, the heater is used to increase the chamber temperature for fluid entering a micronozzle to create a propulsion system for a microspacecraft. In such applications, maximizing the thrust per unit weight of propellant (i.e. specific impulse, I_{sp}) is highly desirable and is enhanced by increasing the chamber temperature by preheating the gas prior to expansion. We also present a simple model to maximize the heater effectiveness while minimizing the pressure drop through the fin array. Varying the number of passages, their length, thickness and spacing accomplishes this goal.
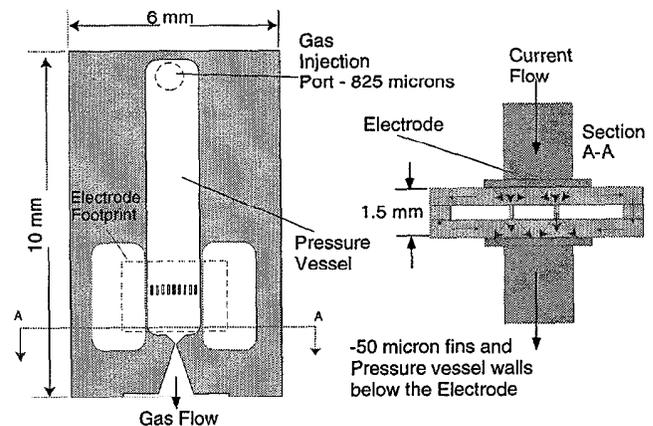
Figure 1. Schematic of device. The thruster is deep etched to provide maximum surface area of the heat exchanger and a high-aspect ratio nozzle.

HEATER OPERATION

The overall heater design is shown in Figure 1. Gas enters the port in the top wafer and flows through the heat exchanger, which is defined by narrow parallel fins upstream of the nozzle. The entire structure is fabricated from heavily doped P-type silicon. An electric current flows from top to bottom through the device, and is focused through the fins, heating them resistively.

The properties of silicon as a semiconductor are used to maintain stable operation of the device at high temperature. By fabricating the heater using heavily doped P-type silicon wafers, the dopant holes will be the primary charge carriers at low-temperatures. As power is dissipated and the temperature of the device increases, the electrons bound in the the in the silicon valence bands become thermally excited to the conduction band. When the number of intrinsic carriers becomes dominant, the resistivity of

the material is governed by statistical mechanics, and decreases exponentially with temperature.

This trend in resistance is shown in figure 2, and indicates the intrinsic point for a silicon sample doped to 10^{19} atoms/cc [3]. The key feature of figure 2 is the change in the sign of the temperature coefficient of resistivity at the intrinsic point.

The heater is operated in a constant current mode. As the current increases, the dissipated power, temperature, and resistance, all rise in response. However, when the intrinsic temperature is reached, any increase in dissipated power and temperature results in a *drop* in resistance. Consequently the dissipated power *falls* and the device returns to operation at the intrinsic point. Thus, a stable operating temperature is maintained with the feedback provided by the resistive properties of silicon.
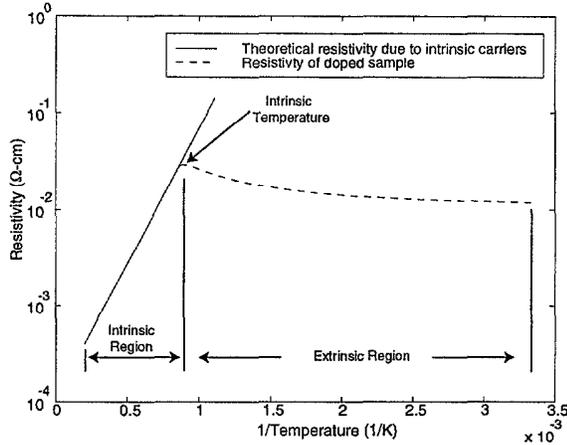
Figure 2. The resistivity of doped silicon is peak when intrinsic and dopant charge carriers are equal. This plot is for a P-type 10^{19} atom/cc silicon wafer. [3]

HEATER MODEL

With a heater design selected, the geometry (shown in figure 3) needs to be optimized to yield the highest heater effectiveness while maintaining the lowest pressure drop. Heat transfer in this instance is a convective transport problem governed by the bulk motion of the fluid. Fin spacing (w_{gap}), the height of the channel (h_o), the length of the channel (L), and the number of channels (N) form the heater design space. The dissipated power is driven by the fin width, and the silicon resistivity. The operational parameters of the device are the mass flow rate ($\dot{m}$), the current (i), and the gas supply temperature (T_f). With these parameters specified, a numerical simulation is used to model the steady-state operating conditions of the power dissipated, the pressure drop across the fins, and the exit temperature. From this, the effectiveness can be computed.

A 1-D model (along the heater length) is used and illustrated in figure 3. Variations at each local cross-section are neglected (valid for moderate fin spacing as in the Graetz problem derived in White [4]). By summing the energy inputs into each control volume and non-dimensionalizing, the following system results:

$$\frac{d^2 T_w}{dx^2} + Bi(T_w - T_f) - s^* = 0 \quad (a \qquad [3]$$

$$\frac{dT_f}{dx} = St(T_w - T_f) \qquad (b$$

where $\quad Bi = \dfrac{hPh_o^2}{k_w A_c}; \quad s^* = \dfrac{\dot{q} h_o^2}{T_i k_w}; \quad St = \dfrac{hh_o PN}{\dot{m} c_p} \qquad [4]$
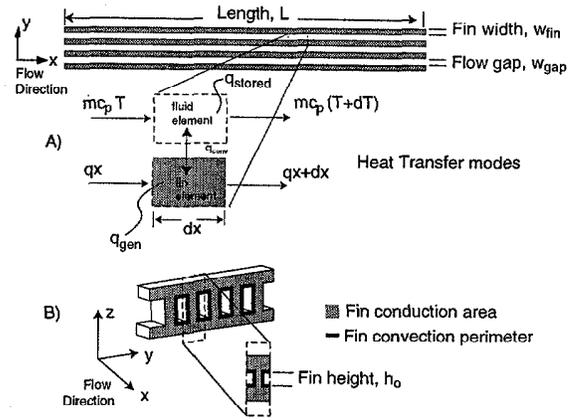
Figure 3. Design simulations are based on a lumped-parameter model that considers the heat generation and transfer along a symmetrical channel element.

The model assumes that heat is generated at a rate set by the electrical model, which is a constant current supply, and temperature-dependant heater resistance depicted in figure 2. The dissipation is treated as a volumetric heating source within the fin ($\dot{q}$). The heat transfer is computed by performing an energy balance on differential fin and fluid elements as shown in figure 3a. The heat is transferred to the fluid convectively and into the up and downstream portions of the fin through conduction (equation 3a). The heat is received by the fluid from the fin and advected downstream (equation 3b). In order to enhance the fidelity of the model, and to take advantage of the heater symmetry, the analysis is performed for the fin segment shown in figure 3b. The channel symmetry plane is the boundary one-half of the channel width on each sides of the fin. The conduction area is the full cross-section, and the convection area is the wetted perimeter (P) outlined. By taking into account the upper and lower surfaces, a closer approximation to the heater geometry is achieved. The heat transfer is governed by the fluid mechanics in the channel. Thus, the convection parameter h is a function of Reynolds number and the distance along the fin. The correlation for computing the heat transfer coefficient was developed by Stephan and is reported by Kakac et al. [5] for developing flow in a two-dimensional duct.

The coupled system requires three boundary conditions. The fins are assumed to be insulated at each end ($dT/dx = 0$), and the inlet fluid temperature (T_i) is specified (300 K).

As denoted by their symbols, the coefficient of the wall-fluid temperature difference in the fin equation is a form of the Biot (Bi) number, and in the fluid equation the coefficient of the wall-fluid temperature difference is a form of the Stanton (St) number. The non-dimensional source term is represented by s^*. It is important to observe the relevance of each of the parameters. The Biot number governs the ratio of the conduction along the fin to the convective heat transfer. High Biot numbers combined with high fin length-to-width aspect ratios result in large temperature differences between the entrance region and exit region of the fins. Because of the high heat transfer in the entrance region, all of the heat generated locally enters into the fluid, and heat generated downstream is conducted along the fin, so it can enter the fluid where gradients are highest. The Stanton number is similar, as it governs the convection to advection ratio along the fin. Thus, when Biot number is low and Stanton number is high, the fins are nearly isothermal. Finally, the source term merely scales the volumetric heat dissipation, relative to that conducted along the wall. When the ratio of Stanton number to the source term is high, then the heater effectiveness will be high. This is due to a majority

of the dissipated heat being convected into the fluid, which is indicative of a high Stanton number flow.

DESIGN STUDIES

In order to minimize the pressure drop and maximize heater effectiveness, a number of test cases were explored. To best understand the effect of the governing non-dimensional parameters, we consider four cases and determine how they affect the heater performance. All cases are operating at a mass flow rate of 378 sccms, and have 50 micron wide fins. This flow rate was chosen to keep the operating power below 10 watts, and maintain as high a Reynolds number through the nozzle as possible to minimize viscous losses (550 at this condition, which has provided reasonable performance [6]). The heater design and performance parameters are summarized in table 1.

Table 1. Heater Design Cases

Case	# of Fins	L (μm)	W (μm)	ε	ΔP (atm)	Bi	S*	St
1	5	500	100	99	0.21	1.1	0.5	6.5
2	**10**	**125**	**50**	**97**	**0.02**	**1.8**	**1.5**	**13.0**
3	5	125	50	85	0.04	1.8	2.5	6.6
4	1	7000	50	90	0.01	0.9	1.5	0.15

Case 1 is a reasonable design. The pressure drop is 7% of the total, and a very large temperature rise is achieved. In case 2, the ratio of the Stanton number to source term has decreased, and as expected, the effectiveness has dropped, but only by 2%. However, the pressure drop has decreased by a factor of 10. Thus, this is a better design due to the higher injected pressure. For all cases, the dissipated power is 7.7 Watts, which is set by the flow rate and temperature rise. The temperature profiles along the fins are depicted in figure 4, and the silicon is almost isothermal.
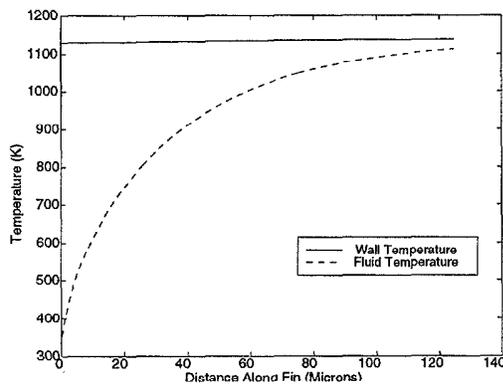
Figure 4. Temperature profile along the channel. Wall temperature is effectively constant due to high conductivity.

Case 3 is the same as Case 2, but with 5 fins. By halving the number of fins, the mass flow through a given channel is doubled. This reduces the temperature rise, as well as, increases the pressure drop. This is the lowest Stanton number to source term ratio, and the lowest effectiveness, yet.

Finally, the case representative of a metalized snaking resistor (a flat plate heater) is analyzed as a case with 1 fin (split between the sidewalls). Even with a chamber 7 mm long, the heater effectiveness is only 90%. By using the design with fins, a more compact arrangement can be attained with a similar pressure drop.
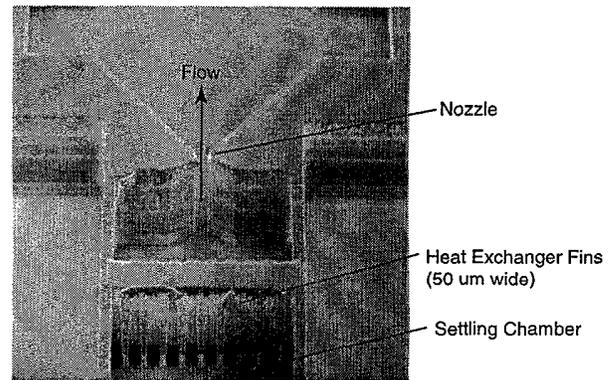
Figure 5. SEM of thruster with integrated heater after first bonding is complete. Bus bar is holding fins together.

HEATER/THRUSTER FABRICATION

The heater/nozzle system is fabricated in silicon using DRIE. A halo mask is used to simultaneously outline the large cavities as well as define the small heater passages and nozzle throat. By matching these widths, a constant loading is maintained during the etch. In addition, a nested mask is used which allows the through-wafer etch to proceed ahead of the heater-fins. This results in the heater fins being connected by a 50-micron high "bus-bar", which maintains their mechanical integrity and provides electrical functionality. After etching, the cleared flow channels are encapsulated by fusion bonding silicon wafers to the upper and lowers surfaces. Figure 5 is a SEM of the 8.25:1 expansion ratio nozzle with a throat width of 65 microns, a nozzle depth of 491 microns, and the heater analyzed in case 2 of the previous section.

Current is supplied to the heater through electrodes clamped to the top and bottom of the die (see figure 1). The large etched regions on either side of the flow chamber reduce parasitic current paths through the bulk silicon, and direct a larger fraction of the total current through the fins, which are in contact with the flow. The working fluid (Nitrogen) is delivered through Kovar tubing brazed to a stainless steel plate and fixed to the silicon via a glass frit. The plate is clamped with an O-ring to a supply manifold. Further fabrication and assembly details are found in Bayt [2].

EXPERIMENTAL RESULTS

Figure 6 shows the temperature and power with time for a heater operated with no gas flow. The current is increased in a series of discrete steps, as indicated by the dotted line.

The temperature increases at about 25°C/sec immediately after power is applied, and then asymptotes to a steady value during the next 300 seconds. This is due to the increasing power dissipation with temperature. At a current of 6.06 Amps, sufficient power was dissipated to reach the intrinsic temperature (~700°C). For subsequent times the temperature is seen to rise briefly above the intrinsic temperature before dropping back down due to the inherent stability of the heater architecture. This behavior is observed independently of the driving current, which was continually increased.
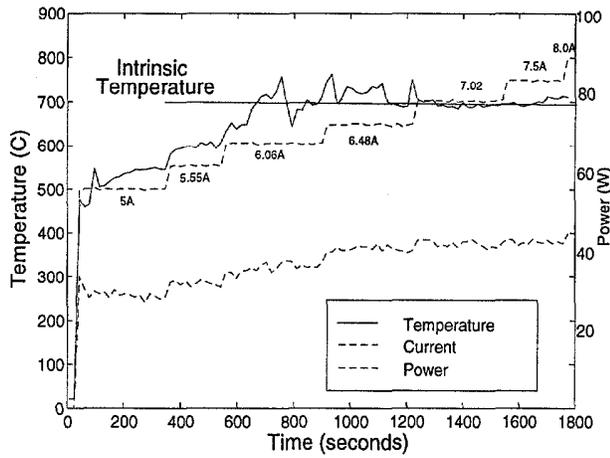
Figure 6. *Heater temperature variation with current. 700°C intrinsic point is observed, as predicted by the doping level.*
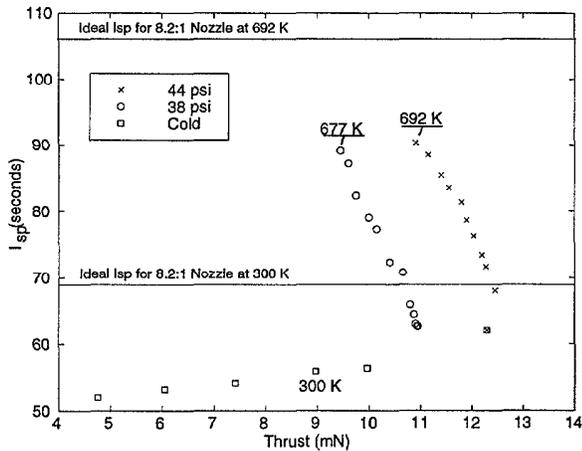
Figure 7. Thrust per unit weight flow (I_{sp}) is an indication of fuel economy, and thrust indicates propulsive authority.

Thrust tests were performed with the thruster depicted in figure 5 to determine thrust and propellant efficiency. The thrust stand was calibrated using known trial weights and had an accuracy of ± 0.5 mN. The entire apparatus was housed in a large vacuum chamber with a background pressure of 50 mTorr.

Figure 7 shows the specific impulse versus thrust for this system operated over a range of heater powers. This illustrates the fuel economy for a given control authority. For fine control, small thrusts are desired, but as shown here, performance decreases for cold flows. For the cold-flow results (shown by square symbols), the thrust is varied by controlling the operating pressure, and hence the mass flow rate and Reynolds number. For two values of the chamber pressure (38 & 44 psia), the pressure is held constant and the heater power is varied by increasing the current through the device. As the temperature rises, the specific impulse rises (as expected) and the thrust is observed to decrease. The thrust is composed of contributions from the exit velocity and pressure:

$$T = \rho U_e^2 A_e + P_e A_e = \frac{P}{RT}\gamma RTM_e^2 A_e + P_e A_e = \dot{m}gI_{sp} . \quad [7]$$

Note that this predicts that the thrust should be independent of temperature. However, as temperature increases, the Reynolds number decreases, and the exit Mach number (M_e), which ideally is only a function of nozzle geometry, decreases due to the

boundary layer blockage effects described in Bayt et al. [6]. For this reason, the observed thrust reduces with temperature, and even more importantly, the I_{sp} relative to the theoretical also decreases significantly. This causes the deviation of the experimental data points in Figure 7 from a vertical line, which would be the ideal for a constant pressure.

Though the I_{sp} is less than the theoretical value, it is 50% larger than that achieved with the cold gas flow. Thus the spacecraft designer trades electrical power for propellant efficiency, which is desirable since power can be replenished via the solar arrays, whereas propellant cannot be re-supplied.

The overall propulsive efficiency (thrust for a given propellant and electrical power) of this demonstration device is quite low (18%) primarily due to parasitic electrical losses in the electrical leads and test structure. However, with optimization of the design and better thermal insulation in the packaging, the efficiency could potentially be raised to 40%.

CONCLUSIONS

The present design has two key features that make it attractive for integrated micro-heat exchangers. First, the use of bulk silicon as both the structural and electrical material simplifies the system architecture and allows for high fluid/heater contact area. Second, the properties of silicon at the intrinsic point are very attractive for stable operation, particularly for gases where thermal runaway can be problematic. By increasing the chamber energy in a microthruster, the mass flow required for a given thrust level is greatly reduced. This translates into increased satellite life for a given propellant supply. Although the device efficiency is low in this initial example, little attempt has been made to minimize parasitic losses through the leads and test structure and one expects the efficiency could be improved in future designs.

ACKNOWLEDGEMENTS

REFERENCES

1. S.W. Janson, "Batch-Fabricated Resistojets: Initial Results". In *International Electric Propulsion Conference*. IEPC-97-070, 1997.

2. R. L. Bayt *Analysis, Fabrication and Testing of a MEMS-based Micropropulsion System* Ph.D. Thesis, M.I.T., 1999.

3. G. L. Pearson, and J. Bardeen, "Electrical Properties of Pure Silicon and Silicon Alloys Containing Boron and Phosphorous", *Physical Review, 75,* 1 (1949).

4. F.M. White, *Viscous Fluid Flow,* 2nd Edition, 1991.

5. R.K. Kakac, S. Shah and A.E. Bergles. *Low Reynolds Number Heat Exchangers.* Hemisphere Publishing, 1st edition, 1983.

6. R. L. Bayt, K. S. Breuer, and A. A. Ayon "DRIE-Fabricated Nozzles for Generating Supersonic Flows in Micropropulsion Systems", *Technical Digest of the 1998 Solid-State Sensor and Actuator Workshop,* Hilton Head Isl., SC, 6/8-11/98, Transducer Research Foundation, Cleveland (1998), pp. 312 - 316.

BRINGING MEMS TO MARKET

Kurt E. Petersen
Cepheid
Sunnyvale, CA 94089

ABSTRACT

MEMS innovations are revolutionizing a number of existing and emerging high technology engineering disciplines. However, many problems and delays have been encountered in the intro-duction of MEMS innovations and potential products into commercial market-places. We present here some practical guidelines to evaluate product marketability and to facilitate the transition of a MEMS concept to commercial viability. These guidelines have been gleaned from an analysis of many successful, as well as some unsuccessful, commercialization efforts, a few of which are briefly described here.

INTRODUCTION

As of the year 2000, we have completed over 25 years of MEMS research, and introduced numerous MEMS products. The field has passed through several phases;

- largely unknown for much of the 70's
- recognized as having a lot of potential in the 80's
- having many commercial success stories in the 90's
- now → becoming a commonly employed and well-accepted solution over a wide range of today's technical industries

Few high tech fields have not been influenced by MEMS; if not through the direct introduction of MEMS products, then through compelling R&D activities and developments which challenge and defy virtually every traditional/conventional mech-anical, opto-mechanical, electromechanical, chemical, and fluidic engineering discipline. The Boeing 747 has 6 million mechanical parts, about 50,000 moving parts. TI's micro-mirror display chip has 1 million moving mechanical parts. Analog Devices' accelero-meter chip detects and accurately controls motions of on the order of nanometers. ACLARA and Caliper microfluidic chips control fluid volumes in the range of picoliters. MEMS has fundamentally enhanced and vastly expanded the limits of the modern practice of engineering.

However, the very revolutionary nature of MEMS has created real difficulties in the transfer of MEMS innovations to commercial products. Most of these problems have been extensively discussed by many workers in the field at many Hilton Head rump sessions: the unique complications of MEMS packaging, the question to integrate or not to integrate, the problem of MEMS foundries, the unique difficulties of MEMS testing, the unique processing techniques of MEMS, the inter-disciplinary nature of MEMS, the comparison of MEMS chips to integrated circuit chips. In the same way that integrated circuit technology has fundamentally redefined the practice of electrical engineering, MEMS is fundamentally redefining many other engineering disciplines. Such momentous changes do not come easily or quickly.

INTRODUCING MEMS PRODUCTS

Introducing a new, revolutionary technology across a variety of markets requires us to carefully consider the commercial accep-tance of that technology from several viewpoints.

One succinct criteria for successfully introducing, manu-facturing, and marketing a MEMS product is to ensure that the MEMS device brings a "compelling, enabling, and rational technology into a *starved market segment*".

This means that the device must provide a compelling function which enables a major performance advantage over existing products, must have a very robust manufacturing process, and must be readily adopted into a very receptive marketplace which is anxiously searching for a solution to the problem which the device addresses.

During the past 25 years of MEMS, this formula for success has been repeated again and again. There have also been examples which did not (or could not) follow this formula, and failed. One frustrating characteristic of the formula is that, many times, a compelling, enabling, rational technology has been forced to wait many years for that very crucial *starving market segment* to develop. Without the starving market, however, the application will either fail, will be only marginally successful, or will be forced to wait longer for the market to develop.

Another factor which greatly influences time to market for an enabling device concept can be characterized as a necessary "convergence of technologies". For example, surface micro-machined optical modulators were first demonstrated almost 30 years ago. However, in order to make use of such an optical element in a practical projection display product, a low-cost manufacturing infrastructure must also be in place for fabricating at least 500K optical modulators on a single chip with near perfect yield. Such a large number of elements is required to achieve a practical display. No manufacturing technology with this capa-bility existed for *any* type of chip until the late 80's (see Figure 1). The convergence of optical modulator array technology and fabrication technologies for producing large numbers of elements on a single chip converged in the early 90's, making it possible to launch products based on TI's micromirror technology.

Another important consideration is that product introduction plans which treat MEMS devices as simple components in larger systems are a mistake. It takes an enormous investment in effort and *time* to design, develop, optimize, and move into production a MEMS chip. In many instances, the complete product system incorporating the MEMS chip is designed, developed, optimized, and moved into production only after the performance character-istics of the chip are finalized. The sequential nature of such product development effort can postpone actual sales for many years. Whenever possible, the MEMS developer/manufacturer should strive to develop the entire product in a parallel fashion, thereby dramatically reducing time to market.
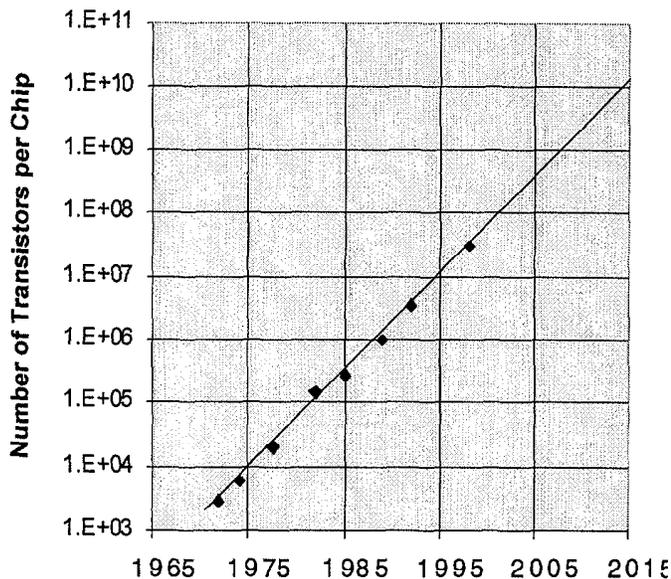
Figure 1. The number of electronic devices per microprocessor chip has followed "Moore's Law" for 30 years. Even though micromechanical optical display elements had been in existence since the 70's, it was not until after about 1990 that the capability of manufacturing chips with sufficient optical elements (about 1 million) to create a viable display product became possible and practical for the first time.

IC-like batch fabrication of MEMS devices is often touted as a major advantage over "one-at-a-time" fabrication. This is a mis-conception for many mechanical devices. Very rapid, fully-automated, one-at-a-time manufacturing can be extremely cost-effective, easily producing 10's of millions of relatively complex parts per year for capital investments of less than $1M. It can be a mistake to lightly dismiss such simple, well-established manu-facturing processes and devices as serious competition to batch-produced MEMS chips.

The implication that market pull dominates successful MEMS products is, of course, only partly true. Market pull is very influ-ential and its importance cannot be under estimated. However, there is no "market pull" without the fundamental progress, promise, and inspiration of advanced technological innovations. Basic research and advanced technical demonstrations always set the stage and establish the expectations for future commercial products.

SUCCESS STORIES (AND FAILURES)

Pressure Sensors

Until about 1982, MEMS pressure sensors were employed only in very limited, high-end, expensive aerospace and industrial applications. During the early 80's, new federal regulations man-dated reduced fuel emissions for automobiles. The solution found by the automotive industry required a pressure sensor to help con-trol the fuel/air ratio. Suddenly, there was a huge demand for low-cost, electronic-based, reliable pressure sensors; a *starving market* was born and MEMS pressure sensors easily dominated the market. Furthermore, MEMS dominance in this important market was readily transferred to other large markets including disposable medical blood pressure sensors and consumer applications.

Accelerometers

The first MEMS acceleration sensors (demonstrated at Stanford and at Westinghouse in the 70's) languished on the shelf for nearly 20 years. As American quality of life continued to grow during the 70's and 80's, however, automobile safety became a dominate political issue. The furious introduction of air-bag crash-protection systems in the early 90's created a *starving market* for low-cost, high reliability, electronic-based crash sensors accelero-meters). These MEMS devices now completely dominate the air-bag market and are spinning off into other areas such as low-cost gyroscopes, navigation equipment, and virtual reality sensors.

Projection Displays

Micromechanical optical modulators have a renowned history going back to the early 70's. Texas Instruments began working on large two-dimensional modulator arrays in the late 70's. This technology barely survived 2 decades of corporate politics, an unimpressed marketplace, and the complete dominance of CRT and liquid crystal displays. The proliferation of computers, however, has generated a *starving market* for high performance, digitally-compatible displays in the late 80's which continues today. The recent success of TI's micromirror display was also helped by a convergence of technologies. Before about 1990, it was impossible to even consider the idea of fabricating cost-effec-tive display chips with the required millions of active elements. Advances in IC fabrication technology, exemplified by the 64 megabyte memory chip, has changed all that. In the early 90's, the convergence of MEMS optical modulation technology with manu-facturing techniques capable of creating millions of elements on a single chip suddenly made the micromirror display commercially feasible.

Ink Jet Nozzles

First developed at IBM in the 70's, it took about 20 years for micromachined ink jet nozzles to become nearly commodity items. Today, over 100M ink jet chips are produced every year. Despite the belief that we have been supposedly moving toward the "paperless" world for 30 years, the market is still starving for simple, low-cost printers with more and more demand for high quality. In fact, the only practical technology for achieving the demands of low-cost, photographic-quality, color printers for today's graphics-oriented computers, continues to ink jet printing.

Chemical Sensors

Throughout many years of development during the 70's and 80's, ISFETs and CHEMFETs suffered from serious performance problems caused by drift and instability. Finally, companies such as i-STAT recognized a *starving market* for chemical sensors and took practical advantage these unstable devices. Instead of trying to solve the intractable problems with drift and instability, i-STAT created a business in which the MEMS sensors were very low-cost, used only once, then disposed of. While the i-STAT technology is a model for efficient and elegant integration of MEMS and plastic packaging, it took i-STAT over ten years to convince medical practitioners that the market was starving for fast, disposable blood chemistry sensors. There are still many other market segments for such devices and increased developments in microfluidic technol-ogy will allow further integration and greater demand for chemical analysis and chemical sensors. Since many of the applications for this technology are medical and biomedical, the move toward dis-posable sensors and microfluidic systems will grow dramatically.

Fiber Optic Switches

Since the dawn of MEMS, numerous workers in the field (the author included) have continuously predicted that this technology would become important for optical fiber switching devices. Suddenly, within the last year, this application area has exploded with intense activity. Why? The basic MEMS technologies now being applied to fiber switches have been around for over 10 years. What has changed is the market. With the advent of the internet, cable television, and video-on-demand, the requirement for communication bandwidth has become insatiable. We are now experiencing a *starving market* for any technical innovations which will enable the implementation and rapid expansion of fiber optic communication systems.

Micro-valves

At the other end of the spectrum for MEMS success stories is the elusive micro-valve application and product. Numerous types and versions of MEMS micro-valves have been demonstrated, developed, and introduced into the market during the past 15 years or longer. None of these have become notable successes because the market is not starving for such products; numerous low-cost traditional valves are readily available which perform perfectly adequately for every conceivable application today. Are there any compelling technical reasons why a MEMS valve should be preferred over a conventional valve? Of course, there are excep-tions and these exceptions can guide us to future successful valves. The Microsensor gas chromatograph system, for example, employs an array of micro-valves in a system where low dead-volume is a crucial performance parameter. In this case, the accomplishment of a high density, integrated array of valves provides the unique operating advantage required in this application. MEMS micro-valves can provide extremely important cost and performance advantages when used in arrays. In the future, in market applica-tions which require highly integrated fluidic systems, the specific advantages of MEMS valves will become valuable. As of today, however, MEMS micro-valves, independent of such highly integrated fluidic systems, have not yet become commercially successful.

Microelectromechanical Switches

The first microelectromechanical switches were fabricated at Westinghouse and IBM in the 70's. Since then, numerous un-successful efforts have been undertaken to commercialize these devices. The lack of success to date has been the result of two problems. Most importantly, there is not a starved market for electromechanical switches. Today, these devices are basically commodities. Numerous, high-reliability, low-cost devices are readily available from many suppliers. (Most of these are manufactured with highly automated on-at-a-time assembly lines.) MEMS switches do not provide a compelling incentive for replacing conventional relays in traditional applications for relays. At the same time, the most common applications for conventional relays are for the switching of relatively high DC currents -- on the order of 1 ampere or more. It has been extremely difficult to design and demonstrate MEMS switches capable of reliably controlling such high-current DC levels. In recent years, however, there has been a trend to design MEMS switches specifically for RF applications. This new application area may allow the MEMS switch, invented over 20 years ago, finally successful. In addition, new applications for arrays of electromechanical switches may provide the basis for successful MEMS switch products in the future.

These examples do not mean that every application is driven solely by market pull and that research in MEMS does not also provide a high level of technology push. Indeed, no product would ever be possible if the early research and feasibility phases were not conceived, initiated, and successfully completed. However, no product will ever become commercially successful without

- the proven, compelling concept
- the enabling device performance capability
- the rational manufacturing process established
- and the *starving market* application

THE FUTURE

So, can we look into the future and anticipate some of the MEMS success stories over the next 10 years? One interesting exercise we can do is extrapolate what has been happening over the past 10-20 years with computing power and data recording density, two powerful manufacturing disciplines which drive advances in modern technology. From Figures 1 and 2, we can expect factors of improvement in these areas of between 100 and 1,000 over the next 15 years. Memory chips can be expected to hold a quantity of data equivalent to yesterday's hard disk drives. If current trends continue, transistors will average less than $0.2\mu m$ on a side. If current trends continue, hard disk drives (currently at about 15 bits/μm^2) will approach 5×10^{13} bits/in^2, or 100K bits/μm^2, or 10 nm^2/bit -- that's 3nm or 30Å on a side) ! This implies that two million elements (electronic OR mechanical) photolithographically produced on a single chip, will be trivial. On the other hand, accurate location and tracking of individual bits of data will require extremely high precision mechanical devices. All of these trends indicate an increased *requirement* for MEMS devices (such as micromechanical actuators) as well as an enhanced *ease of manufacture* of micromechanical arrays (such as MEMS displays, or actuator arrays). The future of MEMS is promising in many ways and will continue to grow along with the expected progress in photolithographic fabrication technology.

Microfluidics Related to Biomedical Applications

A significant revolution is currently underway in the develop-ment and application of microfluidic techniques to a variety of medical, biomedical, and chemical processing and delivery requirements. In a sense, the ink jet nozzle was the initial predecessor of this new microfluidic field. However, microfluidics is only a single area in a much broader and more influential movement in which integrated circuit based technologies are being increasingly applied to biomedical and chemical applications. The Affymetrix Gene chip and other DNA hybridization arrays are a part of this effort. This trend will continue to have an enormous impact on the future of biomedical and biochemical development as well as every other possible area of bioanalysis and diagnostics. Not only will MEMS technologies be employed in pharmaceutical development and genomic discovery, but also in the practical implementation of this knowledge on a daily basis for individual patients in a hospital or a doctor's office. For example, future therapeutic drugs will only be prescribed to patients who have been genetically tested to ensure that they have the proper genetic profile to respond efficaciously to the drug. This new area is called pharmacogenomics. In a few years, it will be common to have your blood genetically analyzed in your doctor's office, by MEMS devices, within just a few minutes, to determine your

genetic compatibility with a variety of drugs, allowing the physician to appropriately prescribe the correct drug for you. These markets are *starving* for fast, miniaturized, automated techniques for performing these and other bioanalytical functions.
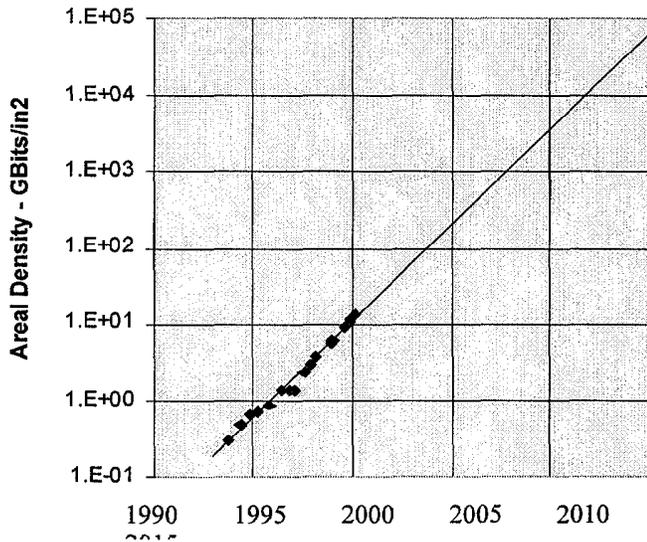
Figure 2. *The areal density of magnetic recording products continues to increase exponentially in the same way as the number of transistors/chip, predicted by Moore's Law.*

Data Storage

Continued expansion of the internet and demand for video is driving storage technology. As fiber optic and other high band-width capability becomes increasingly common, demand for storage will continue to grow exponentially for many years. This market is always *starving* for higher data density and lower cost. Because the current technical approaches have not yet reached their technical limits, MEMS has yet to make a commercial impact on this ubiquitous technology – no *starving market*, yet. However, it is virtually inevitable that future data storage devices will employ MEMS devices for various purposes: micro-positioning and tracking devices for accurately accessing sub-micron magnetic, optical, or surface profile (AFM) data tracks; micro-mirrors for scanning and positioning of optical beams; micro-beams for scanning and controlling Atomic Force Microscope (AFM) recording and reading tips.

Microsurgical Instruments

Many advances in surgery are guided by the trend toward non-invasive techniques, intravascular devices, and laprascopic procedures. It may be possible for MEMS to contribute to this important and growing technical direction. Microsensors, scan-ners, surgical devices, and implantable devices are all biomedical candidates for MEMS development.

RF Communications

As a essential feature of the movement toward absolute portability, wireless communication has many opportunities for improvements, enhancements, increased electronic integration, and cost reductions. One recent start-up has begun to design dispos-able cell-phones – like disposable cameras. Although much work and many innovations are still needed, resonant MEMS structures have the potential to replace conventional electronic filters and tunable frequency standards. Tunable mechanical filters, integ-rated with high

frequency electronics could achieve the ultimate goal of a single RF transceiver chip. MEMS electronic switches, conceived and demonstrated almost 25 years ago, are already close to commercial implementation in RF applications. For many years, the development of these electromechanical devices were targeted toward traditional DC switching applications. However, DC switches are very low-cost commodities without a *starving market*. Current RF switches, on the other hand, are expensive components with sub-optimal performance, which are increasingly in demand for the *starving* RF communications market.

Energy Storage

All the full-function capabilities of future computation and communication must be completely portable. The primary limit-ation to this portability is energy storage. Today's batteries are simply inadequate and will become even more problematic. In addition, the long-anticipated mobile robotic machines of the future will not be possible without quantum improvements in mobile power generation and storage. Orders of magnitude increases in energy storage will revolutionize technology. Can MEMS contribute to such a revolution? Aggressive and forward-looking projects such as the MIT microturbine device have the potential to create power sources more than 10 times more efficient in weight and size than the best batteries. However, other (possibly simpler) approaches using integrated gas and liquid fuel-flow systems together with micro-electric generators might be possible. Energy storage is a market which is always *starving* for improvement.

Displays

Cell phones, PDAs, and other hand held electronic devices have terrible displays today. Even the impressive displays on lap-top computers are power hungry and perform very poorly in sunlight. Promising micro-displays which fit into glasses or goggles have been developed and are close to commercialization. The demand for enhanced performance in every display market creates a demanding, constantly *starving market* for new and improved approaches. MEMS will continue to contribute to all these display applications, micro-displays, projection displays, as well as for new, effective direct view displays.

ECONOMIC CONSIDERATIONS

In bringing new MEMS developments and products to market, it is crucial to understand the economic implications of our guiding criteria - "applying a compelling, enabling, and rational technology into a *starved market segment*". The first obvious economic axiom is that we must be able to manufacture the device and sell it at a cost which is consistent with market requirements. Lets look at an example. Suppose we have a product concept which we anticipate can be totally developed and moved to production for $10M (including the cost of the manufacturing equipment and facility). This is not an unreasonable amount – 15 or 20 people working for 2 years. Simplistically, in order to make a profit on this product and grow the company, the part must be sold for a price which includes the materials and manufacturing costs, plus overhead, plus a profit, plus amortization of the $10M investment. If sales are only 200K parts/year, at least $10 in amortization costs must be added to the price of each part. If sales are 2M parts/year, about $1 must be added to the price of each part. The amortization costs should only be about 10-20% of the price of the part.
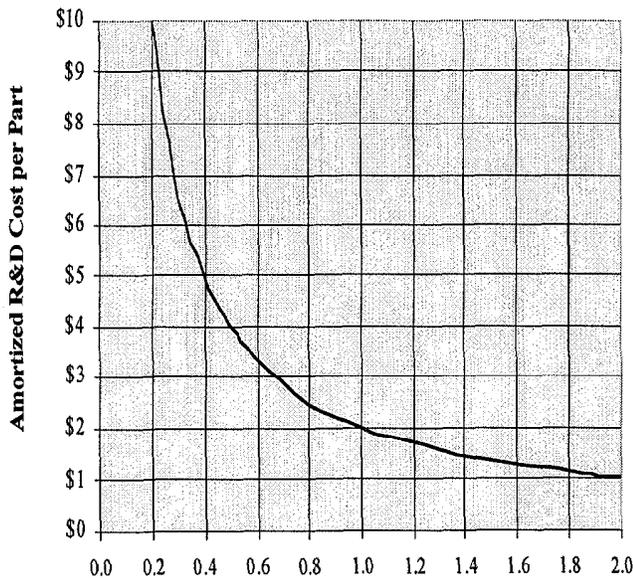
Figure 3. *This graph shows the additional amortization cost which should be added to the cost of a product depending on the number of parts sold per year. It assumes that the total costs for product development and production facilities are $10M and the amortization period is 5 years.*

There is an important lesson here. If the required price target for the part is on the order of $10, sales quantities must be more than about 1-2M parts/year to justify the development costs. At the same time, the target price must be on the order of $100 if the sales quantities are only about 100-200K parts/year. For these reasons, low- and medium-cost, commodity-type, replacement products should be avoided.

CONCLUSIONS

The opportunities for developing and marketing successful products, based on MEMS technologies, are enormous. The infra-structure now in place for MEMS, including university research and education programs, equipment manufacturers, modeling and design software, foundries and services, government R&D funding, conference and journals, acceptance by venture capital funding groups, all provide a technical and informational foundation which is invaluable to MEMS entrepreneurs. This large base of support, together with the current high degree of commercial success, acceptance, and credibility of MEMS is making it easier to bring important new MEMS products to the market.

However, the past 25 year history of MEMS products suggest some current guidelines which will improve the prospects for the successful introduction of new products based on MEMS. As the field changes and progresses, the guidelines will change. As research continues, innovative new devices and new applications will be conceived. It is an exciting time for a potent and dynamic technology.

DEPENDENCE OF TiO$_2$ FILM SIZE AND SHAPE FOR LIGHT ACTIVATED FLOW CONTROL IN MICROCHANNELS

Jaisree Moorthy[1,2], **Christopher Khoury**[1,3], **Mark A. Stremler**[1], **Jeffrey S. Moore**[1], **David J. Beebe**[1,2]

[1] Beckman Institute for Advanced Science and Technology, University of Illinois, Urbana-Champaign, IL 61801
[2] Department of Biomedical Engineering, University of Wisconsin, Madison WI 53706
[3] Department of Electrical Engineering, University of Wisconsin, Madison WI 53706

ABSTRACT

The surface charge of Titanium Dioxide (TiO$_2$) can be varied with light to control the electroosmotic flow velocity in microchannels. The use and limitations of TiO$_2$ films inside microfluidic channels to control both local and bulk flows are explored. Different patterns of TiO$_2$ can be fabricated over which the surface charge can be modulated. Both modeling and experimental results are presented. In channels containing long patterns, bubble formation was observed due to electrochemical reactions. A threshold potential drop for bubble generation varied from 700 V for a 2.5 cm long film to 1000 V for a 0.5 cm film.

INTRODUCTION

Modulation of local velocity in electroosmotic flow driven systems is useful in applications like mixing at low Reynolds number and improving resolution in capillary electrophoresis. Other workers have reported the use of radial fields to influence the local velocity [1], [2]. However, this is limited by fringing effects [3]. Recently, we have reported active control of the zeta potential with light using TiO$_2$ films [4].

TiO$_2$ is a direct semiconductor, and its surface charge can be modulated by irradiating the interface with light of appropriate wavelength. Patterns of TiO$_2$ can be fabricated giving a definite region over which the surface charge can be controlled. Pattern geometry is limited by electrochemical reactions occurring on the film. Here we describe preliminary experimental and computational work to characterize the parameter space of light modulated zeta potential control.

BACKGROUND

Irradiating a direct semiconductor forms electron-hole pairs (Figure 1) that can then undergo oxidation, reduction or recombination reactions. The relative rates of these reactions determine the nature and magnitude of change in the surface charge. Since such changes can be affected dynamically, the electroosmotic flow velocity can be actively modulated.

Based on this concept, Bard and coworkers [5] have shown that the velocity of TiO$_2$ particles is affected by irradiation with UV light. Motivated by their work and the fact that TiO$_2$ is robust in aqueous solutions, we have used TiO$_2$ films in our study. Preliminary results [4] showing both bulk and local flow control with TiO$_2$ patterns have been demonstrated. In previous work, the bulk velocity was measured using current monitoring techniques in the presence and absence of UV irradiation in different buffer solutions. The change in the velocity was seen to vary with pH. The largest velocity change was at pH 4.7 with smaller changes observed at higher and lower pH values. To modulate the local flow, staggered patches

of TiO$_2$ (500 μm square) were patterned in the microchannel. Upon irradiation with UV light, bending of the flow towards the pattern was observed (Figure 2). The limitations imposed by the
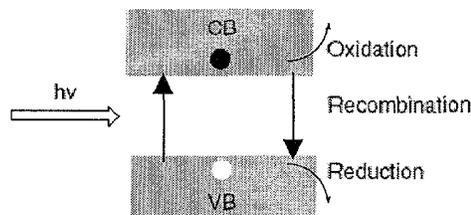
Figure 1 An electron hole pair is formed upon irradiation of a direct semiconductor. The change in the surface charge depends on the relative rates of the oxidation, reduction and recombination reactions.

electrochemical reactions and flow fields on the size and continuity of the pattern are reported in this paper.
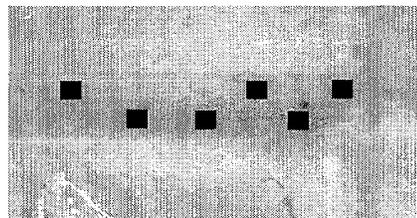
MODEL STUDIES

Figure 2 Flow bending towards the pattern (shown as black squares) due to UV irradiation [4]

Effect of the Pattern length on Electrochemical reactions

A TiO$_2$ film can be prepared from the oxidation of Ti. Incomplete oxidation results in a Ti/TiO$_2$ interface. The films can be modeled along with the electrolyte and electrical double layer at the interface as a combination of resistors and capacitors. For the experimental conditions used here, a simple resistive model is proposed (Figure 3). When there is a defect in the TiO$_2$ film or when the film is contaminated (i.e. not a perfect insulator), the resistance R$_{TiO2}$ is lowered and some current flows through the Ti/TiO$_2$ layer. This creates a potential drop across the film. This is a function of the length of the film. When this drop is large enough, electrochemical reactions can occur resulting in the generation of bubbles and, in some cases, degradation of the film.
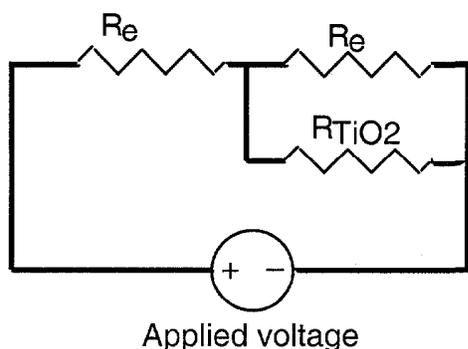
Figure 3 *Simplified circuit model equivalent for fluid filled channels with TiO₂ films.*

with the aqueous electrolyte, the PDMS was made hydrophilic by treating the channel with oxygen plasma. A DC potential was applied across the channel through Pt electrodes placed at the wells. The potential was incremented in steps of 100 V (range 100 V – 3000 V). The applied potential at which the bubbles

Figure 5: *Channel designs with different size TiO₂ patterns (black). The channels were 1000 μm wide, 20 μm deep and 4 cm long. The length of the patterns is indicated in the figure.*

first occur was recorded.

Effect of Pattern Geometry on Flow Field A computational grid was generated on a 2000 μm long section of the channel (1000 μm wide x 20 μm deep), and the velocity field in this domain was computed numerically with the CFD-ACE package (CFD, Research Corporation). An infinite channel was modeled by enforcing periodic boundary conditions at the inlet and exit of the computational domain; *i.e.* by requiring the velocities and the pressure profile to be identical at the inlet and exit. A flow of mean velocity 5400 μm/s was created in the channel by maintaining a pressure difference (•P) between the inlet and the exit. Varying •P controls the net flow rate through the channel. No-slip boundary conditions were assumed on all channel walls. To simulate the TiO₂ patterns a wall velocity of 242 μm/s (obtained from bulk flow experiments) was assumed on the bottom side of the channel. Two patterns, one with staggered 500 μm square patches and the other with a meandering continuous strip were studied (Figure 4). The wall velocity on all other parts of the channel was assumed to be zero.

EXPERIMENTAL

Fabrication A channel was etched on Pyrex 7740 wafers. Inside the channel, TiO₂ films of varying lengths (2.5 cm, 1 cm.

Figure 4 *The TiO₂ patterns (shown as black blocks) that were modeled a) staggered patches b) meandering continuous strip*

0.5 cm, 0.25 cm. 500 μm) were prepared by sputtering (DC magnetron, 1 kW) Ti followed by thermal oxidation (450 °C for 2 h). Polydimethysiloxane (PDMS) slab with holes punched for the inlets was used as the top cover to seal the channels.

Procedure The electrolyte (50 mM Sodium Phosphate buffer at 4.7 pH) was degassed before the experiment. To facilitate filling

RESULTS AND DISCUSSION

Constraints due to electrochemical reactions We tested different pattern lengths over a range of voltages. Preliminary results show that above a certain applied potential, electrochemical reactions occur. This was observed to vary with the film length (Figure 6). From the simplified circuit model presented earlier, a potential drop is generated across the film. When this crosses a threshold, the film acts as an electrode for the electrochemical reactions. For a film of a given quality, the threshold is reached at lower applied potentials as the length increases. In previous studies of local flow control [4] with small 500 μm patterns, no reactions were observed.
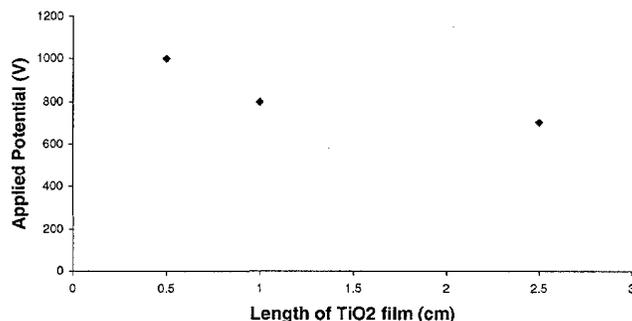
Figure 6 *Applied potential at which the threshold potential is reached for the different film lengths*

The possible electrochemical reactions occurring on the film are:

Cathode

$$2H_2O + 2e^- \rightarrow H_2 + OH^- \qquad\qquad 3$$

$$2H^+ + 2e^- \rightarrow H_2 \qquad\qquad 4$$

Anode

$$H_2O \rightarrow (O)_g + 2H^+ + 2e^- \qquad\qquad 5$$

$$Ti^{2+} + 2H_2O \rightarrow TiO_2 + 4H^+ + 2e \qquad 6$$
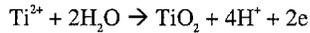
The electrolysis of water (equation 3 and 5) results in evolution of bubbles (see Figure 7). Further oxidation of partially oxidized TiO_2 also results in bubble formation (equations 4 and 6). Once these regions are oxidized, the threshold potential for the electrolysis is increased. This is because the resistance of TiO_2 film (R_{TiO2} in Figure 3) increases with continued (complete) oxidation. Salt deposition was observed on the film. This could be due to the consumption of the electrolyte.
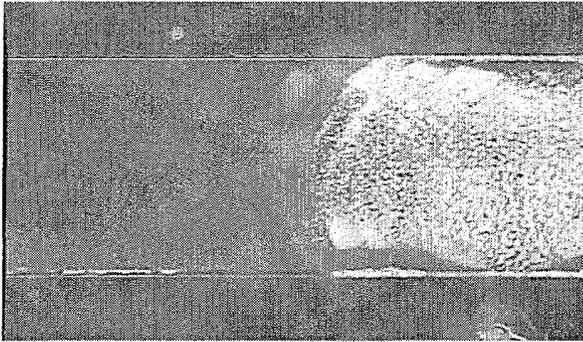
Figure 7 *Bubble evolution on the TiO_2 film due to electrochemical reactions*

In all cases, bubble formation was observed. However, there was no apparent degradation of the TiO_2 film except at the edges, where the largest surface defects of the oxide film exist and adhesion to Pyrex maybe poor. However, earlier studies with films prepared by evaporation of Ti followed by electrochemical oxidation resulted in degradation of the film at lower applied potential (500 V). This could be due to lower resistance (R_{TiO2}) due to defects or contamination

Influence of pattern geometry on flow fields

The computational models show that the presence of the pattern produces spanwise flow, both in the horizontal and vertical directions. Figures 8 and 9 show the velocity fields generated by the patterns. Flow in the horizontal direction is produced at the corners of the pattern. For the continuous strip geometry (Figure 8b) this sets up localized circulation of the fluid at the corners; whereas the patches (Figure 8a) produce a bending effect similar to that observed experimentally (Figure 2). The magnitude of the maximum horizontal velocity is almost double for the patches (see Table 1). The edges of the patterns normal to the streamwise flow direction produce vertical velocity fields (Figure 9). The magnitude of the velocity in this direction is greater for the patches.
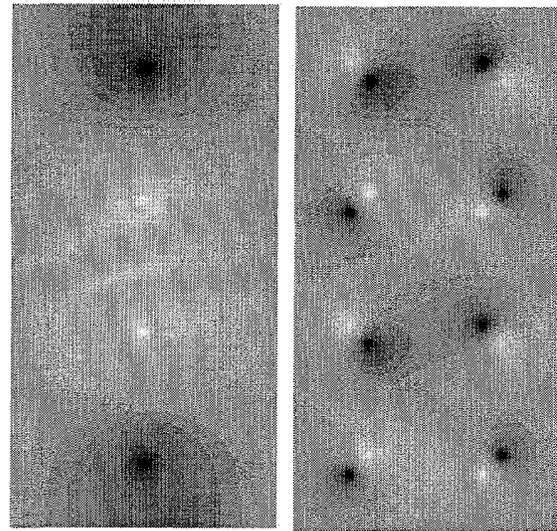
Figure 8 *Horizontal flow fields (at one half channel depth) generated by a) patch and b) continuous. The black regions indicate flow to the left while the white regions indicate flow to the right.*
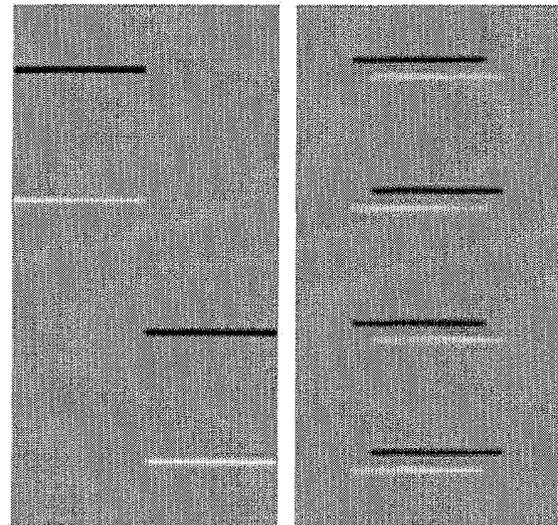
Figure 9 *Vertical flow fields (at one half channel depth) generated by a) staggered patch geometry and b) continuous The black regions indicate flow into the plane while the white regions indicate flow out of the plane.*

Table 1 Maximum spanwise flow velocity generated by the pattern.

Geometry of the Pattern	Horizontal Velocity (um/s)	Vertical Velocity (um/s)
Continuous Strip	66	30
Patches	110	30

Design Criteria for Light Modulated Control

The simulation studies show that the edges and corners of the pattern generate flow in the vertical and horizontal directions respectively. By changing the pattern geometry, different types of effects can be produced for various applications. For example, by increasing the number of corners, localized flow can be produced which would help in mixing. If desired, the flow can be routed by placing the patterns appropriately (Figure 8a). Since the likelihood of the electrochemical reactions occurring depends on the length of the film, a long pattern can be designed by using smaller patches.

CONCLUSIONS

In order to control the electroosmotic flow using light, TiO_2 films can be patterned in the channels. Here, the effect of the geometry of the pattern was explored. A microchannel with different lengths of TiO_2 films was fabricated to study the limitations on the length of the patterns. We observed electrochemical reactions occurring on the TiO_2 films initiated at a threshold voltage. This varied as a function of the film length. A circuit model has been proposed to explain the reactions. CFD flow simulations show that the corner and edge of the pattern produces spanwise flow in the horizontal and vertical directions. These experimental and compuational results provide a design space in which one can design geometries to achieve specific patterns of flow control. The use of light to turn the flow control on and off provides a simplified mechanism for both local and bulk electroosmotic flow control in microchannels.

ACKNOWLEDGEMENT

1. Hayes, M.A.; A.G. Ewing, *Electroosmotic Flow Control and Monitoring with an Applied Radial Voltage for Capillary Zone Electrophoresis*. Anal. Chem., 1992. **64**: p. 512.

2. Lee, C.S.; W.C. Blanchard; C.T. Wu, *Direct Control of the Electroosmosis in Capillary Zone Electrophoresis by Using an External Electric Field*. Anal. Chem., 1990. **62**: p. 1550.

3. Schasfoort, R.B.M.; S. Schlautmann; J. Hendrikse; A. van den Berg, *Field-Effect Flow Control for Microfabricated Fluidic Networks*. Science, 1999. **286**: p. 942-945.

4. Khoury, C.; J. Moorthy; M. A. Stremler; J. S. Moore; D. J. Beebe, *TiO₂ Surface Modifications for Light Modulated Control of Flow Velecity*. in *Micro Total Analysis Systems*. 2000. University of Twente in Enschede, The Netherlands.

5. Dunn, W.W.; Y. Aikawa; A.J. Bard, *Characterization of Particulate Titanium Dioxide Photcatalysts by Photoelectrophoresis and Electrochemical Measurements*. J. Am. Chem. Soc., 1981. **103**: p. 3456.

A HIGH FREQUENCY HIGH FLOW RATE PIEZOELECTRICALLY DRIVEN MEMS MICROPUMP

H.Q. Li, D.C. Roberts, J.L. Steyn, K.T. Turner, J. A. Carretero, O. Yaglioglu,
Y.-H. Su, L. Saggere, N.W. Hagood, S.M. Spearing, and M.A. Schmidt
Massachusetts Institute of Technology
Cambridge, MA 02139

R. Mlcak
Boston Microsystems, Inc.
Woburn, MA 01801

K.S. Breuer
Brown University
Providence, RI 02912

ABSTRACT

A piezoelectrically driven fluidic micropump was fabricated and tested. Microelectromechanical systems (MEMS) fabrication technology was used. Small cylindrical piezoelectric material elements were integrated with microfabricated silicon, silicon on insulator (SOI), and glass chips using eutectic bonding and anodic bonding processes. SOI wafers were used to form the thin membranes within the moving parts (a drive element and two passive valves). The design, microfabrication process, and assembly of the device are described in this paper. Fabrication issues such as control of fillet radii at the feet of the Si membranes for stress reduction and simultaneous eutectic and anodic bonding were discussed. A flow rate as high as 3000 µl/min was recorded. Experimental and simulation results of the dependences of the pump flow rate on the voltage and frequency applied on the piezoelectric material are shown and discussed.

INTRODUCTION

This paper presents the fabrication and testing of a high pressure, high flow rate piezoelectrically driven fluidic micropump with passive valves. In comparison to other piezoelectrically driven micropumps in the literature [1-4], this device incorporates the integration of bulk piezoelectric material with MEMS microfabrication techniques to allow for high structural driving frequencies in the tens of kHz range and implements novel methods of anodic and eutectic bonding to join silicon, glass, and piezoelectric components, enabling minimal device mass and therefore high power density. It requires a compact design that brings about challenging fabrication issues such as multi-layer wafer bonding, multiple lithography and deep reactive ion etch (DRIE) steps for the SOI wafers, use of nested oxide masks, and integration of small piezoelectric material elements. The need for high frequency and high flow rate calls for precisely micro-machined Si membranes capable of large deflection as the moving components in the pump chamber and the passive valves of the micropump. SOI wafers are good choices for this purpose because they have a uniform thin Si layer, their buried oxide can be used as the etch stop during DRIE processes, and the fillet radii of the membranes that are critical to the strength of the membranes are easier to control in the presence of the buried oxide. Another advantage of SOI wafers is that separate electrodes can be formed in the SOI layer during microfabrication, with the buried oxide as the insulator. Eutectic bonding has been used to integrate the piezoelectric elements into the MEMS devices. For this micropump it is necessary to choose low melting point eutectic bonding alloys so that the eutectic bonding can be performed

simultaneously with the anodic bonding of the Si and SOI layers to the glass layer that encloses the piezoelectric material for insulation. The successful fabrication of the micropump is a major step toward the realization of high power density transducer devices for micro-robotic actuation and heel-strike power generation applications [5, 6].

MICROPUMP STRUCTURE

The micropump has a multi-layer structure, schematically shown in Figure 1. It includes one fluid pump chamber and two identical passive check valves [7]. The volume change in the fluid chamber is achieved by the vertical motion of the double layer cylindrical piston-like drive element (layers 4 and 5) attached to annular Si membranes and driven by a ceramic piezoelectric cylinder (layer 3) 1 mm in diameter and 1 mm in height. When a voltage is applied to the piezoelectric cylinder, it pushes the drive element piston up, producing a high pressure in the fluid chamber that closes the inlet valve and opens the outlet valve so that fluid can flow out of the fluid pump chamber. Next, when the voltage is
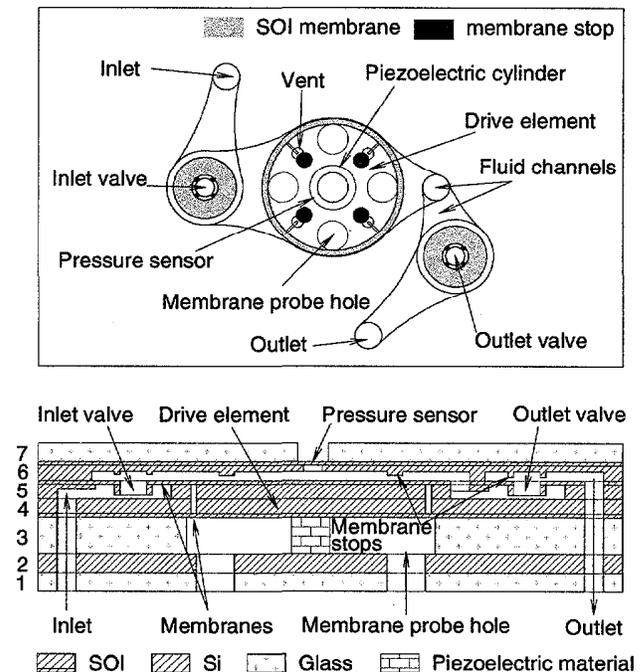
Figure 1. Top view and cross-section view of the micropump. The cross-section view is from a zig-zag cut to show all the features.

reduced the piezoelectric cylinder contracts and the drive element piston moves down, resulting in a low pressure in the fluid chamber that closes the outlet valve and opens the inlet valve so that fluid can flow into the chamber. The pump works by repeating these steps at high frequency.

The drive element piston is 3.6 mm in diameter and each of the two SOI layers 4 and 5 forming the piston is about 400 μm thick. The annular drive element SOI membranes are 150 μm wide and 10 μm thick in the lower layer and 15 μm thick in the upper layer which also includes the valves. SOI wafers (layers 4, 5, and 6) were chosen to construct the membranes in the device for two reasons: 1) to control precisely the membrane thickness by using the buried oxide in the SOI as an etch stop in the DRIE process, and 2) to use the lower SOI (layer4) as an electrode to the piezoelectric cylinder and the buried oxide as an insulator between the piezoelectric cylinder and the remaining part of the micropump. The thickness of the SOI membranes in layer 5 was determined mainly by the design specifications of the passive valves that were subject to much larger deflections than the drive element piston membrane. Venting channels in layer 5 and holes in layer 4 (not shown in Fig. 1) were formed to release pressure in the cavity between the drive element membranes during the wafer diffusion bonding. The passive valves consist of a 0.5 mm center hole and a 15 μm thick annular SOI membrane with an inner diameter of 0.70 mm and an outer diameter of 1.65 mm. When not deflected the valves are only 0.5 μm from layer 4 to minimize back leaking.

The fluid chamber above the drive element piston and fluid channels to and from the valves are located in layer 6. Small cylindrical motion stops about 8 μm from the drive element and about 18 μm from the valves respectively were also formed in layer 6 to prevent the membranes from excessive deflections. A pressure sensor in the shape of a circular membrane 0.8 mm in diameter and 20 μm thick was formed in the fluid chamber for on-site pressure measurement. A ceramic PZT-5H piezoelectric cylinder 1 mm in diameter and 1 mm in height was incorporated in the middle glass layer 3 as the active drive component of the micropump. The Si layer below the piezoelectric cylinder was used as the other electric contacts to the piezoelectric cylinder. 3 mm thick borosilicate glass was used in the top and bottom layers 1 and 7 for structural strengthening of this bench top device.

MICROFABRICATION PROCESSES

To minimize the size of the device, a compact design requiring a minimum number of layers, efficient use of the both surfaces in each layer, and therefore multiple lithography steps for the SOI wafers was implemented. As an example, Figure 2 shows the fabrication process flow for layer 5. The first step is the selection of the SOI wafer that has a 15 μm thick SOI layer, a 0.4 μm thick buried oxide, and a 380 μm thick handle layer. Secondly, positive photoresist was coated on both sides of the wafer and the SOI side was patterned using standard lithography procedures. The photoresist on the handle side is to protect the handle surface from scratching in the lithography steps. This back surface protection method was used throughout the micro-fabrication processes. The wafer was then put in a DRIE machine until the SOI layer was etched through. This step formed the fluid through holes in the SOI layer. Next, the wafer was cleaned and coated with positive photoresist on both sides again. The handle side was patterned with photoresist and etched down about 0.5 μm in a RIE etcher to form the clearance between the valves and their

seating surfaces. In the fourth step the wafer was oxidized at 1100°C for 5 hours to form 2 μm of thermal oxide that is to be used as a nested mask and back surface protection from DRIE etching later on. The thermal oxide was etched in step 5 to form the nested mask and then another lithography process was performed on handle side to define the membranes in the drive element and in the passive valves in step 6. The wafer was etched first by DRIE about halfway though its thickness with photoresist as seen in step 6. Then the photoresist pattern was stripped and the nested thermal oxide mask was exposed for another DRIE step until it stopped at the buried oxide, as seen in step 7. The nested oxide mask was necessary to construct the 200 μm deep venting channels in the drive element and the fluid channels to the valves. Finally, the thermal oxide was stripped using BOE wet etching.

1. Start with a SOI wafer.

2. DRIE SOI and BOE buried oxide.

3. DRIE handle side 0.5 μm.

4. Grow 2 μm thermal oxide.

5. Pattern the oxide film with BOE.

6. BOE and DRIE handle about 200 μm with photoresist pattern on .

7. Strip photoresist and DRIE to the buried oxide with nested oxide mask.

8. Strip the oxide on both sides.

▨ Si   —— Thermal oxide   —— Buried oxide
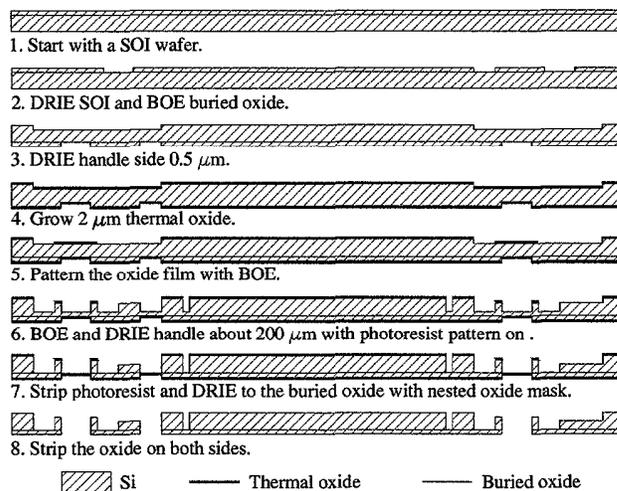
Figure 2. Fabrication flow of an SOI layer (layer 5 in Fig. 1) with the passive valves and half of the drive element piston.

Since all the membranes in this device have to operate under high stress and at high frequency, it is important to have a membrane structure with minimal stress concentration. A significant challenge in the micropump fabrication was controlling fillet radii in the range of 20-40 μm for stress reduction at the feet of the membranes. In the DRIE machine Si is etched at a rate of about 2-4 μm/min. depending on the feature size and exposed Si area. The etch front is usually a gradual curve in cross-section but varies depending on the etch conditions. An SOI membrane fillet radius is formed after the etch front reached the buried oxide and then the Si edges retreat horizontally to the side wall at a rate of about 70-100 μm/minute as more oxide is exposed. There is a time window of only about 2-4 minutes to control the fillet radius in the desired range after about 2 hours etch of the handle side. Therefore careful monitoring in the final stage of the DRIE is critical. Figure 3 shows a 150 μm wide drive element membrane with a gradual fillet about 25 μm long horizontally.

ASSEMBLY OF THE DEVICE

After the final DRIE etch and cleaning, the three SOI wafers were aligned using an Electronic Visions system and fusion bonded at 1100°C in nitrogen for 1 hour. The bonded wafer stack was cut into 16.7 mm by 15.7 mm chips with a die-saw, cleaned in water, acetone, methanol, and isopropanol successively, and finally dried

in air. The glass wafers were processed with diamond drills and die-saw cut into small chips. All the interfaces between Si and glass were anodically bonded at about 300°C. The Si and the piezoelectric material interfaces were eutectically bonded.
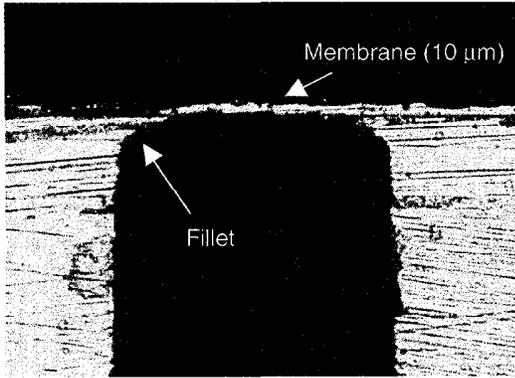
Figure 3. Optical cross-section picture of a drive element membrane, the membrane thickness is about 10 μm.

To prepare for the eutectic bonding, 500 Å of Ti film and 5000 Å of Pt film were deposited on the Si surfaces through a shadow mask by e-beam and 2 μm of an Au-Sn (80-20% wt.) alloy films were deposited on the piezoelectric surfaces by sputtering. The composition of the metal films were chosen for surface wetting (Ti) and low eutectic bonding temperature (Au-Sn) so that the eutectic bonding and the last anodic bonding can be done simultaneously as required by the pump design. The height of the piezoelectric cylinders were designed to be 2-3 μm larger than the height of the middle glass piece (layer 3 in Fig. 1) so that the drive element piston is preloaded and when the chamber is expanded there is no pulling of the drive element piston by the piezoelectric cylinder that might delaminate the eutectic bonding. For this reason, an oxide mask that defines the piezoelectric seating was left on each Si chip below the piezoelectric (layer 2 in Fig. 1) for height compensation by individually trimming etch of the Si chips. Finally wires were soldered on the Ti/Pt electrode pads that were deposited at the same time as the eutectic bonding films were deposited. Figure 4 shows a picture of an assembled micropump.

The key components of this bench top device were built in an area of about 8.6 mm by 7.7 mm, while the electric contacts and the supporting glass pieces take most space of this bench-top device. The inset of Fig. 4 is a cross-section photograph of a passive check valve with 15 μm thick membrane.

TEST RESULTS

The assembled device was tested with an experimental setup that includes fluid reservoirs, valves, pressure sensors, and flow meters, as shown in Figure 5. The inset in Fig. 5 is an picture of the test jig with a micropump sealed in it by rubber O-rings. The system was first evacuated and then filled with 1 centistoke silicone oil. The piezoelectric cylinders were driven by AC signals with peak to peak voltages in the range of 0-1600 V, bias voltages in the range of 300-600 V, and frequencies in the range of 1-12.5 kHz. The inlet pressure was maintained at 850 kPa in all the tests. The highest flow rate recorded is 3000 μl/min at 1200 V peak to peak voltage, 600 V bias voltage, and 4.5 kHz, and zero differential pressure.
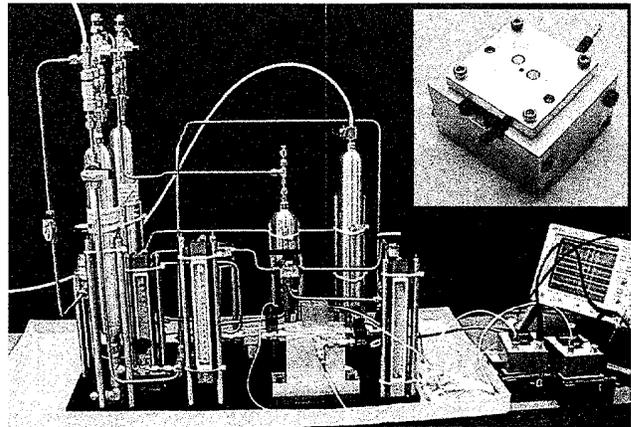
Figure 5. The experimental setup for micropump test. The inset is the package jig holding the micropump.

Figure 6 shows the experimental and quasi-static simulation results of the dependence of flow rate on peak to peak voltage. This test was done at a drive frequency of 3.5 kHz, a bias voltage of 600 V, and zero differential pressure. The flow rate increases with the voltage monotonically but not linearly. In Simulation 1 a rigid chamber structure was assumed. It is believed that the flow rate of Simulation 1 is higher than the experimental results because the micropump has, as part of the fluid chamber, flexible SOI

Figure 4. An assembled micropump with electric leads. Small graduations are in mm. The inset is a cross-section picture of a passive valve with 15 μm thick membrane.
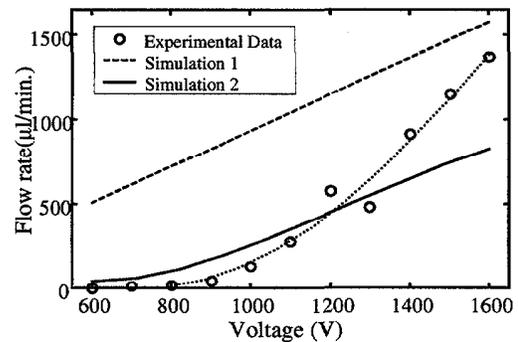
Figure 6. Flow rate vs. piezo peak to peak voltage at 3.5 kHz drive frequency. The results of two quasi-static simulations with different chamber compliance are also shown.

membranes attached to the drive element piston and the two passive check valves (see Fig. 1). An increased fluid chamber compliance was included in Simulation 2 and its results are closer to the experimental data as can be seen from Fig. 6. However, at high drive voltage the experimental flow rate increased with the drive voltage much faster than the results of Simulation 2. This is probably because the membranes deflect more and become stiffer as the drive voltage increases, resulting in a faster increase in flow rate, while in Simulation 2 this increase of membrane stiffness was not accounted for.

Figure 7 shows the relation between the inlet-outlet differential pressure and the flow rate. In this experiment the piezoelectric cylinder was driven at 1200 V peak to peak voltage, 600 V bias voltage, and 3.5 kHz.
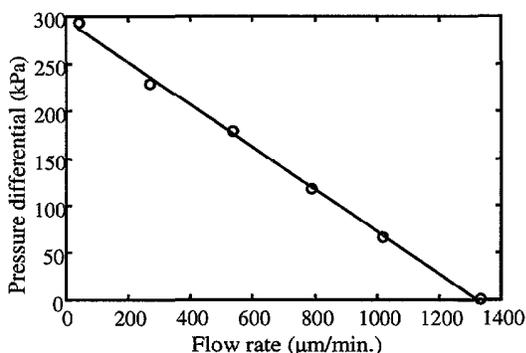
Figure 7. Inlet to outlet differential pressure vs. flow rate.

The flow rate vs. piezoelectric cylinder drive frequency curve of the same device and a quasi-static simulation result are shown in Figure 8. Other experimental conditions of this test were: peak to peak voltage at 1200 V, bias voltage at 300 V, and zero differential pressure. Figure 8 clearly shows some form of resonance effect that is not present in the simulation. We found that changing the bias voltage to the piezoelectric material and changing the pressure difference between the inlet and outlet have an influence on the value of the flow rate but not on the resonance peak positions. The flow resonance could be due to the pressure wave effects at abrupt corners of the package of the micopump or due to the internal structure of the device. A systematic study of the micropumps is under way to better understand their behavior and to improve their performance.
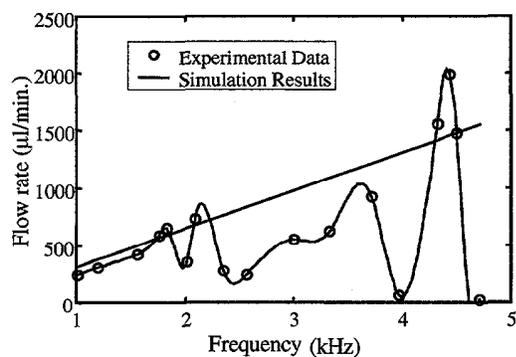
Figure 8. Flow rate vs. drive frequency of the micropump.

CONCLUSIONS

The work in this paper is part of a larger effort to develop high power density transducer devices [5] for micro-robotic actuation and human heel-strike power generation applications. The successful fabrication of the micropump has validated the MEMS process design and bonding techniques to be used for these devices. Since device power is generally proportional to flow rate, higher flow rate micropumps with larger drive elements and multiple piezoelectric cylinders are currently being developed. A further increase in flow rate can be realized by using newly developed single crystal piezoelectric materials [8]. A performance improvement of 10 times is expected in micropumps that combine larger drive elements and single crystal piezoelectric materials. In order to examine the effect of resonance on the performance of the micropump, the internal device structure and external test set-up are under investigation. In addition, it may be possible to use this resonance behavior to maximize the flow rate by tuning the device to a desired resonance frequency.

REFERENCES

1. S. Shoji and M. Esashi, "Microflow Devices and Systems", Journal of Micromechanics and Microengineering, 4, 157, (1994).
2. M. Koch, A.G.R. Evans, and A. Brunnschweiler, " The Dynamic Micropump Driven with a Screen Printed PZT Actuator", Journal of Micromechanics and Microengineering, 8, 119, (1998).
3. M. Esashi, S. Shoji, and A Nakano, "Normally Closed Microvalve and Micropmp Fabricated on a Silicon Wafer", Sensors and Actuators, vol. 20, pp 163-169, 1989.
4. D. Maillefer, H. Van Lintel, G. Rey-Mermet, R. Hirschi, "A High Performance Si Micropump for an Implantable Drug Delivery System", Techinical Digest, IEEE International MEMS'99 Conference, pp. 541-6, (1999).
5. Nesbitt W. Hagood IV, et al. "Micro-hydraulic Transducer Technology for Actuation and Power Generation", submitted to SPIE 7th International Symposium on Smart Structures and Materials, Newport Beach, CA, March 5-9, 2000.
6. D.C. Roberts et al. "Design of Piezoelectrically-driven Hydraulic Amplication for High Pressure, High Frequency Applications", submitted to SPIE 7th International Symposium on Smart structures and Materials, Newport Beach, CA, March 5-9, 2000.
7. H.T.V. Van Lintel, F.C.M. Van De Poll, and S. Bouwstra, "A Piezoelectric Micropump Based on Micromachining of Silicon", Sensors and Actuators 20, 163, (1989).
8. Seung-Erg Park and Thomas R. Shrout, "Characteristics of Relaxor-Based Piezoelectric Single Crystals for Ultrasonic Transducers", IEEE Trans. Ultrasonics, Ferroelectrics, and Frequency Control, 44, 1140 (1997).

CONTINUOUS MICROMIXER WITH PULSATILE MICROPUMPS

Ajay A. Deshmukh, Dorian Liepmann and Albert P. Pisano
Berkeley Sensor and Actuator Center
University of California, Berkeley
497 Cory Hall, Berkeley, CA 94720

ABSTRACT

A continuous microfluidic mixer utilizing pulsatile flow micropumps has been created using an SOI process. While individual MEMS mixers, pumps, and valves have been developed, this work presents an integrated, planar system utilizing novel mixing, valving, and bonding processes in order to deliver well mixed fluids at controllable flow rates.

INTRODUCTION

At small scales, using conventional mixing methods (such as turbulence) to mix fluids is almost impossible since even diffusion can be too slow. However effective mixing is still required for many microfluidic devices, with applications ranging from drug delivery to chemical analysis. Towards this need a new process for delivering well-mixed fluids from a MEMS device has been developed.

THEORY

Pulsatile flow can create a greatly lengthened fluid interface and much faster mixing. This process has been modeled numerically with water (white) and water with glucose (black) flowing through a 5 mm long, 400 micron wide channel at an average speed of 6 mm per second (Re = 2.4). Figure 1a shows the result of steady flow with diffusion; very little mixing occurs by the end. (Modeling done using the CFDRC's code, CFD-ACE; the pictures have been stretched in the vertical direction.)
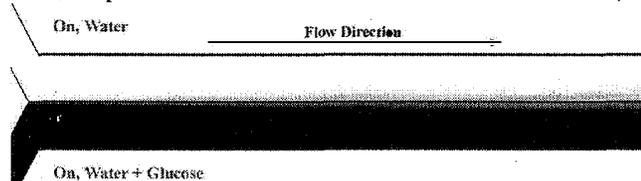
Figure 1a. *Steady flow*

At low Reynolds numbers, if one stream is stopped, the other stream will intrude into the other half of the main channel. In Fig. 1b, the top stream is turned off for 1/6 sec while the dark fluid from the bottom continues flowing.
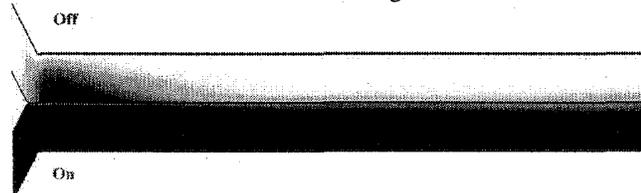
Figure 1b. *The top stream is stopped for 1/6 sec.*

As the fluid moves downstream, the velocity gradient in the channel will distort the intrusion and lengthen the interface even further. In Fig. 1c, the top stream is turned on and pushes the first intrusion downstream. Since the fluid travels faster at the centerline than at the edge, the dark intrusion is further distorted by the velocity profile. This distortion appears as a crescent shaped interface. The velocity profile is similar to, although not exactly like, the parabolic profile obtained from Poiseuille flow.
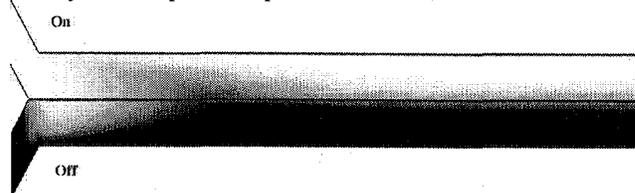
Figure 1c. *The bottom stream is stopped*

Fig. 1d is a picture of the flow when each stream has been turned on and off, out of phase with the other, at 3 Hz. The fluid mixes within a few channel diameters from the fluid junction. While not a steady state solution, this scenario will reoccur every 1/3 sec. The average speed of the exiting fluid is 3.5 mm/sec.
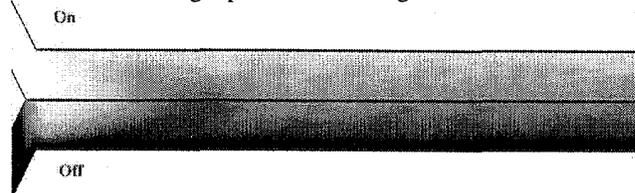
Figure 1d. *Well-mixed fluid after many pulses*

This novel approach to mixing allows for precise outlet composition via control of inlet stream duty-cycles and frequencies as well as continuous fluid delivery; these advantages are not possible using batch processes (e.g. Evans[1]) or passive mixers (e.g. Jaeggi[2]).

Pulsatile flow control is absolutely required for this mixing process. For the current system, positive displacement pumps have been developed. A schematic of the device is shown in Fig. 2a. Fluid 1 (white) is pumped into the mixing channel while Fluid 2 (black) is drawn in from the inlet.
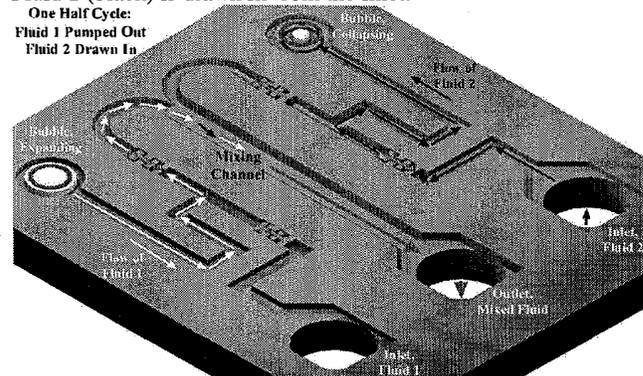
Figure 2a. *Schematic of a Micromixer*

The pumps consist of a bubble chamber and two check valves. When a bubble is created, it acts as a piston and drives fluid out; the check valves allow fluid to exit only in the output direction. When the bubble collapses, fluid is drawn in, but can

only enter from the inlet side. By cycling bubble generation and collapse, a net pumping action occurs. A close-up of this design is shown in Fig. 2b. (A very different piston and check valve pump is also shown in Bardell[3]).
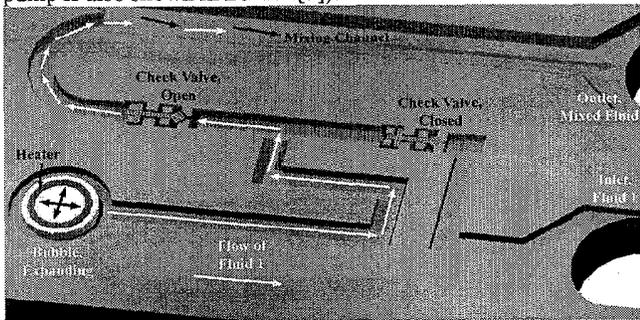
Figure 2b. Close-up of a pump

Steam bubbles can be generated using heat. When the heat is removed, the steam will re-condense and the bubble will collapse. In many cases, boiling the delivered fluid is undesirable. While boiling some of the fluid is necessary, a long channel connecting the bubble chamber to the main flow channel will isolate the working fluid from the delivered fluid.

DESIGN

The micromixer has been fabricated using SOI and quartz wafers bonded together. Bubbles are created by polysilicon resistors on quartz which act as heaters. In Fig. 3a is shown an early version of a heater and bubble chamber: the diameter of the bubble chamber is 800 microns. In Fig. 3b is shown a close-up of a new heater design that is more robust.
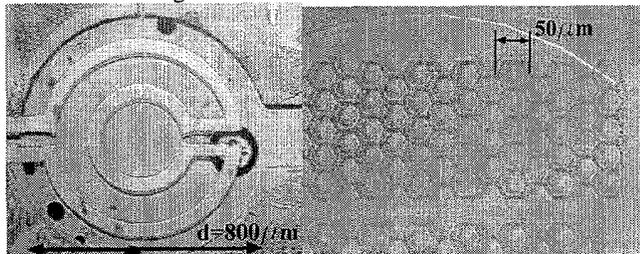
Figure 3a. Old heater   *Figure 3b. New heater*
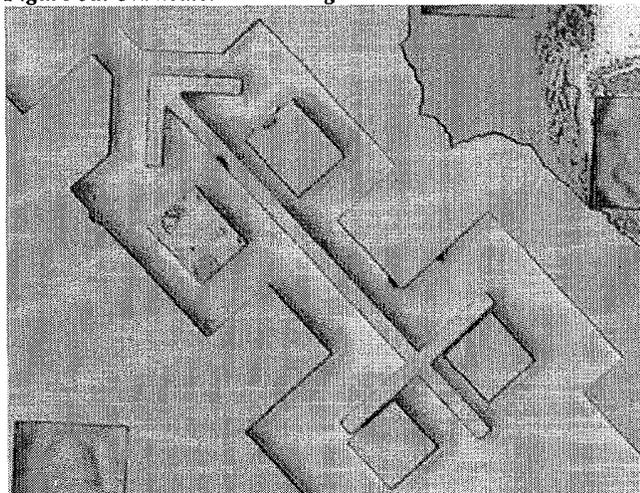
Figure 4. A check valve

The check valves are a new design consisting of free-floating silicon pieces that are actuated by viscous drag. The check valve

in Fig. 4 is free to move back and forth; flow to the lower right will continue unimpeded, whereas flow to the top left will close the valve. Note that these valves operate in the plane of the wafer; most previously disclosed MEMS valves (e.g. Wroblewski[4]) control flow that is out of the plane of the substrate. The channel width is 200 microns; the valve body length is 450 microns.

FABRICATION

The mixer is made of Silicon on Insulator (SOI) and quartz dice. The SOI wafer has a 2 micron thick buried oxide layer and a 75 micron device layer. Two sequential Deep Reactive Ion Etches (DRIE) are performed; the first going all the way through the wafer to form throughholes, while the second going only through the device layer to create the channels. A 1.3 micron thick wet oxide layer provides the mask for the second etch while a layer of 9.5 micron thick photoresist over the oxide serves as the mask for the first etch. After the DRIE processes, the valves are almost completely freed from the substrate in HF. They are then placed in H_2O_2 to generate a thin oxide layer on the bottom of the valve. Afterwards, a probe tip is used to gently free the valves.

The quartz wafer has doped polysilicon and sputtered aluminum applied, both of around 0.3 micron thickness. The layers, after being patterned into heaters, are passivated with silicon oxide and/or silicon nitride, leaving openings for the electrical connections. A novel process for bonding is used to attach the quartz cover plate. A low viscosity epoxy (Epotek 301) is spun on at 10,000 rpm, with tape applied to the electrical contacts to prevent epoxy from covering them. The two dies are then 'flip-chipped', with pressures of between 100 – 300 kPa per die. The epoxy bonds and seals the dies but also traps the valve bodies. In Fig. 5a, the covering die has been bonded on top and an epoxy layer fills the gap between the two surfaces.
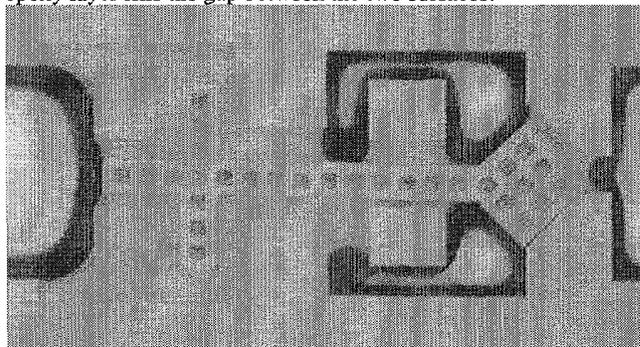
Figure 5a. Check valve after epoxy bond and before ashing
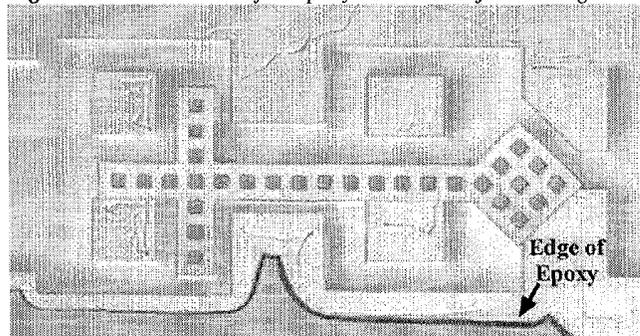
Figure 5b. Valve from Fig. 9a after ashing (note that it has moved)

An oxygen plasma (200 W in 600 mtorr of O_2) is used to remove the epoxy in the fluid channels. Throughholes allow the plasma to access the fluid channels. The oxygen first removes the epoxy on the top of the channels. The plasma then ashes the epoxy in between the silicon and the quartz. It then moves slowly outward from the channel. The plasma etching is timed to free the valves without penetrating far from the channel edge. The valves are freed, but sufficient epoxy remains outside the channels to bond and seal the device.

In Fig. 5b the edge of the epoxy layer can be seen. The epoxy has been etched farther above the channel than below it, possibly due to non-uniformity in the epoxy layer. Once the valves are free, the devices are placed in an H_2O_2 and surfactant ultrasound bath in order to create an oxide layer surrounding the valve bodies. Polyamide tubing is then epoxied into the throughholes.

A descriptive cross section is shown in Fig. 6; it does not correspond to any actual cross section in the device but is a combination of the heater and valve structures.
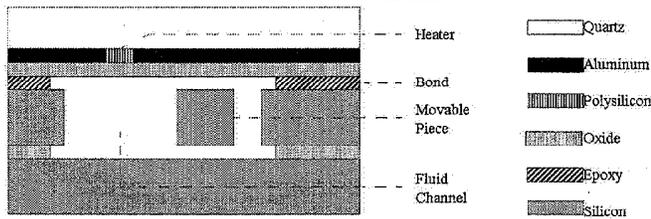
Figure 6. Fabrication cross section

EXPERIMENTAL DETAILS

In the following sequence of photographs, a bubble in the fluid can be used to visualize fluid flow.
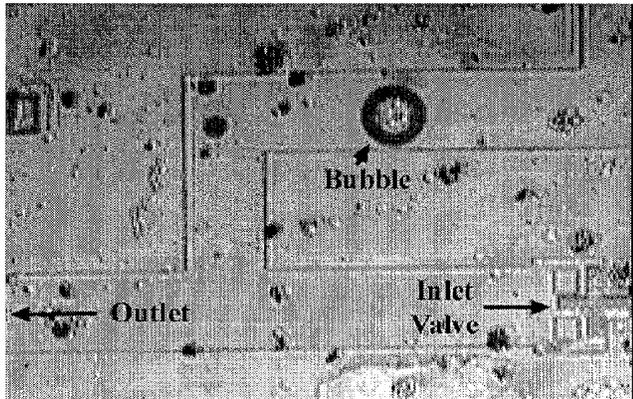
Figure 7a. Flow channel with a bubble inside
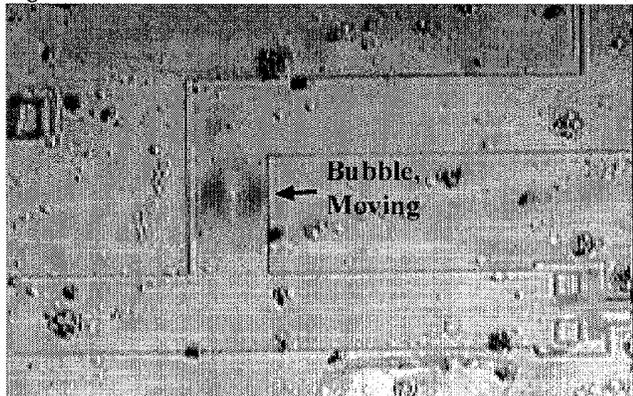
Figure 7b. Motion of the bubble after power has been supplied
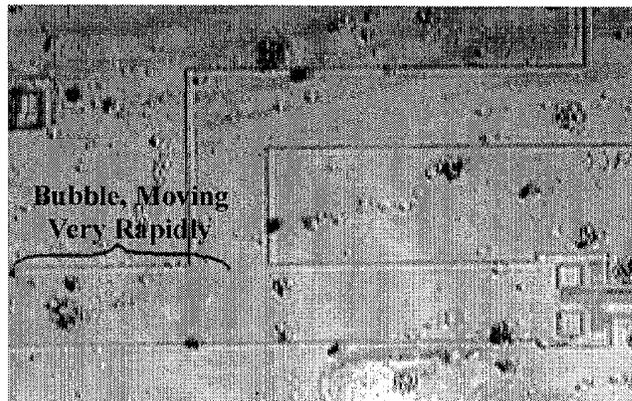
Figure 7c. Acceleration of the bubble as power is maintained

The bubble is originally stationary, but when the heater is activated, the fluid starts moving and travels through the outlet. The velocity of the bubble in Fig. 7c is approximately 15 mm/sec. This represents a peak and not an average velocity for the pump. Figures 8a and b isolate the outlet check valve.
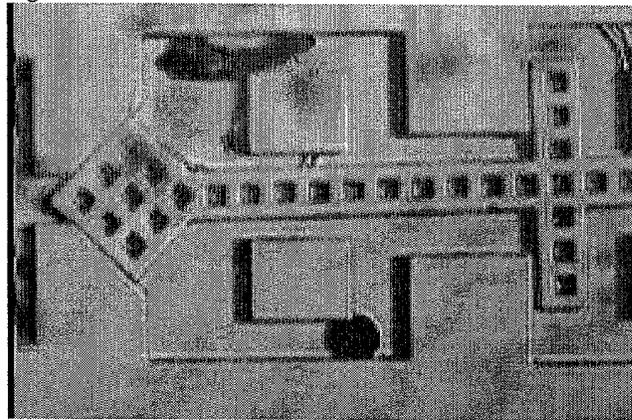
Figure 8a. Outlet valve with power off
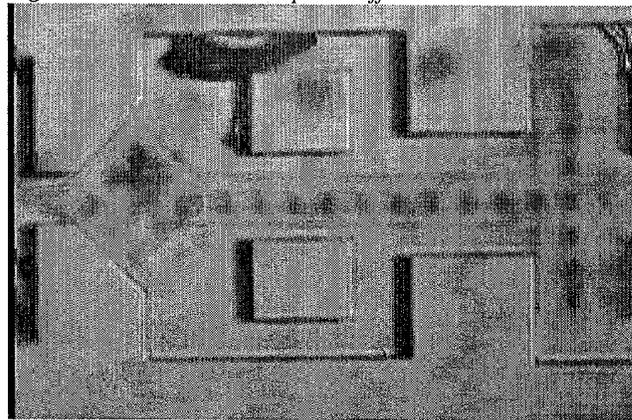
Figure 8b. Valve with power on

In Fig. 8b, the valve vibrates instead of simply staying open. This vibration is caused by the existence of a secondary oscillation in the flow. This oscillation is due to an instability that causes the bubble to fluctuate in size when the power is on. This secondary oscillation not only causes the valve body to move as seen but also actually provides additional pumping force.

Flow rates of around 5 microliters/min have been attained using pumping frequencies of approximately 0.5 Hz. The pump heaters require 1 to 4 Watts to thermally generate a bubble, but the average power is lower. While this is higher than desired, several optimisation strategies to lower power use are possible.

For example, if the desired flow rate is low, the pumps can be operated more slowly, with the heaters off longer than on. Better thermal isolation of the heater and bubble chamber will reduce heat loss to the substrate. Electrolysis bubbles might be usable as well; see Papavasiliou[5].

Mixing has been demonstrated with water containing either fluorescein (white) and food coloring (black). In Fig. 9 is shown four frames from a video recording taken 0.1 sec apart. The fluid is being pumped. The light colored fluid enters the mixing channel and curls into the lower half. The large dark bulge is consequently forced downstream. The interface between the two fluids stretches and distorts as the flow develops further. This distortion is highlighted in Fig. 10. The boundary between the two fluids becomes longer, more crescent shaped, and less distinct towards the right (outlet) end of the channel. Mixing takes place rapidly because of the increase in the interfacial area across which diffusion occurs.
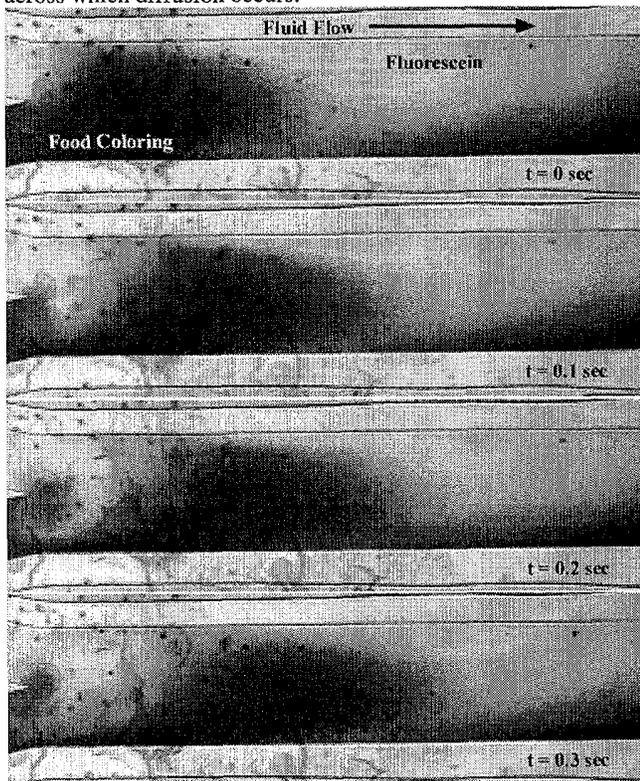
Figure 9. Mixing channel (4 frames)

The channel is 400 microns wide and a little more than 2 mm of length is visible. The flow rate here is estimated at about 1.5 microliters/min; the Reynolds number is approximately 0.06.
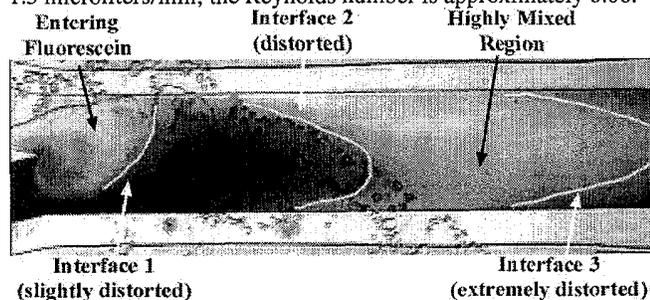
Figure 10. Mixing channel, annotated

CONCLUSIONS

An integrated microfluidic system for mixing has been developed. This system includes several novel fluid control components: free floating micro check valves, crack resistant thermal actuators and planar positive displacement micropumps. In addition, these components have been integrated into a single MEMS system capable of mixing and delivering any two fluids with arbitrary composition.

This system has demonstrated that unsteady forcing of a two-dimensional flow can efficiently mix two fluids in a continuous process (c.f. Evans [1]). While much work remains to be done to improve reliability and efficiency, a practical MEMS microfluidic mixer has been created. Moreover, the mixer is planar, which allows for easy integration with other microfluidic components.

REFERENCES

1. Evans, J.D., Liepmann, D., Pisano, A. P., "Planar Laminar Mixer", MEMS97, Nagoya, Japan, (1997), pp 96-101.
2. Jaeggi, D., Gray, B. L., Mourlas, N. J., van Drieënhuizen, B. P., Williams, K. R., Maluf, N. I., and Kovacs, G. T. A., "Novel Interconnection Technologies for Integrated Microfluidic Systems," Proceedings of the Solid-State Sensor and Actuator Workshop, Hilton Head, SC, (1998), pp. 112-115.
3. Bardell, R., Sharma, N.R., Forster, F.K., Afromowitz M.A., and Penny, R. "Designing high-performance micro-pumps based on no-moving-parts valves," Micro-Electrical-Mechanical Systems (MEMS), 1997 ASME International Mechanical Engineering Congress and Exposition, Dallas, TX, DSC-Vol. 234, (1997), pp. 47-53.
4. Wroblewski, D. E., Horenstein, M., Vandelli, N., Velonis, M., Bifano, T. "MEMS Micro-Valve Arrays for Fluidic Control" Micro-Electrical-Mechanical Systems (MEMS), C.J. Kim, et al., eds., 1998 ASME International Mechanical Engineering Congress and Exposition, Anaheim, CA, DSC-Vol. 66, (1998), pp. 145-151.
5. Papavasiliou, A., Liepmann, D., Pisano, A. P., "Fabrication of Free Floating Silicon Gate Valve," Proceedings of the ASME MEMS Division, 1999 IMECE, vol. 1 pp. 435-440.

POST-CMOS PROCESSING FOR HIGH-ASPECT-RATIO INTEGRATED SILICON MICROSTRUCTURES

Huikai Xie*, Lars Erdmann*, Xu Zhu*, Kaigham J. Gabriel*† and Gary K. Fedder*†

*Department of Electrical and Computer Engineering and †The Robotics Institute
Carnegie Mellon University, Pittsburgh, PA 15213, USA

ABSTRACT

We present a new fabrication sequence for integrated-silicon microstructures designed and manufactured in a conventional CMOS process. The sequence employs a post-CMOS deep silicon backside etch, which allows fabrication of high aspect ratio (25:1) and flat MEMS devices with integrated circuitry. A comb-drive actuator and a beam resonator were fabricated using this process sequence. Electrical isolation of single-crystal silicon was realized by using the undercut of the reactive ion etch (RIE) process. The fabricated devices were actuated electrostatically and flatness characterized using interferometric measurements.

INTRODUCTION

Micromachined devices fabricated by CMOS-compatible fabrication processes are attractive because of the possibility to integrate high-performance on-chip signal conditioning circuits and digital readouts, and expected multi-vendor accessibility and short design cycle times. Currently, most CMOS-compatible micromachining processes are polysilicon or poly-Ge surface micromachining processes that use silicon oxide as the sacrificial material and typically involve a wet etch for releasing micromechanical structures [1,2]. Even though vapor HF is often used, protection of integrated circuits and sticking problems during release are still major concerns [3]. Moreover, the relatively large parasitic capacitance in many integrated polysilicon processes degrades performance of capacitive sensor designs.

In prior work, post-CMOS surface micromachining processes use composite microstructures made from combinations of aluminum, silicon oxide, and silicon nitride thin films [4]. For these processes, the silicon substrate acts as the sacrificial material and is undercut for release (Fig. 1). The resultant multi-conductor microstructures provide wiring flexibility and low parasitic capacitance to the substrate. Both lateral and vertical CMOS accelerometers [5,6] with fully differential full bridge capacitive interface circuits have been fabricated by using this post-CMOS micromachining process. A tri-axial microstage has also been realized [7].

However, all of the microstructures discussed above are made of thin films, either homogeneous or multi-layer. Thin-film deposition processes generate residual stress and stress gradients which cause curling. This limits the maximum thickness and size of microstructures, which are critical dimensions for inertial sensors. Moreover, release holes and unwanted curvature of microstructures degrade their applications in optics.

Deep RIE (DRIE) technologies have advanced significantly in recent years. By alternating passivation and etching cycles, the Bosch advanced silicon etch (ASE) process [8] can typically achieve high aspect ratios of 20~30:1. For example, a bulk silicon comb-drive actuator with 100 μm deep comb fingers and 15 μm gap spacing has been fabricated by using the ASE process [9]. Because silicon etch rate of the ASE process is relatively high (~3μm/min), it is feasible to release microstructures by through-wafer etching, instead of by undercutting the silicon substrate [10]. However, given a typical wafer thickness (~500 μm), a minimum gap spacing of 20 μm would be needed to retain anisotropy. Such a spacing is too large for either effective sensing or actuation.

Our solution is to combine the maskless post-CMOS micromachining process [4], ASE and backside etch. The backside etch is used to control the thickness of the final, released, microstructures, which allows optimization of design parameters and simplifies the packaging process. The new process sequence (Fig. 2) provides high-aspect-ratio and flat microstructures. *It incorporates all the advantages of CMOS composite microstructures with the excellent mechanical properties of single crystal silicon.*

In this paper, a new CMOS DRIE micromachining process will be described and contrasted with the previous post-CMOS surface micromachining process. Characterization results of two example devices fabricated using the new process flow will be presented and some process issues will be discussed.

CMOS MICROMACHINING PROCESSES

CMOS surface micromachining process

The previous post-CMOS micromachining process [4] consists of an anisotropic dielectric etch for defining the microstructures and an isotropic silicon etch for release. Fig. 1 shows the cross-sectional view after the process is completed. The etching masks are interconnect metals from preceding CMOS processing, so no post-CMOS photolithography is required. An aspect ratio as high as 4:1 can be achieved. The microstructures consist of up to three metal layers that are electrically isolated by dielectrics, which
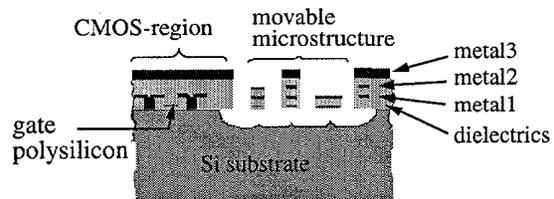
Figure 1. Cross-sectional view of the thin-film post-CMOS micromachining processing.
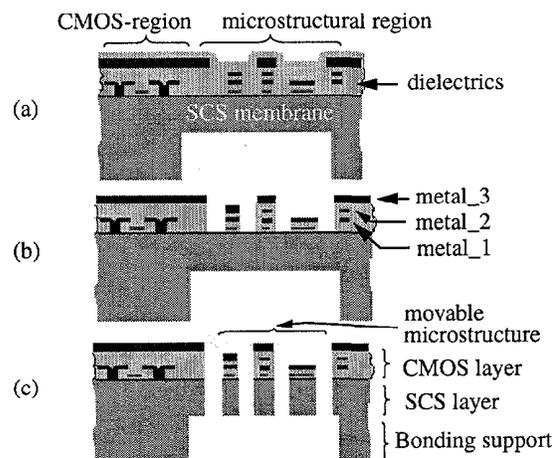
Figure 2. The process-flow for the modified CMOS micromachining. (a) CMOS-chip with backside etch. (b) Anisotropic dielectric etch. (c) Anisotropic silicon etch for release.

enables a variety of wiring configurations within the moving microstructures [5-7]. One issue with CMOS-MEMS is that the residual stress and thermal coefficient mismatch in the embedded layers cause curling or bending [11]. Vertical curling can be compensated to first order through a curl matching frame [5], but the maximum size which can be compensated is limited and design complexity is increased.

CMOS backside micromachining process

In order to overcome some of the drawbacks of thin-film microstructures without losing the advantages of the multi-conductor structures, we propose a new process sequence building on the previous post-CMOS micromachining process. The basic idea is to introduce a single-crystal silicon (SCS) layer underneath the CMOS multi-layer structures in such a way that the mechanical properties are dominated by the SCS layer and electrical connections are provided by the CMOS microstructure layer.

The process flow is given in Fig. 2. We start with a deep anisotropic backside etch leaving a 10 to 100 μm thick SCS membrane (Fig. 2 (a)). This backside etch step is used to control the thickness of microstructures as well as to form a cavity to simplify the packaging process. Next an anisotropic dielectric etch is performed from the frontside (Fig. 2 (b)). Then in contrast to the prior work on post-CMOS processing, we use an anisotropic instead of an isotropic silicon etch for release (Fig. 2 (c)). Therefore, a thick SCS layer remains underneath the CMOS layer, resulting in a totally flat released microstructure. Thus, no curling compensation is needed and a microstructure can be designed to an arbitrary shape. Furthermore, the proof mass and comb-finger capacitance for CMOS-MEMS based inertial sensors is significantly increased. Fig. 3 shows the SEMs of a comb-drive actuator fabricated by using the new process. Both sides are flat (no visible curling). In a later section, we will discuss the optical and electrical characterization of this device in detail.
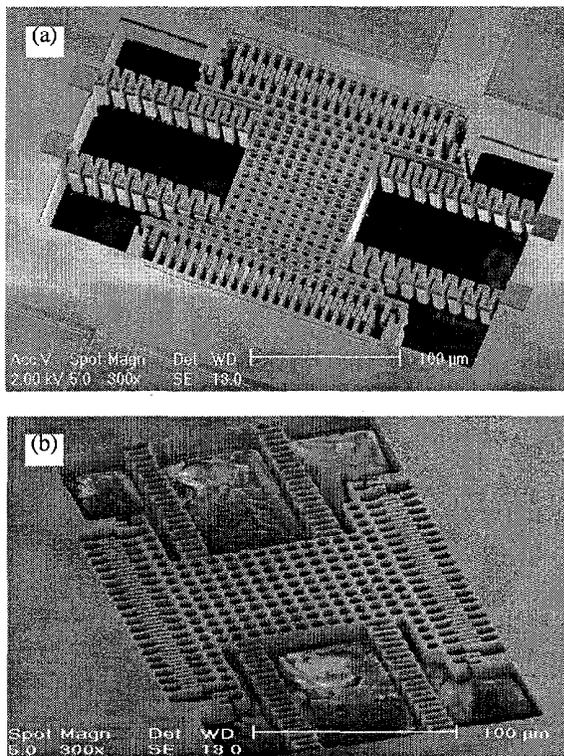
Figure 3. SEMs of the comb-drive actuator fabricated by using the new process. (a) Front-side view. (b) Backside view.

Design considerations

Since there is no fixed electrode underneath the microstructures, this technology is mainly suitable for sidewall capacitive sensing and actuation. Suppose there are two side-by-side beams, e.g., a pair of comb fingers, as shown in Fig. 4. The sidewall capacitance change versus displacement in the longitudinal, transverse and vertical directions have the following relationships,

Figure 4. Dimension definition of a sidewall capacitor

$$\frac{\partial C}{\partial x} \propto \frac{h}{g} \cdot \frac{L}{g+w}$$

$$\frac{\partial C}{\partial y} \propto \frac{h}{g} \cdot \frac{l}{g} \cdot \frac{L}{g+w}$$

$$\frac{\partial C}{\partial z} \propto \frac{l}{g} \cdot \frac{L}{g+w}$$

where l, g, h and w are the engaged length, gap, height and width of the comb fingers, respectively, and L is the total length of the comb-finger area. The total number of comb fingers is thus $L/(g+w)$. The gap aspect-ratio is h/g, which is fixed for the ASE process. If a fixed area is assumed, i.e., L is fixed, then the smaller the gap, the larger the capacitance change in all three directions. The height of the comb fingers may be set by the minimum possible gap. Thicker comb fingers would require a larger gap and result in a reduction in capacitive sensitivity. However, there is a minimum SCS layer thickness that just eliminates the stress-induced curling. Note that this minimum thickness value depends on the size of the designed microstructure as well as the thickness control accuracy of the deep RIE process. For inertial sensors, mass is a critical parameter that dictates resolution. For inertial applications, the SCS layer should be as thick as possible and there is a trade-off between the gap spacing and the structural thickness.

There is a small undercut in the deep RIE, which actually can be advantageous. If the width of a beam is smaller than two times the undercut, the SCS layer underneath this beam will be etched away. Through complete undercut of the silicon, released areas of bulk silicon can be electrically isolated. In addition, soft springs can be obtained by using the same principle. An electrically isolated SCS block is shown in Fig. 5, where the beam width is 0.8 μm and the CMOS bridge provides mechanical support as well as electrical wiring. To further guarantee the electrical isolation, the to-be-undercut silicon region is implanted as n-well (for p-silicon substrate) to form a p-n-p junction in case there is still a remaining thin layer of silicon.

Fabrication

We used single-chip processing to demonstrate the DRIE release sequence. The square CMOS chips (2 mm by 2 mm) are made in the Hewlett Packard 0.5 μm three-metal n-well process available through MOSIS.

The chips come with an unpolished backside. Prior to the deep etching process the backside was patterned in a photolithography step with a backside release mask using a 10 μm layer of a high viscosity photoresist (Shipley AZ 4620). The exposure was performed with a Karl-Suss MA 56 mask aligner. There is no need for a high backside-frontside alignment accuracy. The patterning has to guar-

CMOS bridge
(silicon underneath
is undercut)

CMOS layer

silicon island
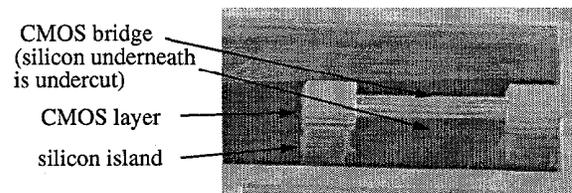
Figure 5. SEM of an electrically isolated SCS block

antee that a sufficient silicon frame for mechanical support of the silicon membrane remains after the deep etch. The chips were mounted on a 4" silicon wafer covered with photoresist. The backside ASE etch for defining the silicon membrane was performed in a Surface Technology Systems (STS) ICP etcher. The etching-process employed is a typical ASE process [8] with an etching rate of 2.9 μm / min for a small silicon load. The main plasma parameters for the etching part of the cycle are: 600 watt coil power; 12 watt platen power; 130 sccm SF_6 flow; and 23 mT chamber pressure. The passivation is performed at an identical plasma power and 12 mT chamber pressure with 85 sccm of C_4F_8 with no platen power. The duration of the etching cycle is 12 seconds and in the passivation cycle we deposit the fluorocarbon polymer for 8 seconds. The major demands for the etching process are good stability of the etching rate, low surface roughness and a sufficiently uniform structural thickness. A polishing of the chip backside prior to the backside deep etch improves the smoothness of the resulting silicon membrane surface but is not necessary. The backside etch step defines the thickness of the remaining silicon membrane and high accuracy measurements of the thickness of the chips and the etch depth are necessary. We employ an optical microscope to perform these measurements. By using a 50x microscope objective, accuracies of ±3 μm can be achieved. After removal of the chips from the carrier wafer and an oxygen plasma clean of the chip frontside, the oxide RIE etch of the CMOS-micromachining process (Fig. 2(b)) is performed in a Plasma-Therm 790 reactor.

For performing the anisotropic silicon release step (Fig. 2 (c)) in the STS ICP etcher, the chips were mounted on a 4" photoresist-covered silicon wafer. The top metal of the CMOS layers served as the mask. We observed no degradation of the top metal layer even for etching durations of several hours. We could not detect any performance decrease of the ICP etcher due to use of the Al mask. Only die-sized samples were processed, which limited the amount of aluminum exposed in the chamber.

The SEM in Fig. 6 shows the depth differences for three trench-widths due to well known micro-loading effects of the ASE process [12]. The trenches with a larger width are etched faster. Therefore the large open areas are etched through to the backside first before the narrow gaps. This results in polymer-deposition on the remaining silicon backside during the passivation steps of the ASE.

Fig. 7 (a) shows a close-up of the backside of a comb drive actuator after performing
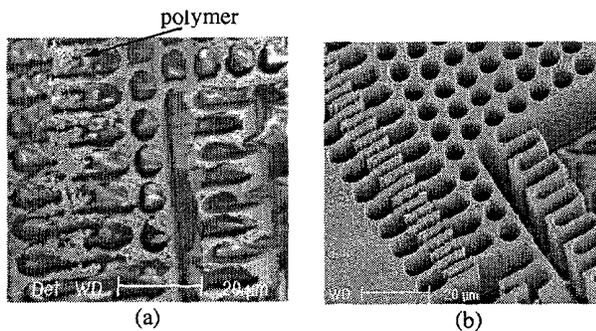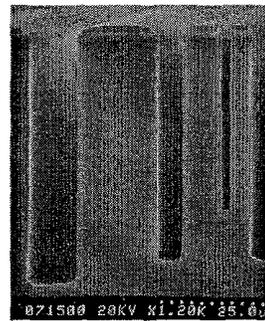
Figure 6. *Microloading effect of the ASE process.*

the anisotropic etch step. The polymer deposited on the backside of the SCS layer due to the relatively early etch-through of the wider gaps forms a residual structural layer that cannot be etched by continuing the ASE process. The polymer layer is removed by using a zero-bias oxygen plasma clean at the end of the process flow after the narrow gaps are etched through to the polymer film. Fig. 6 (b) shows the backside of a clean device after an oxygen plasma step is used at the end of the release procedure.

CHARACTERIZATION

Beam resonator

In order to investigate the mechanical properties of the composite microstructures with stacked CMOS and SCS layers, a 110 μm-long and 3.0 μm-wide cantilever beam was designed. Fig. 8 shows the released beams. The gap is 2.1 μm, which limits beam thickness to a maximum of 55 μm for an aspect-ratio of 25. The fabricated beam thickness is 45 μm. In the figure, the typical scallops of the ASE process and an undercut of about 0.4 μm were observed. The measured resonant frequency is 254 kHz, which is in good agreement with the finite element simulation of 249 kHz. The Young's modulus of the SCS layer is assumed to be 168 GPa [13] and the CMOS layer has an effective Young's modulus of 63 GPa [11].

Comb-drive actuator

Fig. 9 is a close-up of the comb-drive actuator shown in Fig. 3. The gap of the comb fingers is 1.2 μm and thus the underlying SCS layer has been thinned down to 25 μm thick for an aspect-ratio of ~25. The lateral displacement versus the applied DC voltage squared is plotted in Fig. 10. The plot shows that the stiffness of the serpentine springs is increased by a factor of 4 by the underlying thick SCS layer. Note that the underlying SCS beam width is ~0.8 μm narrower than the CMOS microstructure due to the 0.4 μm
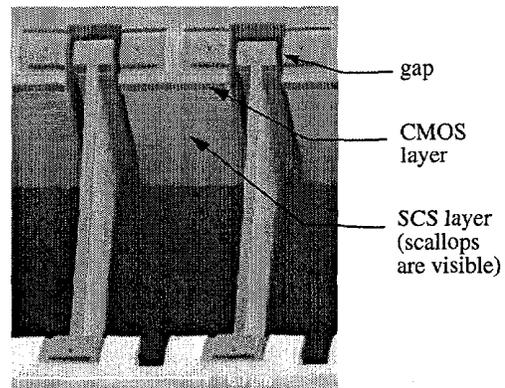
Figure 8. *SEM of the beam-resonators*

Figure 7. *SEMs of the backside of the comb-drive actuator.*
(a) No oxygen plasma clean step during the anisotropic release.
(b) Smooth and clean backside after oxygen plasma cleaning.
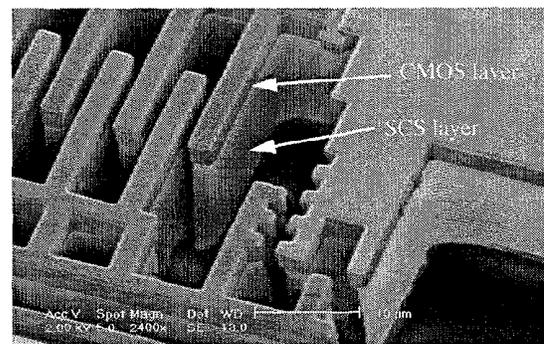
Figure 9. *Close-up of one corner of the comb-drive actuator (see Fig. 3).*
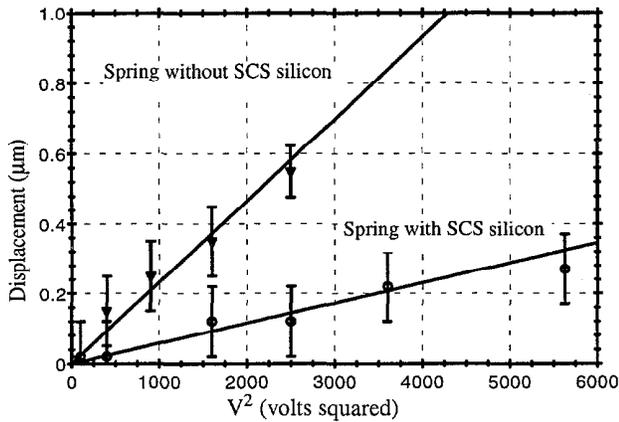
Figure 10. *Measured displacement of comb drive actuators versus squared voltage.*

undercut. Also note that the underlying SCS is not electrically isolated from the silicon substrate in this comb-drive actuator.

To evaluate the topography of the released structures, we employed a Linnik-type Michelson interferometer with LED illumination (λ=610 nm). The topography is calculated from a series of fringe patterns using a phase shifting technique. We achieved a measurement accuracy of better then 40 nm for the z-curling. Fig. 11 (a) shows the measured topography of a comb-drive actuator with no underlying SCS layer. The peak-to-valley curling measured on the whole device is 1.2 μm. For a device with underlying SCS, we found a curling of 0.15 μm -- nearly an order of magnitude decrease (Fig. 11 (b)).
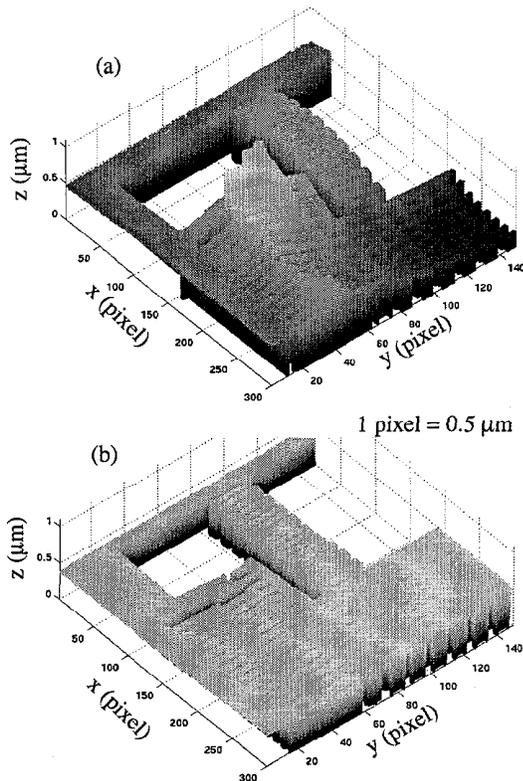
Figure 11. *The topography of the released comb-drive actuator (see Fig. 3) (only one quarter is shown) obtained by using phase shifting interferometry. (a) Conventional release. (b) Backside release.*

CONCLUSIONS

The new post-CMOS micromachining process is demonstrated by using a beam-resonator and a comb-drive actuator. This process is CMOS compatible and can generate flat and thick microstructures with optimized gap spacing. The unique electrical isolation of the silicon and multi-conductor features provide high expectations for capacitive sensing and actuation. The resultant large mass and increased comb-finger capacitance will give rise to high resolution and high sensitivity for inertial sensors as well as larger forces for microactuators.

The microloading effect is crucial in this process, which can be used to realize electrical isolation of silicon and soft springs. The silicon membrane thickness control is another important issue which is under further active investigation.

ACKNOWLEDGEMENT

REFERENCES

[1] R.T. Howe, B.E. Boser, A.P. Pisano, "Polysilicon integrated microsystems: technologies and application", *Sensors and Actuators A (Physical)*, vol.A56, no.1-2, p. 167-77.

[2] A.E. Franke, D. Bilic, D.T. Chang, P.T. Jones, T.-J. King, R.T. Howe, G.C. Johnson, "Post-CMOS integration of germanium microstructures", *MEMS'99*, Orlando, FL, USA; 17-21 Jan. 1999, p630-637.

[3] C. Hierold, "Intelligent CMOS sensors", *MEMS'2000*, Miyazaki, Japan, 25-28 Jan. 2000, p1-6.

[4] G.K. Fedder, S. Santhanam, M.L. Reed, S.C. Eagle, D.F. Guillou, M.S.-C. Lu, L.R. Carley, "Laminated high-aspect-ratio microstructures in a conventional CMOS process", *Sensors and Actuators A*, vol.A57, p.103-110.

[5] G. Zhang, H. Xie, L. deRosset, G.K. Fedder, "A lateral capacitive CMOS accelerometer with structural curl compensation", *MEMS'99*, Orlando, FL, USA; 17-21 Jan. 1999, p606-611.

[6] H. Xie, G.K. Fedder, "A CMOS z-axis accelerometer with capacitive comb-finger sensing", *MEMS'2000*, Miyazaki, Japan; 25-28 Jan. 2000, p496-501.

[7] H. Xie, L. Erdmann, Q. Jing, G.K. Fedder, "Simulation and characterization of a CMOS z-axis microactuator with electrostatic comb drives", *MSM'2000*, San Diego, CA; 27-29 Mar. 2000.

[8] F. Laermer and A. Schilp of Robert Bosch GmbH, "Method of Anisotropically Etching Silicon", US-Patent No. 5501893.

[9] Chris Seung-Bok Lee, Sejin Han, N.C. MacDonald, "Multiple depth, single crystal silicon microactuators for large displacement fabricated by deep reactive ion etching", *Technical Digest. Solid-State Sensor and Actuator Workshop*, Hilton Head Island, SC, USA; 8-11 June 1998, p.45-50.

[10] E.M. Chow, et al., "Fabrication of high-density cantilever arrays and through-wafer interconnects", *Technical Digest. Solid-State Sensor and Actuator Workshop*, Hilton Head Island, SC, USA; 8-11 June 1998, p.220-4

[11] M.S.-C. Lu, X. Zhu, G.K. Fedder, "Mechanical property measurement of 0.5-μm CMOS microstructures", *Symposium on Microelectromechanical Structures for Materials Research*, San Francisco, CA, USA; 15-16 April 1998, p.27-32.

[12] J. Jiao, M. Chabloz, T. Matsuura and K. Tsutsumi, "Mask layout related area effect on HARSE process in MEMS applications", *Transducers'99*, Sendai, Japan; 7-10 June. 1999, p546-549.

[13] J.C. Greenwood, "Silicon in Mechanical Sensors", *J. Phys. E, Sci. Instrum.*, 21,(1998), p.1114-1128.

MULTI-DEGREE OF FREEDOM DYNAMIC CHARACTERIZATION OF DEEP-ETCHED SILICON SUSPENSIONS

D.A. Horsley, P.G. Hartwell, R.G. Walmsley, J. Brandt, U. Yoon
Hewlett Packard Laboratories, 1501 Page Mill Road MS 2U-20, Palo Alto, CA 94304-1126
S. Hoen
Agilent Laboratories, 3500 Deer Creek Road MS 26-M2, Palo Alto, CA 94304-1392

ABSTRACT

Modern plasma etching techniques enable the fabrication of micromachined devices with extremely high height-to-width aspect ratio. This paper examines the mechanical performance of high aspect ratio silicon flexure suspensions. Flexures were etched with aspect ratios varying from 22:1 to 45:1 through the Bosch deep silicon etching process. Using non-contact measurement techniques and magnetic actuation, the ratio of the out-of-plane to in-plane suspension stiffness was measured as a function of in-plane displacement and flexure aspect ratio. This stiffness ratio was shown to be proportional to the square of the flexure aspect ratio, and diminished parabolically with in-plane displacement. A maximum out-of-plane to in-plane stiffness ratio of 700:1 was observed in a device with 45:1 flexures.

INTRODUCTION

The advent of semiconductor wafer bonding and deep reactive ion etching (DRIE) has enabled the fabrication of micromachined structures with extremely high height-to-width aspect-ratio. Such structures may be used to produce inertial sensors with large cross-axis mode separation or actuators with the ability to bear large out-of-plane loads[1],[2]. These properties come about because linear beam theory predicts that a rectangular cross-sectioned beam will have a ratio of out-of-plane to in-plane bending stiffness that is proportional to the aspect-ratio squared:

$$\frac{k_z}{k_x} = \frac{I_x}{I_z} = \left(\frac{h}{w}\right)^2 \qquad (1)$$

where w and h are the beam width and height, respectively. Real-world devices may not match the predictions of Eq. (1), however. In reality, plasma etched beam profiles are never exactly rectangular. Additionally, linear beam theory does not capture such effects as the dependence of the in-plane and out-of-plane stiffness on in-plane deflections. Finally, because typical micromachined suspensions consist of folded or hairpin flexures connected by short linkages, losses in the folds and linkages may severely reduce the out-of-plane suspension stiffness.

To take a closer look at the mechanical performance of high aspect-ratio suspensions, 100 μm thick suspended test structures were etched into silicon-on-insulator (SOI) substrates using the Bosch DRIE process. A SEM image of a typical device is illustrated in Fig. 1. Each test structure consisted of a 1.4 mm × 1.4 mm × 0.1 mm mobile stage, flexurally suspended on four sides to allow two degree-of-freedom motion in the x-y plane while limiting out-of-plane (z-axis) motion. To allow the effect of varying aspect ratio to be investigated, flexures were designed with widths of 3 μm, 4 μm, and 5 μm. Mask undercut from the DRIE process varied from 0.5 μm to 1.0 μm depending on the device location on the wafer, resulting in fabricated aspect-ratios (AR) from 22:1 to 45:1.
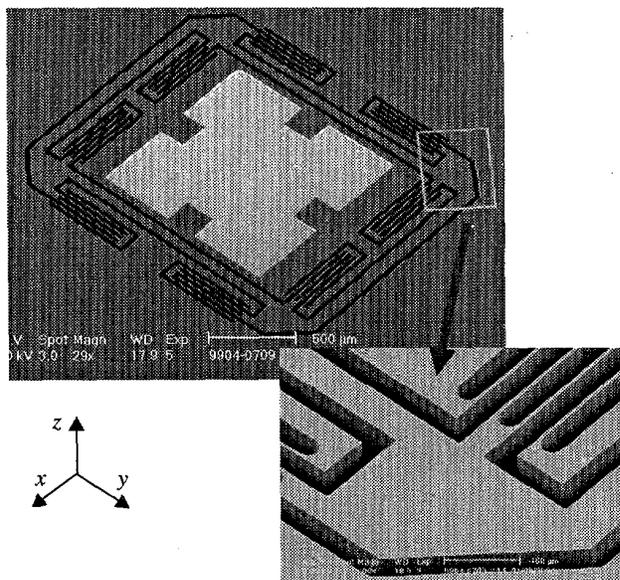
Figure 1. SEM images of completed test structure. Folded flexures provide x and y axis compliance, while limiting out-of-plane (z-axis) displacement. The center of the device is coated with NiFe to permit magnetic actuation.

THEORETICAL SUSPENSION PERFORMANCE

The x-y stages described in this paper utilize cascaded suspension elements to produce freedom of motion in two axes. The x and y axis stiffness of these suspensions were equal and will be denoted by k_x. The effective z-axis stiffness of the entire suspension, k_z, is the series combination of the z-axis stiffness of the x and y suspension elements. Thus, the maximum ratio of out-of-plane to in-plane stiffness is half that predicted by Eq. (1), or:

$$\frac{k_z}{k_x} = \frac{1}{2}\left(\frac{h}{w}\right)^2 \qquad (2)$$

The difference between the stiffness ratio, k_z/k_x, achieved by a particular device and the value predicted by Eq. (2) can be considered to be a figure of merit. Additional losses in out-of-plane stiffness are contributed by bending in the linkages connecting the individual flexures, and by torsional bending of the flexures themselves. Careful suspension design is required to minimize these effects.

Even the best suspension design is subject to limitations from the cross-sectional geometry of the fabricated flexures. Although the side-wall profile obtained by DRIE is very close to vertical, an angular deviation of 0.1° results in nearly 0.2 μm of undercut at a depth of 100 μm. Beam width variations have a more pronounced effect on in-plane stiffness than on out-of-plane stiffness. For small $\Delta w/w_0$, the change in the stiffness ratio is approximately:

$$\frac{k_z}{k_x} = \frac{k_{z0}}{k_{x0}}\left(1 - 2\frac{\Delta w}{w_0}\right) \tag{3}$$

where k_{x0} and k_{z0} are the in-plane and out-of-plane stiffesses of a rectangular beam of width w_0. Assuming that the beam profile is independent of width, the stiffness loss predicted by Eq. (3) will be more pronounced at narrower beam widths, since Δw will be a larger fraction of w_0. Although variations in beam height can also result in stiffness losses, the use of SOI substrates guarantees less than 0.5% variation in the flexure height from wafer to wafer.

In addition to the stiffness ratio at zero displacement, it is important to examine the effects of in-plane displacements on flexure stiffnesses. In general, both k_x and k_z can be effected by in-plane displacements. Increased torsional stress from in-plane displacement has the effect of reducing k_z. As a consequence, the stiffness ratio is in turn reduced. For a rectangular cross-sectioned beam, this stiffness loss can be calculated from the inverse of the stiffness ratio:

$$\frac{k_x}{k_z} = \left(\frac{w}{h}\right)^2 + \alpha\left(\frac{x}{L}\right)^2 \tag{4}$$

where x is the in-plane displacement, L is the beam length, and α is a constant that is approximately equal to 0.14 for aspect ratios greater than 5:1. At displacements greater than 10% of the flexure length, tensile stiffening effects begin to increase the in-plane stiffness. However, the displacements in this study were sufficiently small to prevent tensile stiffening effects from occuring.

FABRICATION

Deposit 2 μm NiFe on 100 nm Ta adhesion layer.
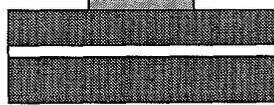
DRIE flexures, stop on buried oxide.
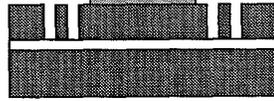
DRIE backside cavities, stop on buried oxide. HF etch to release suspended test structure.

Figure 2. *Simplified fabrication process flow.*

A simplified fabrication process flow is shown in Fig. 2. Processing began on SOI wafers with a 100 μm thick device layer bonded to a 450 μm thick handle wafer via a 1 μm thick buried oxide layer. A 2 μm NiFe layer was sputtered onto the front surface and patterned with a lift-off technique. Next, 100 μm tall flexures were deep etched into the device layer using the Bosch silicon etching process. The silicon etch rate was observed to vary by approximately 3% from the edge to the center of each wafer. This etch-rate non-uniformity necessitated a slight overetch to clear the silicon at the wafer center. A slower etching recipe was used for this overetch in order to minimize undercutting of the silicon layer at the Si/SiO$_2$ interface. Although this approach improved device yield, it did

result in significant variations in the across-wafer sidewall profile of finished beams. Following the flexure etch, 450 μm deep cavities were etched into the back-side of the wafer. The buried oxide was used as an etch stop layer for the back-side cavity etch, leaving the stage supported by a thin oxide membrane. The final processing step is to remove this oxide with hydrofluoric acid, releasing the finished stage from the wafer.

A cross-sectional view of the deep etched flexures before removal of the oxide membrane is illustrated in Fig. 3. The cross-section was taken near the end of the first, third and fifth trench, and some silicon residue is visible at the base of these flexures. Although this residue was typically observed only near the flexure ends, the base of each fabricated flexure was either slightly narrower or slightly wider than the as-drawn linewidth, depending on the location of the beam on the wafer.
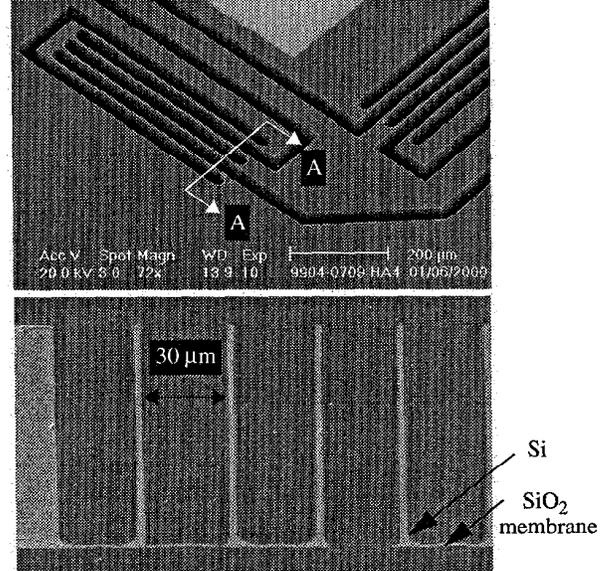
Figure 3. *Cross-sectional view of deep-etched flexures. This cross-section taken at A-A reveals some silicon residue at the base of the 1st, 3rd, and 5th trenches.*

TESTING APPARATUS

It is possible to directly measure flexure stiffness using load cells and displacement transducers[4]. However, this technique requires careful positioning of the applied load and suffers from low precision when applied to devices with a limited range of travel or very high compliance. For this reason, resonant frequency measurements were used to produce higher precision measurements of suspension stiffness. The relationship between the resonant frequencies and the stiffness ratio is simply:

$$\frac{k_z}{k_x} = \left(\frac{f_z}{f_x}\right)^2 \tag{5}$$

where f_x and f_z are the in-plane and out-of plane resonant frequencies, respectively. Natural frequency can be easily measured with an accuracy of better than 1%, resulting in a 2% accuracy in the calculated stiffness ratio. In addition, the dependence of the stiffness ratio on in-plane displacements can be studied by measuring the local natural frequencies at a fixed displacement. Calculation of the spring constants themselves, k_x and k_z, requires knowledge of the effective mass of the stage, but this value can be accurately esti-

mated based on the SOI wafer thickness and the stage geometry.

Each device was tested by exciting and measuring the in-plane and out-of-plane vibration modes at selected points over a ± 30 μm range of in-plane displacements. As illustrated in Fig. 4, out-of-plane vibration was induced using a vibration exciter (shaker), and external magnets were used for in-plane displacement. These external magnets consisted of a wire coil for ac stimulation and a permanent magnet assembly for dc inputs. The magnets generated a maximum force of approximately 800 μN. Static in-plane displacements were accomplished by moving the permanent magnet assembly on a micrometer driven stage. A frame-grabber and CCD camera were used to measure the resulting in-plane displacement of the stage with 0.25 μm accuracy. A microscope mounted laser-doppler vibrometer was used to measure the out-of-plane natural frequency, while the in-plane natural frequency was measured from the blur envelope of the video image[5].
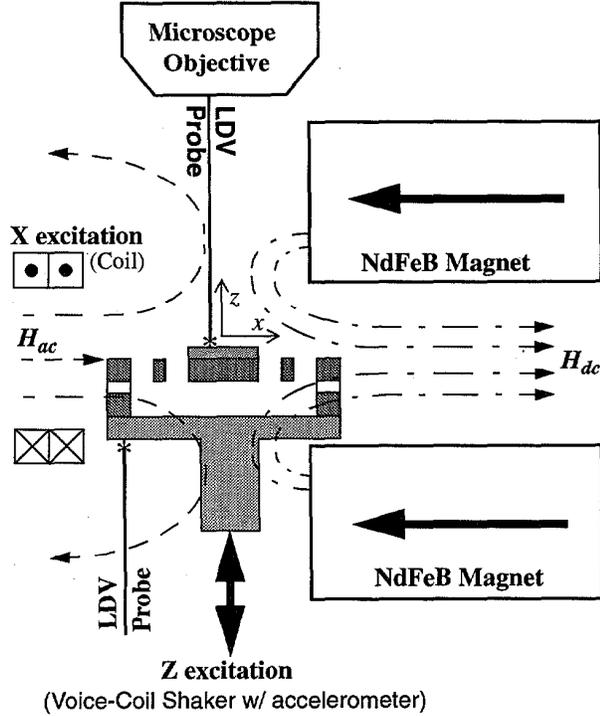
Figure 4. Excitation and measurement setup. The sample is mounted on a z-axis vibration exciter. NdFeB magnets may be moved relative to sample to vary the dc magnetic field, while ac fields are applied with a coil.

The measurement system allowed accurate measurement of both f_x and f_z at a given displacement. However, nonlinearities in the magnetic force were the dominant contributor to the measured variation in f_x as a function of displacement. Denoting the NiFe saturation magnetization as M_s, the magnetic field, $H(x)$, produces an in-plane force on the NiFe layer that is proportional to the magnetic field gradient:

$$F(x_0) = M_s S \left[H\left(x_0 - \frac{L}{2}\right) - H\left(x_0 + \frac{L}{2}\right) \right] \approx M_s SL \frac{dH}{dx}\bigg|_{x_0} \quad (6)$$

where x_0 is the distance from the edge of the permanent magnet to the center of the test sample, and L and S are the length and cross-sectional area of the NiFe film. However, a magnetic spring effect

is produced by the derivative of Eq. (6):

$$\frac{dF}{dx}(x_0) = M_s S \left[\frac{dH}{dx}\bigg|_{x_0 - \frac{L}{2}} - \frac{dH}{dx}\bigg|_{x_0 + \frac{L}{2}} \right] \approx M_s SL \frac{d^2 H}{dx^2}\bigg|_{x_0} \quad (7)$$

The permanent magnet assembly was designed to produce a large x-axis magnetic field gradient, maximizing the difference in the field at the two edges of the NiFe sheet. An unfortunate consequence of the design is that it also resulted in a large second derivative of the magnetic field. Although it was possible to remove the permanent magnets to measure f_x and f_z at zero displacement, introducing the magnets resulted in a measurable shift in f_x. Since in-plane displacements were induced by moving the magnet assembly, the magnetic shift in f_x was displacement dependent. The measured shift was roughly 5% over the full 30 μm range of travel, in agreement with the value predicted using an electromagnetic finite element analysis (FEA). In contrast, mechanical FEA predicted a frequency shift of less than 1%, suggesting that the magnetic effect dominated over any mechanical stiffening.

This measurement limitation was not observed in the z-axis. Because the NiFe is a thin sheet, the magnets did not produce appreciable out-of-plane force or spring stiffness. In addition, the out-of-plane mechanical suspension stiffness is more than two orders of magnitude greater than the in-plane stiffness, rendering the out-of-plane measurements relatively insensitive to magnetic field effects.

EXPERIMENTAL RESULTS

The measured stiffness ratio (k_z/k_x) of seven devices is displayed in Fig. 5 along with the ratio calculated by FEA. Test samples with both 3 μm and 4 μm as-drawn flexure widths were tested, with the 3 μm beams having a higher stiffness ratio than the 4 μm beams. However, varying beam profile caused the stiffness ratio to vary from device to device. Devices with 5 μm flexures were too stiff to be displaced by more than 10 μm and do not appear on the plot. The FEA model had the same geometry as the test samples and a 3 μm flexure width. As described in the preceding section, it was not possible to measure the variations in the in-plane stiffness versus displacement, so the nominal stiffness at zero displacement, k_{x0}, was used in preparing this plot.

In general, the out-of-plane stiffness diminished parabolically with in-plane displacement, as described by Eq. (4). The data in Fig. 5 correspond to the inverse of Eq. (4). When plotted in this form, the stiffness loss increases with aspect ratio, however, the parabolic loss term, α, is largely independent of aspect ratio. The constant is theoretically equal to 0.14 for a simple beam, but tested suspensions consisted of cascaded beams of two lengths, each of which was subjected to a different deflection to achieve a particular stage displacement. Correcting for these factors, the predicted parabolic loss term for these suspension designs was 0.09. The 3 μm flexures had an average measured parabolic loss constant of α = 0.09, while the 4 μm flexures had a slightly lower measured loss constant of α = 0.07.

While the devices with 3 μm flexures have a higher stiffness ratio than those with 4 μm flexures, there is a great deal of variation from device to device. This variation is due to the edge-to-center beam profile variations. A convenient way of comparing the vari-
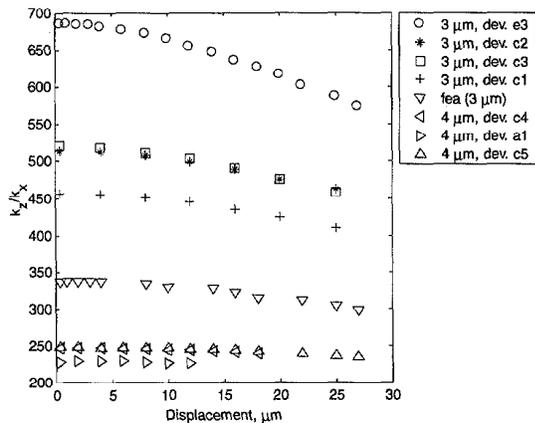
Figure 5. *Measured stiffness ratio versus in-plane displacement. The as-drawn beam width of each sample is indicated on the legend.*

ous devices is to plot the nominal stiffness ratio of each as a function of the flexure aspect-ratio. Direct measurement of the beam profile would require destructive cross-sectional analysis, so a simpler approach was adopted. Assuming a rectangular beam profile, the nominal beam width was estimated from the ratio of the measured natural frequency to that predicted by the FEA model:

$$w = w_{FEA}\left(\frac{f_x}{f_{x,FEA}}\right)^{2/3} \qquad (8)$$

Despite the fact that the fabricated beam profiles deviated slightly from rectangular, the estimates calculated from Eq. (8) proved reasonably accurate. Two cross-sections measured using a SEM are shown along with estimated widths in Fig. 6

The nominal stiffness ratio at zero displacement is plotted in Fig. 7 as a function of the aspect ratio computed from Eq. (8). The theoretical relationship between the stiffness ratio and aspect ratio from Eq. (2) is plotted as a solid line, while a linear fit to the data is plotted with a dashed line. The slope of this fit is 0.32, or 36% lower than the theoretical maximum stiffness ratio. The fact that the FEA data also fits close to this curve suggests that the suspension design is the main source of this loss, rather than any nonideality in the flexures themselves.
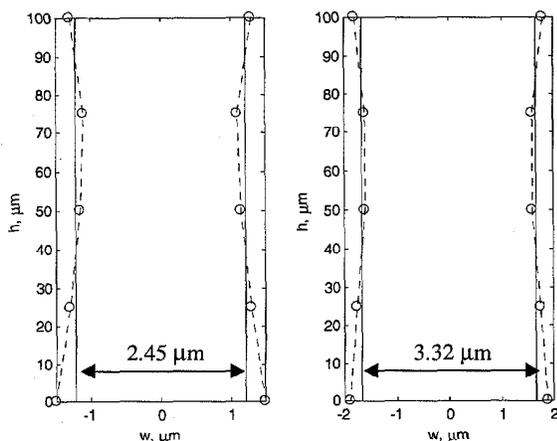
Figure 6. *Measured beam profiles and widths estimated by Eq. (8)*
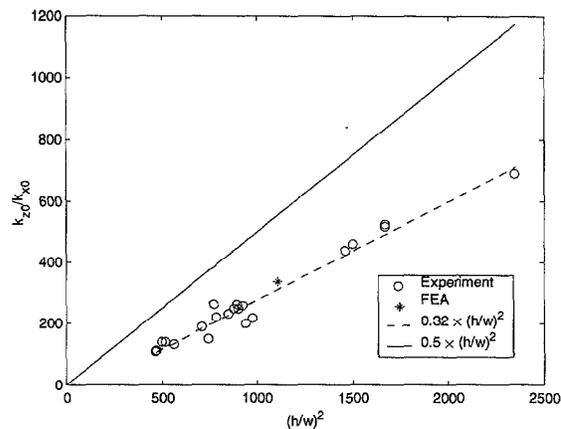
Figure 7. *Nominal stiffness ratio (k_{z0}/k_{x0}) vs. aspect ratio squared. A linear fit to the data is indicated with a dashed line, while the solid line indicates the theoretical limit from Eq. (2)*

CONCLUSIONS

This paper presented a study of the mechanical performance of deep etched silicon flexure suspensions. While the ratio of the in-plane to out-of-plane flexure stiffness was shown to be proportional to the square of the aspect ratio, this ratio was 36% lower than predicted for a single, rectangular cross-sectioned beam. The source of this additional stiffness loss is bending in the linkages connecting the individual flexure elements, as well as torsional bending introduced at these linkage points. Variations in k_x and k_z as a function of in-plane displacements were also measured. Variations in k_x were measured to be less than 5% at a 30 μm displacement, and the measured variation was magnetically, rather than mechanically, induced. Variations in k_z were measured with accuracy of less than 1%, and were shown to have a parabolic dependence on in-plane displacement.

REFERENCES

1. T.J. Brosnihan, J.M. Bustillo, A.P. Pisano, R.T. Howe, "Embedded Interconnect and Electrical Isolation for High-Aspect-Ratio, SOI Inertial Instruments", *Proc. 1997 Intl. Conf. on Solid State Sensors and Actuators*, Chicago, IL, June 1997, pp. 637-640.

2. D.A. Horsley, N. Wongkomet, R. Horowitz, A.P. Pisano, "Design and Feedback Control of Electrostatic Actuators for Magnetic Disk Drives", *Proc. 1998 Workshop on Solid-State Sensors and Actuators*, Hilton Head, SC, June 1998, pp.120-123,

3. R.I. Pratt, G.C. Johnson, R.T. Howe, J.C. Chang, "Micromechanical Structures for Thin Film Characterization", *Proc. 1991 Intl. Conf. on Solid State Sensors and Actuators*, San Francisco, CA, June 1991, pp. 205-208.

4. V. Temesvary, S. Wu, W.H. Hsieh,Y.-C. Tai, D.K. Miu, "Design, fabrication, and testing of silicon microgimbals for super-compact rigid disk drives", *Journal of Microelectromechanical Systems*, V. 4, No. 1, pp. 18-27, 1995.

5. K.L. Turner, P.G. Hartwell, and N.C. MacDonald, "Multi-dimensional MEMS Motion Characterization Using Laser Vibrometry", *Proc. 1999 Intl. Conf. on Solid State Sensors and Actuators*, Sendai, JP, June 1999, pp. 1144-1147.

A MICROMACHINED PENDULOUS OSCILLATING GYROSCOPIC ACCELEROMETER

Todd J. Kaiser

Milli Sensor Systems & Actuators, Inc.
West Newton, MA 02165

Mark G. Allen

School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0250

ABSTRACT

A silicon Pendulous Oscillating Gyroscopic Accelerometer (POGA) was fabricated using deep reactive ion etching (DRIE) and silicon wafer bonding technologies. The accelerometer is composed of three individual layers that are assembled into the final instrument. The top layer uses wafer bonding of an oxidized wafer to a handling wafer to create a silicon-on-oxide wafer pair, in which the oxide layer provides electrical isolation between the mechanical members and the handling layer. The middle layer is a two-gimbal torsionally-supported silicon structure, and is in turn supported by an underlying drive/sense layer. The instrument proved to have better than milli-g resolution and dynamic ranges in excess of 1-g (open loop) and approximately 12 milli-g (closed loop).

INTRODUCTION

The Pendulous Oscillating Gyroscopic Accelerometer (POGA) is the oscillatory analog of the Pendulous Integrating Gyro Accelerometer (PIGA), the most accurate strategic-grade accelerometer to date [1]. Instead of rotating members as in the PIGA, the members of the three orthogonal axis system in the POGA oscillate [2]. The interaction of the oscillations of the inner and outer members provide a DC torque to the middle member to rebalance a pendulous seismic mass. Because the members oscillate rather than rotate, significant design and manufacturing simplifications are possible. The oscillatory nature of the POGA makes it amenable to layered fabrication, which is achievable using micromachining technologies. The operation of the POGA is easiest understood if the operation of the PIGA is first reviewed.

The PIGA operation is based on gyroscopic theory of rotating bodies [3-5]. A spinning mass called the momentum wheel or rotor generates angular momentum about the spin axis. The rotor is supported by a housing called the torque-summing member (TSM) that allows rotation of the wheel about a second axis that is perpendicular to the spin axis. This combination of components is a single-degree-of-freedom gyroscope. A rotation of the gyro about a third axis that is perpendicular to both the TSM and rotor axes, generates a gyroscopic torque on the TSM which acts to rotate the TSM. The third axis is usually referred to as the input axis and the TSM axis is usually referred to as the output axis.

A PIGA is formed from the above gyroscope by making the TSM pendulous and by mounting the gyro onto a member that allows the rotation of the gyro about its input axis. This member is referred to as the Servo Driven Member (SDM). The TSM is made pendulous by adding a mass, m, to the TSM usually along the spin axis at some moment arm, l, from the output axis. The result is an arrangement that sums a pendulous and gyroscopic torque on the same member, the TSM. Since acceleration varies the pendulous torque, the gyroscopic torque can be servoed by rotating the SDM to balance the pendulum

and hence hold the TSM at null. At this null, the rotation of the SDM can be related to the input acceleration. The equilibrium torque equation is given by:

$$mla = I_1 \dot{\psi} \dot{\phi}, \qquad (1)$$

where I_1 is the moment of inertia of the rotor about the spin axis, $\dot{\psi}$ is the angular velocity of the rotor, and $\dot{\phi}$ is the angular velocity of the SDM.
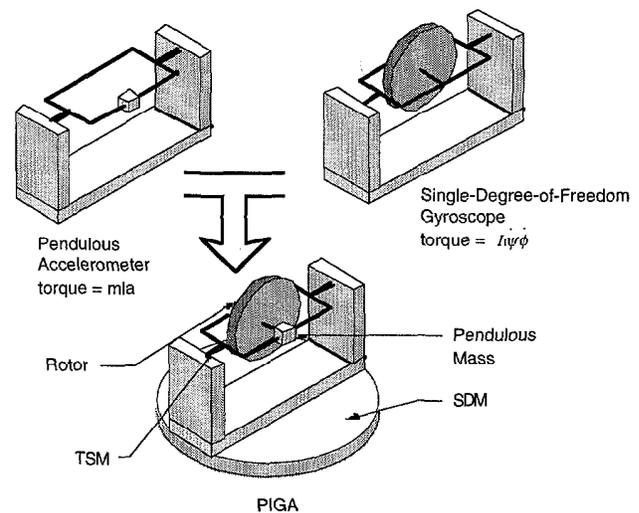
Figure 1. The PIGA is the combination of a pendulous accelerometer and a single-degree-of-freedom gyroscope.

This replacement of the rotary members with oscillatory members is the fundamental design change between the PIGA and the POGA. It makes the device suitable for microfabrication while maintaining the fundamental physics that has produced the highest performing accelerometers.

The POGA is also a three orthogonal axis system. Instead of rotating members as in the PIGA, the members are supported by flexures and designed to oscillate. The inner member designated the rotor-driven member (RDM) generates the angular momentum just as in the PIGA, but instead of a fixed angular momentum, the angular momentum oscillates. The oscillating angular momentum requires an oscillating SDM to generate the gyroscopic torque that nulls the TSM. If the RDM and SDM are forced to oscillate at the same frequency ω, then the equilibrium torque equation can be given by:

$$mla = I_1 \dot{\psi} \dot{\phi} = I_1 \Psi \Phi \omega^2 \sin(\omega t) \sin(\omega t + \beta), \qquad (2)$$

where Ψ and Φ are the amplitude of oscillation of the RDM and SDM, and β is phase difference between their oscillations.

Trigonometric and algebraic manipulations produce the resulting expression for the acceleration:

$$a = \frac{I_1 \Psi \Phi}{2ml} \omega^2 \cos \beta + f(2\omega). \qquad (3)$$

The acceleration is a function of the cosine of the phase difference between the two oscillations plus a second harmonic term that time averages to zero. The phase difference becomes the control signal used to maintain the TSM at null in a closed-loop POGA.
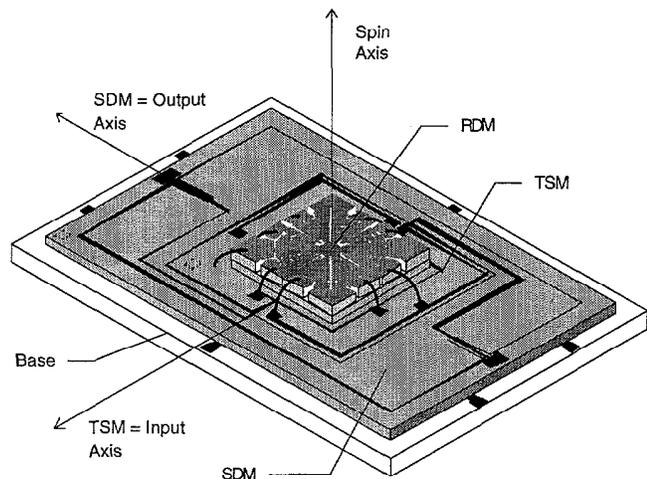
Figure 2. Schematic drawing of the micromachined POGA, showing the principal components and rotational axes.

DEVICE DESCRIPTION

The RDM of the micromachined POGA is a 4.5 mm diameter silicon member formed by etching entirely through a two inch wafer. It has two sets of four rotational electrostatic comb drives for in-plane actuation. The RDM is supported by four 50 micron wide silicon flexures that are anchored to an under layer of silicon. The spacing between the comb fingers is 50 microns. The comb drives and anchors are electrically isolated from each other by thermally grown silicon dioxide.
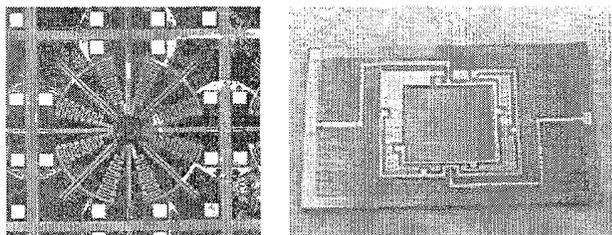
Figure 3. The left image shows a single RDM in an array of devices. Each RDM has a diameter of 4.5 millimeters and is formed on a 6x6 mm silicon die connected by tabs. The right image is an SDM/TSM assembly with a recess for mounting the RDM on the TSM. Aluminum conduits isolated by an oxide layer run on the surface. The SDM frame measures 13x20mm.

The TSM and the SDM are silicon members etched from the same 2 inch silicon wafer as a single unit. The TSM is inset within the SDM. The TSM flexures are 50 microns wide and 1 mm long connecting the TSM to the SDM. The SDM flexure dimensions are varied to match the SDM resonant frequency to the RDM resonant frequency. These flexures attach the SDM to a frame that is mounted to a supporting base. The SDM has 100 micron holes perforating the silicon structure to reduce the squeeze film damping between the SDM and the supporting glass base. The TSM requires damping to reduce the vibration sensitivity of the TSM, so no perforations are necessary. The back side of the SDM and TSM are etched to create the necessary gap to allow motion of these members. This gap was varied between 10 and 50 microns. Smaller gaps required less drive voltages but allow smaller mechanical motion. The front side has electrical conduits patterned on the surface separated from the substrate by a layer of silicon dioxide.
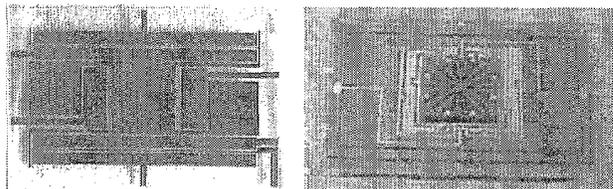
Figure 4. The left image shows the base electrode configuration. The right image is an assembled POGA: RDM mounted on TSM, SDM mounted on base, 15 x 22mm.
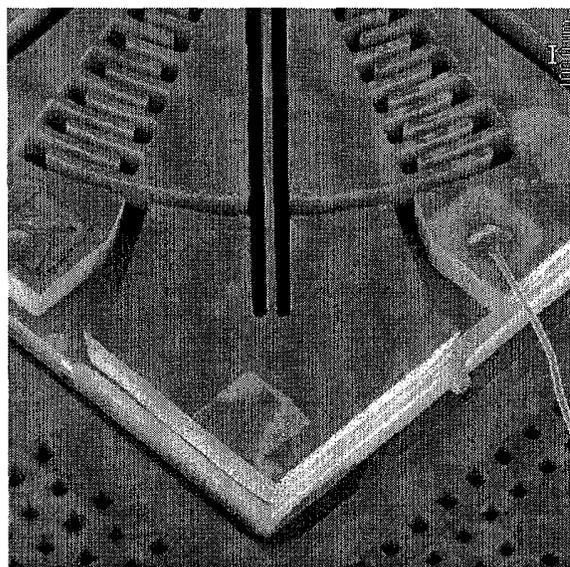
Figure 5. SEM of a quadrant of RDM. The RDM flexure runs vertically in the image between the radial comb drives. A wire bond from the comb drive is visible on the right side of the image. The edges of the two bonded layers and the perforations in the SDM/TSM layer are visible.

Figure 3 shows an RDM still in a device array on the wafer. Small tabs are left during the etching process to interconnect the components. These are easily cleaved to dice out the individual components. The second image in Figure 3 is an SDM/TSM assembly. The SDM axis of rotation is horizontal in the image with electrical connections running to the perimeter. The TSM axis of rotation is vertical in the image with the electrical conduit running across the surface of the flexures.

The SDM/TSM assembly is mounted on a base that provides electrodes for actuation and sensing of the mechanical members. The drive electrodes for the SDM are positioned near the edge of the member to increase the torque generated. The

SDM sensing electrodes are positioned just inside the drive electrodes. The TSM sensing electrodes are placed under the TSM at the edge of the member to maximize sensitivity to rotations. Both sensor systems use differential capacitor readout electronics to monitor the position of the SDM and TSM. A ground plane is serpentined between the electrodes to reduce cross-talk between the capacitors. Figure 4 shows an image of the base with the electrode layout visible. The second image of Figure 4 is an assembled POGA. The RDM mounts on the TSM. The RDM itself creates the pendulosity of the TSM.

A scanning electron microscope image of an assembled POGA is shown in Figure 5. Due to the relatively large size of the device, only a portion of the instrument is visible. The image is an isometric view of the RDM mounted on the TSM.

FABRICATION PROCESS

The micromachined POGA is fabricated in three separate assemblies, the RDM, the SDM/TSM assembly and the base. The RDM and SDM/TSM assembly were etched out of silicon wafers in a Plasma Therm inductively coupled plasma reactive ion etcher (ICP-RIE). Characterization and discussion of the procedure can be found elsewhere [6]. The base is formed in glass with patterned metal electrodes on the surface.

The RDM fabrication sequence is shown in Figure 6. A two inch p-type silicon wafer is oxidized to produce a one micron layer of oxide surrounding the wafer. The oxidized wafer is bonded to another non-oxidized wafer using standard techniques [7]. The bonded wafers are now electrically isolated by the oxide layer between them. The wafers are then oxidized. The top wafer is then patterned and etched down to the buried oxide layer defining the RDM, its flexures, and its radial comb drives. The wafer combination is then etched from the backside to release the RDM mechanical structure, yet retain the anchors for the flexures and comb drive. The oxide is then removed from the exterior surfaces and aluminum is patterned on the comb drive anchors by using a shadow mask.

The shadow masks were also produced using the ICP. Holes were etched through silicon wafers to produce the masks. When the mask is aligned to the wafer, aluminum was allowed to deposit on the component only in the required areas. The shadow masks were produced for both device level metalization and wafer level metalization.

The fabrication sequence of the SDM/TSM assembly is shown in Figure 7. A p-type two inch silicon wafer is oxidized to produce a one micron layer of oxide. Windows are etched in the oxide to provide contact regions to the silicon. Aluminum is deposited on the wafer, sintered then patterned. The backside of the wafer is etched to produce the standoff height between the base metal layers and the silicon mechanical layers. The wafer is then etched entirely through its thickness from the front side to define the SDM and TSM.

The bases are formed in glass. Aluminum is deposited on the glass with a titanium adhesion layer. The metals are then patterned to produce the SDM drive electrodes, and the sensing electrodes for the SDM and TSM.

The three components are then assembled to form the instrument. The assembly procedure is shown in Figure 8. The SDM/TSM assembly is temporarily mounted on a mesa support structure. The mesa is used to handle the component and support the movable members during assembly. The RDM is positioned on the TSM and epoxied in place. A recess in the TSM aids in the alignment of the RDM to the TSM. Electrical connections are made from the RDM by wirebonding. The SDM/TSM assembly is released from the supporting mesa and epoxied to the glass base. The base is then wirebonded to the package.
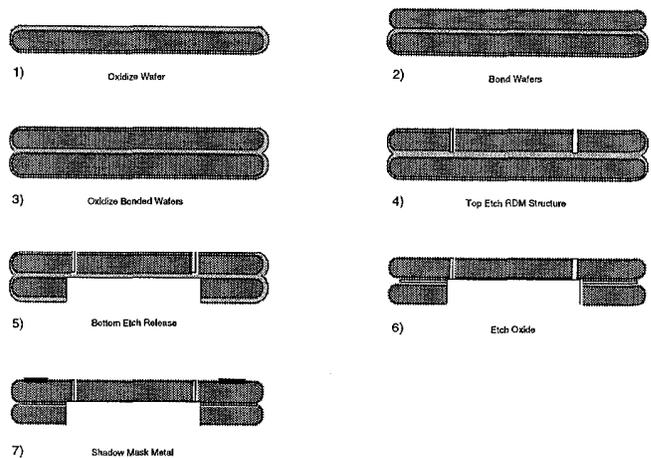
Figure 6. RDM fabrication sequence.
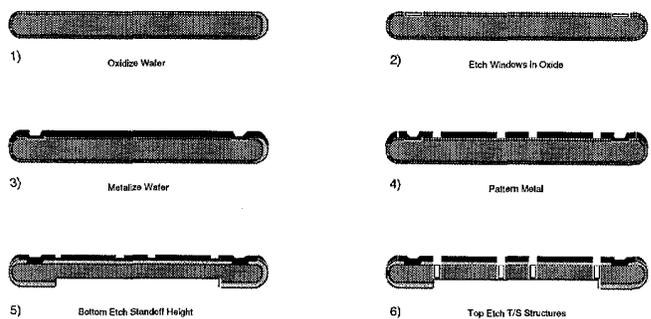
Figure 7. SDM/TSM assembly fabrication sequence.
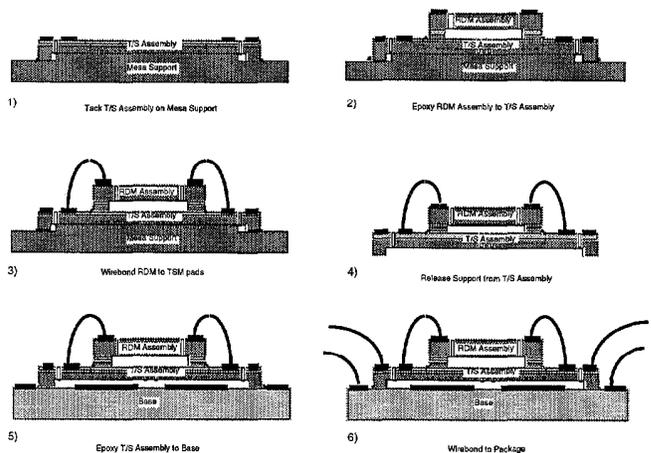
Figure 8. Micromachined POGA assembly procedure.

RESULTS

Open loop, the POGA operates just like any other pendulous accelerometer. Under acceleration, the pendulous mass torques the member against the supporting flexures. This motion is detected with the differential plate capacitors for the TSM. Figure 9 shows an open loop tumble test of the bulk micromachined POGA and an ADXL150, an Analog Devices

50g micromachined accelerometer. The instruments are mounted on a dividing head and rotated at 10° steps through 360°, measuring the component of the earth's gravity along the input axis of the accelerometers.
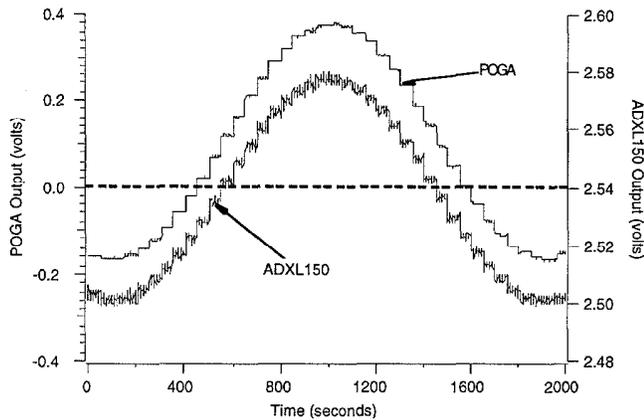
Figure 9. Open loop tumble test of ADXL150 and micromachined POGA. Rotated 10° steps at 50 second intervals measuring a component of earth's gravity.
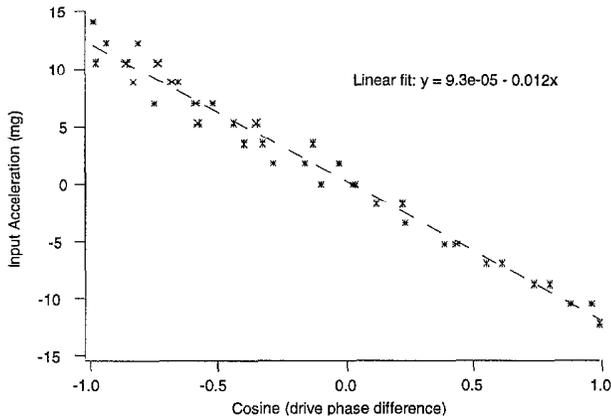
Figure 10. Closed loop scale factor test. Input acceleration nulled by adjusting phase between input drive voltages.

Closed loop, the POGA operation differs from other accelerometers. The innermost member, the RDM, generates an oscillating angular momentum. The drive frequency is set at the resonant frequency of the RDM. A bias voltage is applied to the sinusoidal signal to offset the drive so that it is always applying a positive potential to remove the frequency doubling effect of capacitive electrostatic actuators. The oscillating angular momentum is the functional equivalent to the constant angular momentum of a spinning wheel gyro. In the PIGA, the pendulous mass is torqued back to the null position by spinning the gyro to generate the gyroscopic torque. Similarly, the SDM of the POGA oscillates to create a torque that nulls the TSM. The drive voltage of the SDM is phase locked to the RDM drive. A bias voltage is also applied to remove frequency doubling effects. By controlling the relative phase and magnitude of the RDM and SDM oscillations, the middle pendulous torque summing member can be held at null.

Figure 10 shows the results of the closed loop scale factor test. The POGA was positioned such that the input axis of the accelerometer was vertical on a dividing head. The dividing head was then rotated to produce an input acceleration. The

input acceleration was sensed by the TSM causing a shift in the output voltage. The voltage was then returned to the null position by adjusting the phase difference between the phase locked function generators used to drive the RDM and SDM. The input acceleration was calculated by taking the sine of the angle on the dividing head. Equation (3) shows the input acceleration should be a linear function of the cosine of the phase, experimentally confirmed in Figure 10.

SUMMARY AND CONCLUSIONS

The POGA uses the same fundamental physics that has been demonstrated to produce the highest performing accelerometer, the PIGA. The scale factor is determined only by mechanical quantities, the pendulosity and angular momentum, rather than by a precision voltage supply as in other servoed accelerometers.

The fabrication methods demonstrate the use of the ICP as a machining tool in addition to a micromachining tool. Application specific layers were produced then assembled into an instrument. The stacked layer design in conjunction with bulk micromachining is ideally suited for modern inertial sensors where reduced cost and high performance are required.

ACKNOWLEDGMENTS

REFERENCES

1. M. S. Sapuppo., "Pendulous Oscillating Gyro Accelerometer: POGA," *Joint Services Data Exchange for Guidance, Navigation & Control*, 24th, Anaheim, CA, Nov., 1998.

2. M. S. Sapuppo, "Pendulous Oscillating Gyro Accelerometer," U. S. Patent #5,457,993, October 17, 1997.

3. A. Lawrence, *Modern Inertial Technology*, Springer-Verlag, New York, 1998.

4. G. R. Pitman, (Ed.), *Inertial Guidance*, Wiley, New York, 1962.

5. M. Fernandez and G. R. Macomber, *Inertial Guidance Engineering*, Prentice-Hall, Englewook Cliffs, NJ, 1962.

6. A. A. Ayon, R. Braff, C. C. Lin, H. H. Sawin, and M. A. Schmidt, "Characterization of a Time Multiplexed Inductively Coupled Plasma Etcher," *J. Electrochem. Soc.,* Vol. 146, No. 1, pp. 339-349, 1999.

7. M. A. Schmidt, "Silicon Wafer Bonding for Micromechanical Devices," *Technical Digest of the 1994 Solid State Sensor and Actuator Workshop*, Hilton Head, SC, pp127-131, 1994.

THE POLYCHROMATOR:
A PROGRAMMABLE MEMS DIFFRACTION GRATING FOR SYNTHETIC SPECTRA

G. B. Hocker, D. Youngner
Honeywell Technology Center
Plymouth, MN 55441-4799

E. Deutsch, A. Volpicelli, S. Senturia
Massachusetts Institute of Technology
Cambridge, MA 02139

M. Butler, M. Sinclair, T. Plowman
Sandia National Laboratories
Albuquerque, NM 87185

A. J. Ricco
ACLARA BioSciences
Mountain View, CA 94043

ABSTRACT

We report here the design, fabrication and demonstration of an electrostatically actuated MEMS diffractive optical device, the Polychromator grating. The Polychromator grating enables a new type of correlation spectrometer for remote detection of a wide range of chemical species, offering electronic programmability, high specificity and sensitivity, fast response and ruggedness. Significant results include: (1) The first demonstrations of user-defined synthetic spectra in the 3 – 5 μm wavelength regime based upon controlled deflection of individual grating elements in the Polychromator grating; (2) The first demonstration of gas detection by correlation spectroscopy using synthetic spectra generated by the Polychromator grating.

INTRODUCTION

MEMS technology enables many innovative optical devices. The Polychromator grating is an array of a large number of long, narrow optically reflective elements whose vertical positions are electrically controlled. If the element length is many optical wavelengths, the width a few optical wavelengths, and vertical positions controllable to a fraction of a wavelength, the Polychromator can function as a programmable diffractive optical device. Specifically, as described below, the Polychromator grating elements can be positioned to generate a desired multi-line, or polychromatic, spectrum at a given diffraction angle. Furthermore, this synthetic spectrum can be designed to match the infrared absorption spectrum of a molecule and used in a correlation spectrometer for the spectroscopic detection of that species.

MEMS DEVICE DESIGN

The Polychromator gratings reported here are designed to generate synthetic spectra in the 3-5 μm wavelength range. A grating consists of 1024 individually addressable diffractive elements each 10 μm wide by 1 cm in length. The vertical position of a grating element is to be controlled over approximately a 2 μm range with an accuracy of about 0.1 μm. Because of the difficulty of providing individual inputs to 1024 lines, the first devices fabricated used only 132 pads for signal and ground, compatible with a readily available pin-grid-array package. In order to actuate all 1024 lines with this limited I/O capability, the grating elements are hard-wired on-chip as 8 repeat units of 128 elements, such that a voltage applied to a pad actuates 8 lines, one from each repeat unit. This on-chip hard wiring, while advantageous for device packaging and initial testing, has a deleterious effect on device yield, since if any one of the eight lines connected to a pad is compromised, the entire set of 8 lines is non-functional.

Nevertheless, successful results have been obtained using this configuration.

The deflection of each element is a function of the applied voltage, and the maximum travel is limited by an electromechanical instability called pull-in. For the beam lengths and widths used in this design, and with beam thickness comparable to the free-space gap, actuation via a continuous underlying electrode would result in pull-in occurring at approximately two-thirds of the free-space gap. In order to achieve positioning beyond the pull-in instability, the design for the Polychromator device takes advantage of leveraged bending[1]. The element design shown in Figures 1 and 2 uses a two-beam structure attached periodically to the substrate, in which the lower beam undergoes bending while the top mirror beam deflects vertically without flexion. Leveraged bending of the lower beam by voltage applied to the split actuation electrodes located near the support posts allows controllable positioning of the upper mirror beam through the full gap. Electromechanical analysis is used to design structures capable of near-full-gap actuation with voltages of about 100 volts or less. The gaps and beam thicknesses are about 2 μm. For the devices reported here, post-to-post spacings are in the range of 500 to 700 μm, and electrode lengths are in the range of 80 to 100 μm.

DEVICE FABRICATION

The Polychromator grating device is fabricated by polysilicon surface micromachining. Two polysilicon layers are used for the actuated structure, and two for electrical interconnections and actuating electrodes. The process begins by growing a thermal oxide on an ultra-flat silicon wafer. Two layers of boron-doped polysilicon separated by silicon nitride are used for the interconnection and electrode structures. Next a relatively thick sacrificial SiO_2 layer is deposited, which will define the actuation gap. Support posts for the bending layer of polysilicon are patterned and etched in this SiO_2 layer. Before depositing the bending layer of polysilicon, small "dimples" are etched in the SiO_2. The dimples cause nodules to form on the bottom of the polysilicon, and the nodules prevent stiction during operation should the bending beam pull in. The bending layer of polysilicon is deposited, implanted, and patterned, followed by the deposition of the second sacrificial SiO_2 layer. The second SiO_2 layer has vias etched in it to create support posts for the top polysilicon mirror beam layer. The top polysilicon mirror-beam layer is then deposited, implanted and patterned. Typical thicknesses for the bending and mirror beam layers of polysilicon and for the two sacrificial SiO_2 layers are approximately 2 μm. Next the wafers are annealed to set the polysilicon stress. The process is completed by depositing gold for the mirrors and pads, then removing all of the sacrificial SiO_2. The fabricated element and grating are shown in Figures 3, 4, and 5.

GRATING ACTUATION RESULTS

The individual diffractive elements of the Polychromator grating are actuated electrostatically. Beam deflection depends on the applied voltage, electrode length, distance between support posts, beam thickness, residual stress, and free-space gap. Combinations of mask design and process parameters are selected to result in structures which maximize travel while maintaining moderate actuation voltage. The vertical position of each element is determined by the applied voltage. To evaluate fabricated devices, interference microscopy is used to measure vertical movement of individual beam elements as shown in Figure 6. Using this technique, the deflection of each beam is measured as a function of applied voltage. Results for several grating element designs are shown in Figure 7. Measured maximum beam deflections for these devices are nearly the full gap of two microns, and pull-in occurs at voltages ranging from 65 to 110 volts, depending on device designs.

OPTICAL OPERATION AND SYNTHETIC SPECTRA

When infrared radiation strikes a grating, the phase shift of the wavefront on reflection from a given diffractive element depends upon vertical position of that element. The diffracted radiation is the sum of the wavefronts diffracted from the individual grating elements. Since the vertical position of the elements of the Polychromator grating can be controlled by the voltages applied, it is possible to program the Polychromator grating to diffract infrared radiation in a desired manner. A spectral design algorithm has been developed that calculates the displacement profile for the diffractive elements that will generate desired spectra at a specified angle[2]. This algorithm can be used to design novel diffraction gratings that are complex spectral filters with any spectral content desired. They can be used to synthesize the absorption spectra of specific molecules [3] and to function as the reference spectra in a programmable correlation spectrometer [2]. More complex spectral designs can be used to enable real-time multi-variate analysis which can distinguish between chemical species with overlapping spectra. This development extends the application of diffractive optics from spatial-domain optical processing into the spectral domain.

For the Polychromator grating, the spectral design algorithm was used to design the deflection profiles required to produce desired spectra in the $3 - 5$ μm wavelength regime. Using measured voltage/displacement data for the grating such as those shown in Figure 7, the deflection profile was converted to a set of voltages that are applied to the groups of elements.

We have for the first time generated synthetic spectra with a programmable MEMS diffraction grating. Two such designed spectra in the infrared are shown in Figure 8; their designs and use in detection of CO_2 will be described in the following section.

CORRELATION SPECTROSCOPY AND GAS DETECTION

A correlation spectrometer correlates the spectrum from a sample with a reference spectrum, providing sensitivity and specificity, fast response, and mechanical and optical simplicity. The need for a reference cell containing a sample of the target species limits its usefulness. It is awkward to change target species by changing the reference cell, and the technique is impractical for transient or highly toxic chemical species. In addition, a large amount of background light is transmitted in spectral regions with no optical absorption, adding to the background level without contributing to the desired signal. Use of the Polychromator grating to generate the reference spectrum enables programmable detection of specific gases species including transient and toxic species. Furthermore, the Polychromator grating can be programmed to produce the complement of the optical absorption spectrum of the target species to be detected. This removes the light transmission at frequencies with no optical absorption, thereby enhancing the signal-to-background ratio.

The optical layout for the correlation spectrometer is shown in Figure 9. The infrared source is an 1100K black body with the light passing through a 10 cm gas cell. The light is collected by a CaF_2 lens and focused on the entrance slit. Subsequently, the light is recollimated by a spherical mirror and directed onto the Polychromator surface. The diffracted light is collected at the desired angle by another flat mirror and then focused on the exit slit. The detector is a liquid-nitrogen-cooled indium antimonide detector (EG&G Optoelectronics). A high-voltage analog driver instrument (HVADI), custom designed by Digital Designs and Systems, Inc. (dideas.com), serves as the drive electronics for the Polychromator. The HVADI provides 128 control lines, each capable of delivering $0 - 200$ volts. The HVADI's microprocessor and RS-232 port allow for control of the instrument and for up-loading of files to generate various grating profiles. The HVADI can modulate the Polychromator by switching between two such profiles – a requirement for correlation spectroscopy.

The rate of switching between the two grating profiles can be varied. A flat response is observed for frequencies below 2 kHz with roll-off at higher frequencies. Since the resonant frequency of the beams is in the tens of kHz range, we believe the roll-off is associated with either an RC time constant formed by the capacitance of the beam and the resistance of the lead providing the voltage, or the current capacity of the instrument driving the beams.

The two synthetic spectra shown in Figure 8 were designed for the detection of CO_2. The solid line is a bandpass centered on the CO_2 absorption band at 2350 cm^{-1}, while the dashed line is a double bandpass that samples bands on either side of the CO_2 absorption line. The integrated intensity of the transmitted light for the two spectra is designed to be equal. Switching between these two profiles generates a modulated signal only if CO_2 is present in the gas cell. The modulated output from the detector when the gas cell was alternately filled with pure nitrogen and 1000 ppm of CO_2 in nitrogen is shown in Figure 10. This is the first demonstration of gas detection by correlation spectroscopy using synthetic spectra generated by the programmable Polychromator grating.

The fact that the 1024 beam elements are hard wired in 128 groups of eight lines has two important effects on device operation. First, as already stated, there is a reduction in device yield. The results shown here are from a grating with only 25% of the sets of 8 beams functional. The fact that successful synthetic spectra and chemical detection could be performed with such a device demonstrates the robustness of the Polychromator against localized defects. When all 1024 lines can be independently actuated, functional beam yield will increase significantly because an isolated fault will no longer affect 8 beams. It is estimated that the diffraction efficiency increases as the square of the number of functional elements. Thus, a considerable improvement in performance is expected for a completely functional device.

A second effect of the hard wiring is that the periodicity imposes a sampling in wavelength space for the synthetic spectra, limiting somewhat the spectral features that can be implemented.

This limitation will be removed when a packaging scheme capable of driving all 1024 lines is implemented.

CONCLUSIONS

We have demonstrated the Polychromator, a programmable MEMS diffraction grating capable of generating synthetic spectra. The innovative electromechanical design of the electrostatically controlled grating elements uses leveraged bending to allow controllable actuation over the nearly the full gap. Fabrication utilizes multilevel polysilicon surface micromachining. A design algorithm is used to calculate the displacements of the grating elements that will create a desired diffraction spectrum, along with the voltages required to create these displacements. Synthetic spectra have been generated as designed in the 3 – 5 µm wavelength regime. Two such designed spectra have been used to detect CO_2 in the first demonstration of correlation spectroscopy using synthetic reference spectra generated by a programmable MEMS device.

The Polychromator concept will lead to an instrument for stand-off detection of many chemical species. Advantages include electronic programmability, specificity to the desired species and immunity to interference, and fast response, in a miniature, rugged, low-power instrument.

Future efforts will include improvements to grating element yield to improve diffraction efficiency, and development of Polychromator gratings and control electronics offering individual control of 1024 grating elements for improved spectral resolution.

ACKNOWLEDGEMENTS

REFERENCES

1. E. S. Hung, S. D. Senturia, "Extending the Travel Range of Analog-Tuned Electrostatic Actuators," *Journal of Microelectromechanical Systems, 8,* 497 (1999).

2. M. B. Sinclair, M. A. Butler, A. J. Ricco and S. D. Senturia, "Synthetic Spectra: A Tool for Correlation Spectroscopy," *Applied Optics, 36,* 3342 (1997).

3. M. B. Sinclair, M. A. Butler, S. H. Kravitz, W. J. Zubrzycki and A. J. Ricco, "Synthetic Infrared Spectra," *Optics Letters, 22,* 1036 (1997).
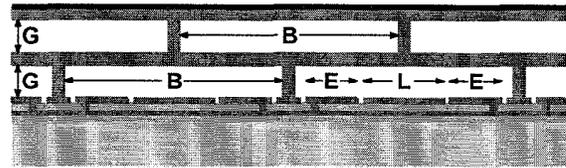
Figure 1. View of Polychromator grating beam structure in cross-section for unactuated state (B is the doubly-supported beam length, E is the length of one actuating electrode, L is the ground pad length, and G is the lower and upper free-space gap).
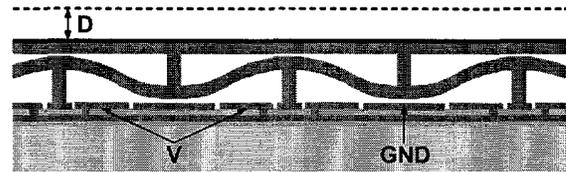
Figure 2. View of beam structure in cross-section for actuated state The lower beam undergoes bending, while the top mirror beam remains flat and deflects vertically (D is the vertical displacement from the original position when voltage V is applied to the actuating electrodes.
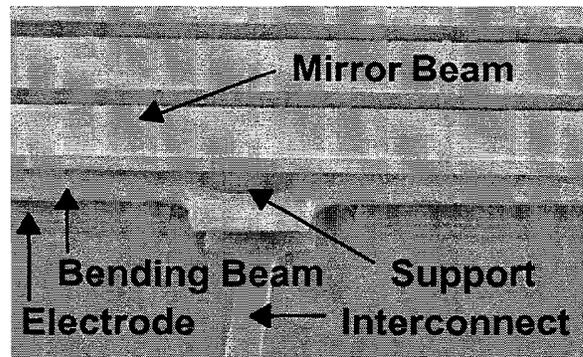
Figure 3. SEM of MEMS grating showing the double beam design, support structure, interconnects, and electrodes.
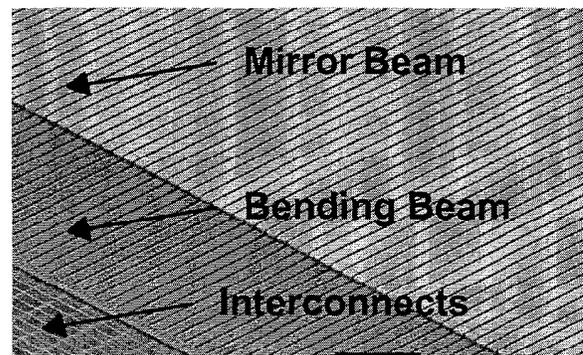
Figure 4. SEM of the edge of a MEMS Polychromator grating showing interconnects, lower bending beam, and upper mirror beam. Beam widths are 10µ, and they extend 1 cm in length.
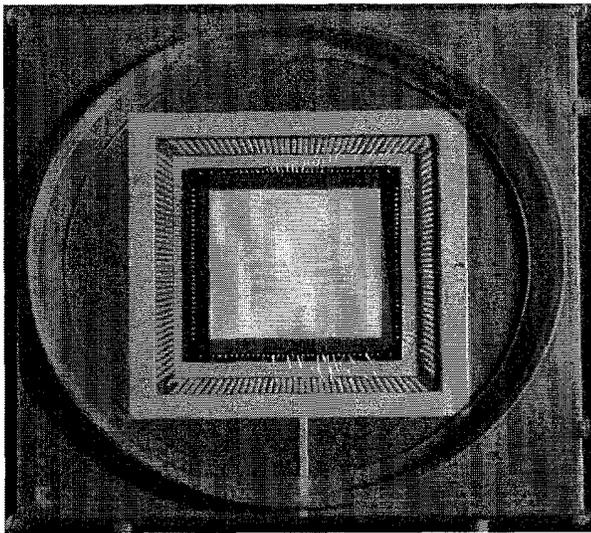
Figure 5. Packaged Polychromator die. Array is 1 cm x 1 cm.
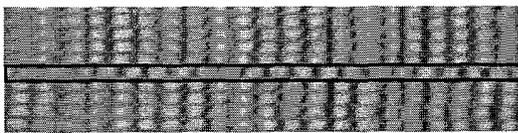
Figure 6. Fringe shift of actuated beam using interference microscopy.
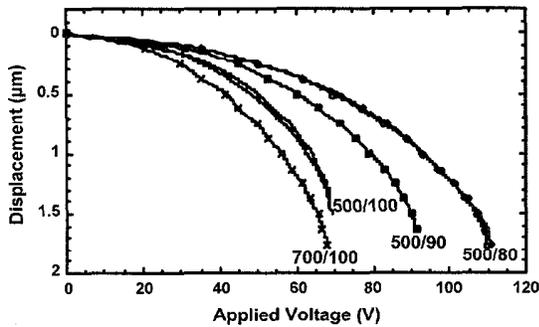
Figure 7. Displacement (μm) versus actuating voltage (V) for a variety of grating element designs (Numbers in the figure indicate the beam length to electrode length relationship B/E in microns for each design).
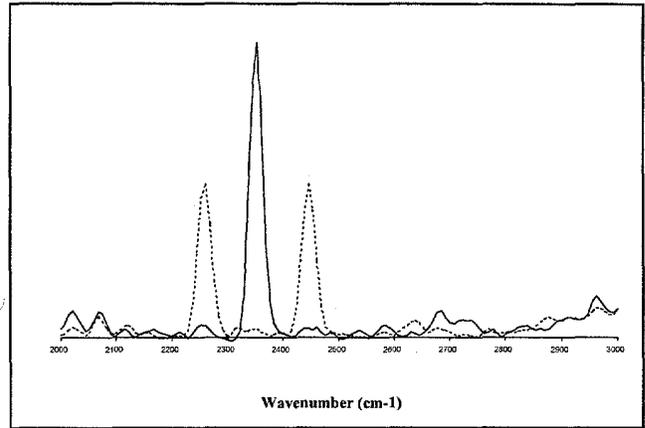
Figure 8. Spectral transmission of the Polychromator for the two grating profiles used for CO_2 detection. The solid line spectrum matches the CO_2 absorption at 2350 cm^{-1}, while the double-bandpass dashed line spectrum samples an equal area outside the CO_2 absorption. These spectra were measured with a Nicolet 750 FTIR spectrometer.
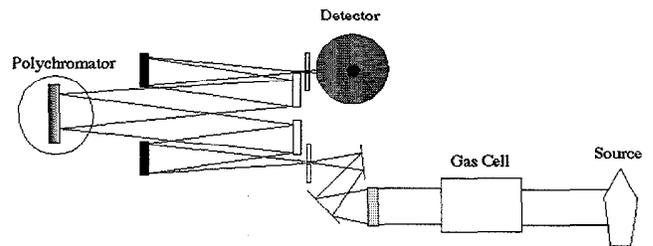
Figure 9. Diagram showing the optical configuration of the Polychromator-based correlation spectrometer. Further details are given in the text.
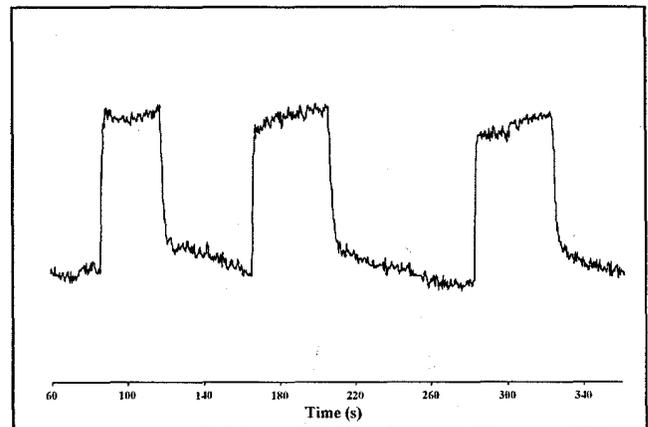
Figure 10. Response of the correlation spectrometer to 100 ppm-meters of CO_2 admitted and purged from the gas cell 3 times.

92

DEVELOPMENT OF A SILICON TWO-AXIS MICROMIRROR FOR AN OPTICAL CROSS-CONNECT

Andrew S. Dewa and John W. Orcutt
Telecommunications Business Unit, Materials & Controls Group
Texas Instruments
Dallas, TX 75243

Marshall Hudson and David Krozier
Telecommunications Business Unit, Materials & Controls Group
Texas Instruments
Attleboro, MA 02703

Alan Richards and Herzel Laor
Astarté Fiber Networks
Boulder, CO 80301

ABSTRACT

This paper describes the development of a two-axis silicon micromirror for a 480 × 480 optical cross-connect (OXC). The micromirror is an electromagnetically actuated, analog silicon mirror. The range of motion for the micromirror is ±8°. Full range of motion open-loop moves have been demonstrated in under 5 ms, which enables a full 480 ×480 switch reconfiguration in 10 ms. The micromirror and OXC have been optimized for low loss. The average optical insertion loss is −2.5 db at 1550 nm and −2.9 db at 1310 nm, the two primary wavelengths in fiber optic data transmission. SONET data at 2.4 Gbits/sec has been transmitted through the OXC with no loss penalty.

The preliminary life and environmental validation data for the NxN Micromirror is reported. Bare micromirror samples have passed both the GR-63-Core vibration specification (1-100 Hz 0.5G, 3-axis; 100-500 Hz, 3G, 3-axis) and temperature extreme specification (-40°C to 70°C) for devices in their shipping containers. We also have shock tested a small sample of micromirrors to failure. These fail at 800-900G shock. Micromirrors have been cycled to full deflection for 30 million cycles (6× specified life for the product) with no change in performance.

INTRODUCTION

The need for all-optical switching is being driven by exponentially increasing demand for bandwidth and the use of dense wavelength division multiplexing (DWDM) in all optical networks. As the number of wavelengths carried on each fiber grows, current methods of optical/electrical/optical (OEO) switching become unfeasible. In OEO switching, the light from one fiber must be first separated into its individual wavelengths. The time domain multiplexed signal on each wavelength is demulitplexed down to lower data rate electrical signals. The electrical signals are then switched. The signals are mulitplexed back up in data rate and the wavelengths are recombined for transmission to the output fiber. As the number of wavelengths per fiber grows, the space and power consumption of OEO quickly becomes too great.

In an all-optical cross connect (OXC), no electrical conversion is needed. All wavelengths are switched simultaneously. At present, all announced, large-port-count (N≥256) OXCs are MEMS-based, because of the ability of this technology to produce high-precision, micromechanical moving mirrors. Typical single-mode optical fiber cores are 9 µm in diameter, so that the emerging optical beam is compatible with the size of typical MEMS optical components. The actual size of the micromechanical mirror depends on a complicated relationship of angle of view, the beam diameter, optical magnification, and the length of the optical path. Furthermore, since each mirror needs individual control and an optical fiber interface, the control electronics together with optical connector interfaces to support hundreds to thousands of micromirrors can dwarf the optical MEMS. In addition, the telecommunications industry requirements on system reliability mandate redundant power supplies and system controllers, so that the final size of the OXC system can be several equipment racks.

Texas Instruments (TI) and Astarté Fiber Networks (AFN) are jointly developing a 480 × 480 OXC. The TI/AFN approach utilizes individually-packaged, micromachined silicon mirrors to perform the switching function and active optical feedback to maintain low-loss, long-term connections. This approach provides a high degree of modularity, so that the customer can purchase the number of ports needed now, and further populate the switch frame as demand increases. Modularity also makes the OXC field serviceable–a failed module can be replaced without disrupting traffic on other modules.

This paper will discuss the development of the two-axis silicon micromachined mirror. It will include a general description of the OXC architecture, the fabrication and operation of the micromirror, preliminary life and environmental validation testing results, and the optical performance of the micromirror in the switch matrix. For clarity, the silicon mirror will be referred to as the NxN Micromirror throughout the rest of the paper.

OPTICAL CROSS CONNECT

Texas Instruments and Astarté Fiber Networks are jointly developing a 480 x 480 free-space, optical cross connect (OXC) [1,2]. The OXC is single-stage and non-blocking, allowing any-to-any connectivity. TI is developing the Optical Unit (OU) capable of switching the optical signals from one fiber to another, while AFN has developed the overall OXC architecture. The NxN Micromirror provides the beam steering function for the OU. The OU also contains the focusing optics and the servo control emitters and detectors. AFN combines 4 OU's with a control electronics board into an Optical Module. The optical module is the minimum granularity of the OXC. Two hundred and forty modules, arranged in a 5 by 24 array make up a fully populated switch. All the optical units are pointed at a fixed mirror one meter away to give a 2-meter optical path length. A schematic showing four Optical Modules is given in Fig. 1.

The fundamental optical path of the OXC is shown in Fig 2. The light exits the fiber and goes through focusing optics. The expanded beam is reflected off a fixed mirror to the NxN Micromirror. The beam is steered by the NxN Micromirror to the
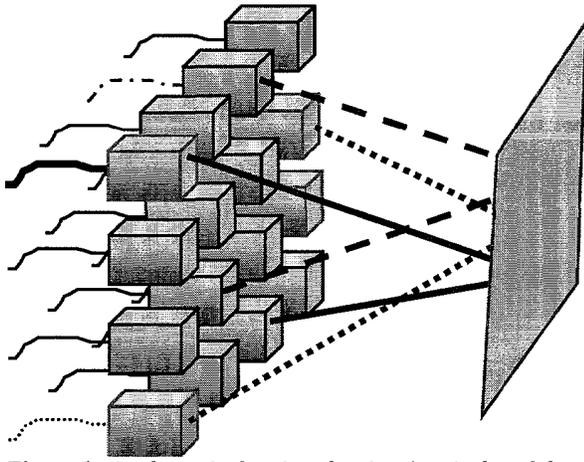
Figure 1. A schematic drawing showing 4 optical modules.

target OU via a large fixed mirror. This folding of the optical path allows for a smaller OXC footprint. It also allows for the doubling of the size of the switch, by replacing the mirror by another array of optical modules.

The two Micromirrors are servo controlled by emitters and detectors using the same optical path as the light from the fibers. Active servo control is key to achieving and maintaining very low insertion loss. The active servo control means that the connection is not susceptible to vibration, shock and long term drifts.

NxN MICROMIRROR FABRICATION

The NxN Micromirror is an electromagnetically-driven, analog, two-axis silicon mirror. A photograph of the mirror die is
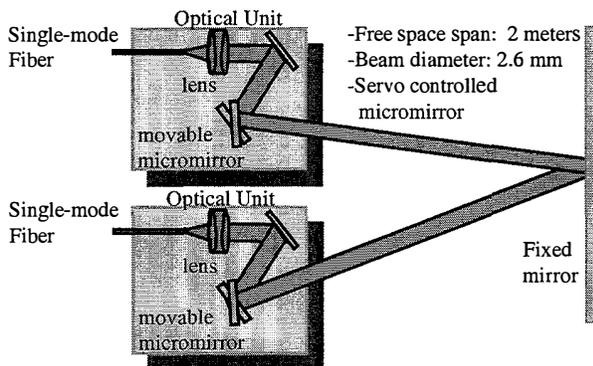
Optical emitters and detectors for servo control not shown

Figure 2. The optical path of the 480 × 480 OXC.

shown in Fig. 3. The active part of the silicon mirror is a 3.2 × 3.6 mm oval which is supported by 2 sets of silicon torsional hinges, one set between the mirror and the gimbals, and the second set between the gimbals and the frame. The size of the mirror is set by the beam diameter in the optical path. The two sets of torsional hinges allow the independent movement of the mirror in two axes. The overall die size is 7.2 mm × 9.0 mm. The mirror is about 0.1 mm thick and is gold coated to maximize its reflectivity at the wavelengths of interest for telecommunications, 1.25 μm to 1.65 μm.

The NxN Micromirror is fabricated using photolithography and reactive ion etching via the Bosch Process [1]. The etch recipe was optimized for vertical sidewalls with little undercut. The typical difference between the width of the torsional hinges at the
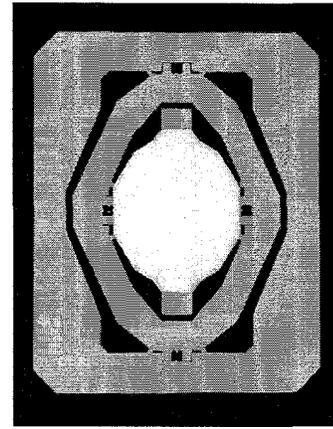
Figure 3. Optical photograph of an NxN Micromirror die.

top of the wafer and the bottom of the wafer is 0.2 μm for the 100μm etch depth.

The packaged NxN Micromirror is shown in Fig. 4. The package was partitioned into a hermetic portion, the ceramic header and window that holds the NxN Micromirror, and a non-hermetic portion, where the coils for the actuator, wiring harness and mounting bracket are located. The micromirror is hermetically packaged to guarantee the 20-year life expected by the
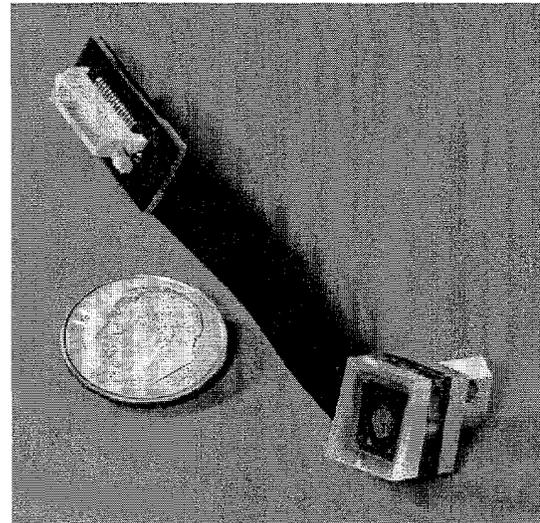
Figure 4. A photograph of the packaged NxN Micromirror. A US dime is shown for scale.

telecommunications industry.

MICROMIRROR ACTUATOR

The electromagnetic actuator for the mirror utilizes drive coils below the micromirror and permanent magnets mounted on the mirror and gimbals frame, as shown in Fig. 5. This actuator was chosen to allow the tips of the mirror and the gimbals to move several hundred microns providing the required deflection of ±8° in both axes with the millimeter-size of the micromirror.

Figure 6 shows actual deflection data from an NxN Micromirror prototype plotted on the magnetic model curves. Because there is no ferromagnetic material in the magnetic circuit other than the permanent magnets, analytical methods could be used to simulate the actuator [4]. This model, along with finite element analysis, was used to optimize the actuator design. The
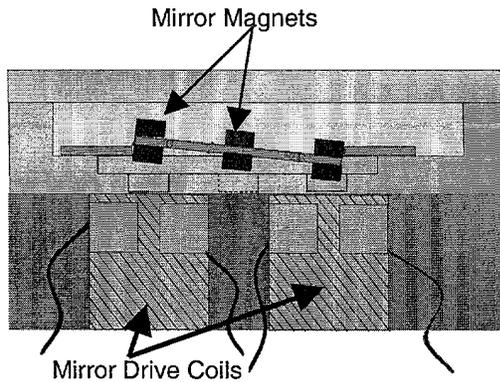
Figure 5. *Schematic drawing of NxN Micromirror actuator mechanism.*

Figure 7. *Measured position and current waveform for an open-loop move of 15.6° in under 5 ms (Note: on the Y-axis the angle is in radians and the current is in amperes).*

mirror and gimbals hinges are identical in geometry and fall on the same straight line. Perfectly linear behavior is what is expected for single crystal silicon hinges. The model fits the experimental data to within the experimental errors of the measurement.

The OXC makes a connection by first steering the beam to the new location as fast as possible with open-loop control. Then an acquisition takes over to quickly move the remaining distance to the center position, at which point the closed-loop servo control takes over. Since the hinges are perfectly elastic, the micromirror is an underdamped system. Therefore, the dynamics of the system were modeled and a specified waveform is generated for each move. A move of 15.6° of the mirror plus gimbals under open-loop drive is shown in Fig 7.

The mirror plus gimbals moved 15.6 degrees in under 5 ms and was stable. The stability shown in Fig. 7 is more than enough for the acquisition and servo control to take over and optimize the connection. The under 5 ms move time is also sufficient to guarantee a full OXC reconfiguration in 10 ms.

ENVIRONMENTAL AND LIFE VALIDATON TESTING

Key to deploying a product in the fiber optic telecommunication market is demonstrating a reliable product which meets the Telcordia specifications. We have begun validation tests on the environmental and life of the NxN Micromirror. As one would expect with single-crystal silicon hinges, cycling mirror over its full range of motion in both axes for over 30 million cycles results in no measurable change in the

micromirror performance. Thirty million cycles is six times the product life specification of one reconnect every minute for 20 years. Since each cycle is move forward and back, one represents cycle two connections.

We have tested the NxN Micromirror to the GR-63-Core shipping temperature specification -40°C to +70°C. The initial test samples have passed this test with no change in operating characteristics. Further testing is underway with additional samples to improve out confidence levels.

Upon seeing the NxN Micromirror, the initial reaction of most engineers is to question its shock and vibration sensitivity—because there are relatively large magnets at the ends of long moment arms supported by thin silicon hinges. To validate the design, we have begun testing the NxN Micromirror to the more severe GR-63-Core specification for devices shipped in turboprop aircraft. It should be emphasized that the GR-63-CORE specification is for devices in their shipping containers, including all padding, etc. We tested 4 samples of the NxN Micromirror bolted directly to the shaker table to this specification: 1-100Hz at 0.5G and 100-500 Hz at 3G in all three principal axes. (The standard GR-63-CORE shipping test is 100-500 Hz at 1.5G.) All samples passed the test. Again, further testing is underway to improve our confidence levels.

We have also tested NxN Micromirrors to failure under shock. Three mirrors passed a 500G shock test with a 6 ms duration, in all three axes. We took 2 of the mirrors to failure under 800-900 G shock tests with a duration of 2.5 ms. The duration of the shock at higher G levels was reduced due to equipment limitations. Further shock testing is underway.

The environmental and life validation testing results to date are summarized in Table 1.

Figure 6. *Measured hinge force vs. mirror deflection data plotted on the magnetic force model.*

Table 1. *Environmental and life test results*

Life Testing	30 million cycles (3× expected life)
Shipping Thermal Extremes	-40°C to 70C with no detectable change in performance
Shock survivability	Survives 500G 3-axis, 6 ms duration
Shock failure	800-900G 2.5 ms duration
Vibration test (GR-63-Core in shipping container)	1-100 Hz 0.5G, 3-axis; 100-500 Hz, 3G, 3-axis

OPTICAL PERFORMANCE

The OXC is designed to operate at the two major wavelengths utilized in optical fiber communications, 1310 nm and 1550 nm. The coatings on the optical unit lens, optical elements and the micromirror window are optimized for the band-pass from 1250 to 1650 nm so that the OXC is fully compatible with dense wavelength division multiplexing.

The measured optical insertion loss at 1310 nm averaged –2.9 db with a maximum value of –3.2 db. At 1550 nm the insertion loss averaged –2.5 db with a maximum value of –3.0 db. Further testing of the optical loss is ongoing.

One key metric for an optical switch is to measure any excess data loss in a real telecommunications data channel other than that due to the loss of optical power measured by its insertion loss. This is done by comparing the data loss rate for the OXC to that of a passive optical attenuator with the same attenuation as the measured insertion loss of the connection. This is called the optical loss penalty. The OXC was able to transmit 2.4 Gbits/sec. of SONET data with no measurable loss penalty.

The measured optical performance of the NxN Micromirror is

Table 2. *NxN Micromirror OXC Performance*

Optical Loss (1310 nm)	-2.9 db ave.; -3.2 db max.
Optical Loss (1550 nm)	-2.5 db ave.; -3.0 db max.
2.4 Gbits/sec SONET Data transmission	No Measurable Loss Penalty
Time for a full range move	5 ms for 15.6° move (measured)
Time for a full switch reconfiguration	10 ms (projected)
Deflection	±8° in two independent axes

summarized in Table 2.

SUMMARY

We have demonstrated a two-axis silicon micromirror for a 480 × 480 optical cross connect. The micromirror is electromagnetically actuated, and can move, open-loop, 15.6° in under 5 ms to enable a full switch reconfiguration in 10 ms. The average optical insertion loss for the optical switch is –2.5 db at 1550 nm and –2.9 db at 1310 nm. The optical switch has transmitted 2.4 Gbits/sec SONET data with no loss penalty.

We have begun life and environmental testing of the NxN Micromirror. Bare micromirror samples passed the GR-63-Core vibration specification and temperature extreme specifications for devices in shipping containers. We also shock tested a small sample of micromirrors to failure. These failed at 800-900G. We tested micromirrors to 30 million cycles (3× expected life) with no change in performance. Further life and environmental testing is ongoing to get enough data to fully predict the life and reliability of the NxN micromirror.

REFERENCES

1. D. Krozier, M. Hudson, J. D'Entremont, H. Laor, A. Richards, and E. Fontenot, "Performance of a 576 × 576 Optical Cross Connect", *Technical Proceedings of the National Fiber Optic Engineers Conference*, Chicago, Illinois, September 26-30, 1999, vol. 1, p276.

2. H. Laor, A. Richards, E. Fontenot, "576x576 Optical Cross Connect for Single-Mode Fibers", *Conference Proceedings of the 1999 Annual Multiplexed Telephony Conference*, San Diego, CA, July 19-22, 1999, pp. 343-349.

3. F. Laermer and A. Schilp, Robert Bosch GmbH, "Method for Anisotropic Plasma Etching of Substrates", *U.S. Patent 5498312* (1996).

4. E. P. Furlani, "A Three-dimensional Field Solution for Axially Polarized Multipole Disks," *Journal of Magnetism and Magnetic Materials, 135*, 205-214 (1994).

A LARGE-DEFLECTION ELECTROSTATIC ACTUATOR FOR OPTICAL SWITCHING APPLICATIONS

John D. Grade * **and Hal Jerman**
Seagate Research
San Jose, CA 95131

Thomas W. Kenny
Department of Mechanical Engineering, Stanford University
Stanford, CA 94305

ABSTRACT

Fundamental improvements in the design of comb-drive ac-tuators have allowed their use in a 1xN optical switch requiring 200 um static deflection and a switching speed of less than 1 ms. The main obstacle to the development of large-displacement comb-drive actuators has been electromechanical side instability, which increases with forward deflection and may cause catastrophic failure. In this work, the stable deflection range is increased by employing prebent suspensions and linearly engaging comb teeth. Nonlinear analytical models, as well as Mechanica^{TM} finite element models and Simulink^{TM} numerical models have been developed to fully characterize the performance of the actuators and facilitate their design. A simple, high-yield, deep-reactive-ion-etching fabrication process has produced consistent performance that closely matches FEA and numerical models of the electrostatic actuators. Using separately-fabricated, highly-reflective mirrors, the actuators have been utilized in a compact optical design that features high-speed, low optical-loss switching.

INTRODUCTION

There has been considerable interest in devices for switching light from one optical fiber to another or from a free-space optical beam to one or more optical fibers, particularly for telecommunications. Several switch configurations are of interest, including 1x2, 1xN, and NxN, where N is a number from 2 to above 64. The dominant optical switch designs have utilized optical waveguides, electro-optical effects, or electromechanical actuators, and working switches are commercially available using these techniques.

Of the 1xN electromechanical optical switches, the most common design uses a single rotary actuator to "steer" the optical beam to a particular output location. An alternate switch architecture uses multiple linear actuators to intercept and redirect the optical beam into a selected output fiber [1]. Figure 1 shows an example of such a switch with four 45 degree mirrors attached to independent linear actuators, each capable of diverting the collimated beam to a particular output. While this switch architecture does not require the precision servo-control of the beam-steering designs, it does require each linear actuator to travel at least the diameter of the collimated beam, which is greater than 100 um for practical optical layouts.

Traditional designs for linear comb drive actuators show limited forward displacement due to electromechanical side instability. Side instability occurs when the parallel-plate electrostatic forces between the sides of the combs overwhelm the suspension stiffness orthogonal to the desired motion [2-5]. This problem is exacerbated during large displacements, because with increasing forward displacement, a traditional folded-beam suspension's side stiffness decreases while the electrostatic side forces increase.

* e-mail: John_D_Grade@notes.seagate.com   Phone: (408)952-5036
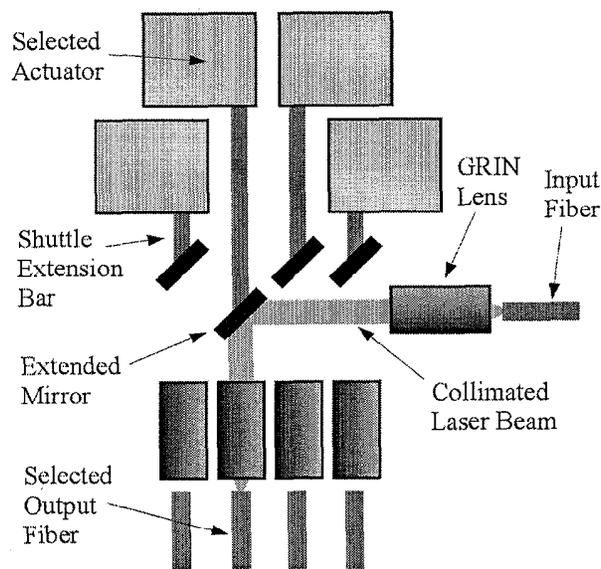
Figure 1. *Schematic diagram showing a 1x4 optical switch. Light leaving the input fiber is collimated by a GRIN lens and travels left until it intercepts an extended mirror and is focused onto an output fiber.*

This paper describes a linear actuator, fabricated using deep reactive ion etching, that has been optimized for use in optical switching systems such as that shown in Figure 1. In an example application for an optical data storage system, the actuator must be as small as possible with the ability to move a mirror a total distance of 200 um in less than 1 ms while using less than 150V.

SWITCH DESIGN

In the example data storage application, at least 80% coupling efficiency is required with switching speeds below 1 ms. Few existing optical switches meet these specifications, and none are available that also fit within the 18 cm^3 device volume. The proposed optical switch requires an actuator for each output fiber. A collimated laser beam passes down the switch's central axis, as shown in Figure 1. During operation, one actuator is extended while all others are retracted. A mirror attached to the actuator serves to redirect the laser beam into the selected output fiber. The mirror width should be twice the collimated beam width to reduce diffractive losses, and the actuator must have a total deflection that is at least the width of the mirror. The difficulty in designing large-deflection actuators provides a motivation to reduce the collimated beam diameter. However, as the beam diameter is reduced, the distance over which the beam remains collimated is also reduced. The divergence of the optical beam results in a reduction

in the achievable coupling efficiency. The proposed optical switch uses 250 um diameter GRIN lenses to produce a 100 um diameter collimated beam. The divergence of the beam over the length of the switch is then small enough allow up to 80% coupling efficiency.

ACTUATOR DESIGN

Several novel features facilitate the design of the large-deflection electrostatic actuators that are required for the proposed optical switch. As shown in Figure 2a, conventional comb drives are typically symmetric about the translation axis, having four folded-beam suspension springs [6]. The proposed actuator is not symmetric about the translation axis, allowing it to consume only half the area of a traditional design. Since a one-sided design has half the mass, half the springs, and half the combs (see Figure 2b), both the suspension stiffness and the electrostatic force are reduced by a factor of two. Thus neither the resonant frequency nor the maximum static deflection are changed. Alternatively, a one-sided design may occupy the same area as a traditional comb-drive, but use the area more efficiently to achieve a larger stable forward deflection range. Note that in Figure 2b, the suspension has been placed exterior to the combs to improve the low rotational stiffness of the suspension shown in Figure 2a.

(a)   (b)   (c)

(d)   (e)   (f)
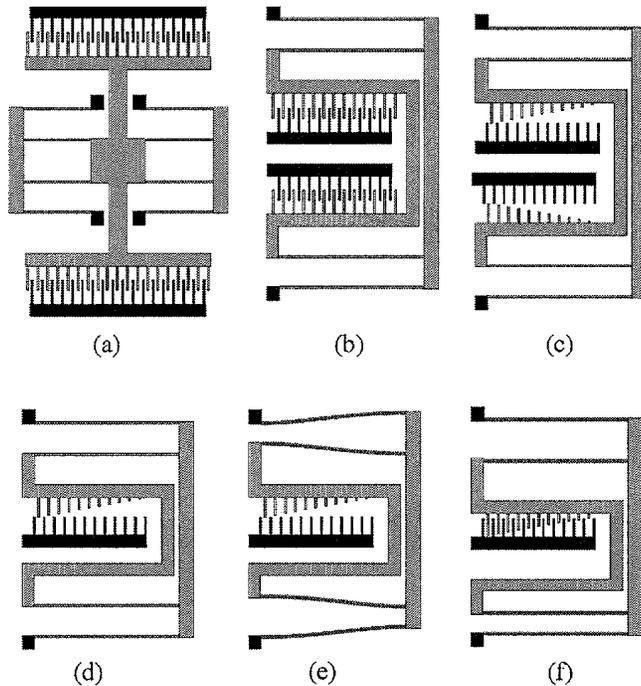
Figure 2. *Schematic diagram of a traditional comb-drive actuator (a), a one-sided actuator (b), an actuator with linear engagement comb teeth (c), an actuator with a parallel-plate electrode (d), an actuator with a prebent suspension as-fabricated (e), and an actuator with a prebent suspension after forward deflection (f).*

In the proposed optical switch, the actuators are only required to fully extend or retract. Switching speed and positional accuracy may be increased by using mechanical stops to define the two actuator states. The actuator then drives the shuttle to press against either forward or rear crash-stops that are fixed to the substrate. The crash stops allow the electrostatic force versus the deflection profile to be customized. While the parametric design of non-rectangular comb teeth has been described [4, 7], a simpler way to

modify the displacement dependence of the electrostatic force is to adjust the lengths of the individual comb teeth. By linearly increasing the length of the combs that form the moving electrode, the number of engaged comb teeth varies linearly with forward displacement (see Figure 2c). The forward electrostatic force is then

$$F = \frac{(N-1)\varepsilon h x}{g x_{\max}} V^2 + \frac{\varepsilon h}{2g} V^2, \quad (1)$$

where N is the number of comb teeth, ε is the dielectric constant of the air gap, h is the device thickness, x is the forward displacement, g is the electrostatic gap, and V is the applied voltage. This linear engagement comb drive achieves the maximum forward deflection, x_{max}, at the same voltage as a corresponding traditional comb drive. However, with only half the electrostatic overlap area, the linear engagement design suffers less from electrostatic instability since it only generates half the detrimental parallel-plate forces between the sides of the comb teeth.

The use of mechanical stops to define the shuttle position also allows the design of an actuator with only one bank of opposing comb teeth (see Figure 2d). In the proposed actuator the motor is retracted 110 um by a set of comb teeth; it is extended by releasing the retracted shuttle and allowing it to glide freely to a set of mechanical stops 90 um in the extended direction, where it is captured by a pair of parallel plate electrodes. By using the back side of the fixed combs to form the fixed parallel plate, the resulting compact actuator (about 1 mm^2) only requires a single electrical drive pad. The actuator changes state when the drive voltage is removed, and the new state is locked in when the voltage is reapplied
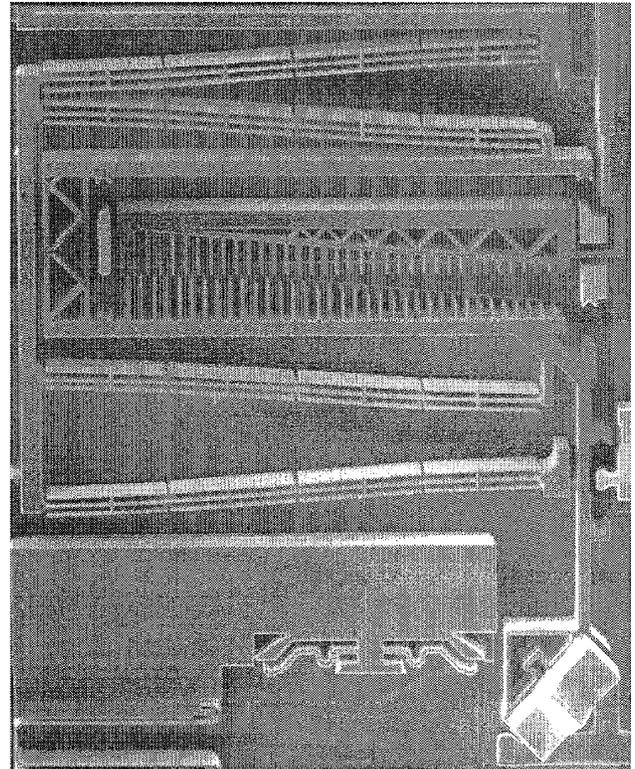
Figure 3. *SEM image of the actuator used in the optical switch. The four "H" shaped fins attached to each suspension beam help control the sidewall profile without changing the spring stiffness.*

Finally, with only one bank of opposing comb teeth, the folded-flexure suspension may be customized. The side stiffness, k_{side}, of a traditional folded-flexure suspension falls with the square of the forward displacement [2, 8]:

$$k_{side} = \frac{Ehw}{L}\left(\frac{3x^2}{8w^2}+1\right)^{-1},\qquad(2)$$

where E is the Young's modulus of silicon, x is the forward displacement, and L, w, and h are the length, width, and thickness of the individual suspension beams. The situation may be dramatically improved by fabricating the suspension beams in an initially bent configuration, as shown in Figure 2e. A cosine with amplitude $x_{pb}/4$ is an admissible function for the initial bend profile. This effectively replaces x by $(x-x_{pb})$ in Equation 2. The side stiffness increases as the shuttle moves forward, reaching a maximum when the beams are straight at $x=x_{pb}$ (see Figure 2f). In a typical actuator of the proposed design, the use of prebent suspension beams increases the minimum side stability over the actuation range by a factor of thirty. A SEM of an actuator similar to that of Figure 2e is shown in Figure 3. Note that a separately fabricated mirror has been attached to a mirror holder on a shuttle extension bar.

FABRICATION

All of the front-end wafer processing for the actuators and optical switches was carried out at TRW Lucas NovaSensor, in Fremont, California. The five masks required for the process were provided to TRW Lucas NovaSensor and the parts were delivered in wafer form. The wafers were then tested and diced prior to assembly and packaging. The simple fabrication process kept performance variations across 4 wafers (more than 500 switches) below 5%, with a defect rate of only 7%. While the proprietary details of the TRW Lucas NovaSensor process have not been disclosed, their generalized fabrication process has been described [9] (see Figure 4).

Etch shallow cavity in carrier wafer to create future movable areas.

Fusion bond device wafer to carrier; polish to final thickness.

Oxidize; open contact holes; deposit and pattern pad metal.

DRIE etch through device wafer.
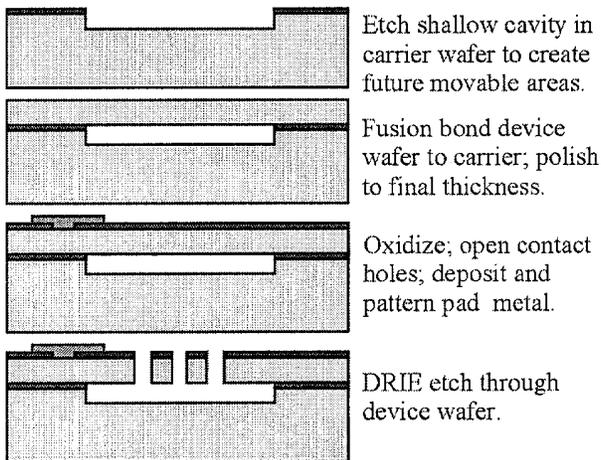
Figure 4. *Schematic diagram showing the fabrication process carried out at TRW Lucas Novasensor.*

ACTUATOR PERFORMANCE

After fabrication, dicing, and assembly, the actuators (see Figures 2e and 2f) were tested both statically and dynamically. The static response shown in Figure 5, was generated by increasing the drive voltage while measuring the forward deflection with an optical microscope at 1000x magnification. Unlike a traditional comb drive, the displacement of a linear engagement design is not

directly proportional to the square of the voltage. The maximum forward deflection is defined by the placement of the forward crash stops.
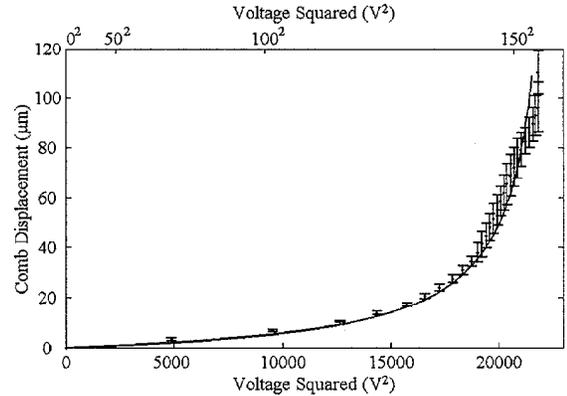
Figure 5. *Comparison of the measured (circles) and predicted (line) quasi-static response of a linear-engagement actuator with a pre-bent suspension.*

A laser vibrometer was used to quantify the resonant behavior of the actuators. Figure 6 shows the swept sine frequency response of an actuator and the response predicted by a Simulink™ model of the actuator as a 2 DOF, second-order system. The fundamental resonance occurred at 995 Hz with a quality factor of 57. The resonance at 4 kHz results from the contrary motion of the shuttle and the suspension truss. These dynamic modes, as well as the additional mode at 9 kHz, match those predicted by a Mechanica™ finite element model.
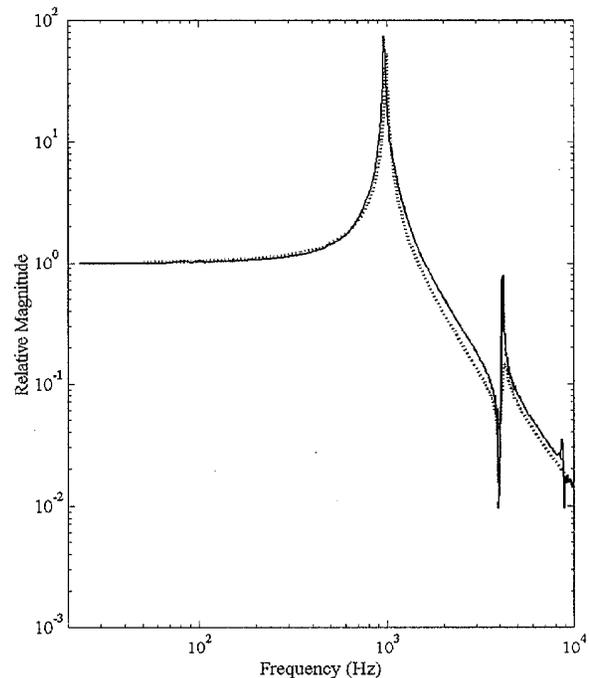
Figure 6. *Measured and predicted frequency response magnitude. The solid line represents data gathered using a laser vibrometer. The dotted line represents the response predicted by FEA and numerical models using Mechanica™ and Simulink™.*

SWITCH PERFORMANCE

Optical switches from 1x2 to 1x16 have been fabricated, and a portion of an assembled switch is shown in the SEM of Figure 7. Additional polarization detection optics are present in this particular switch to facilitate its use in an optical data storage system.
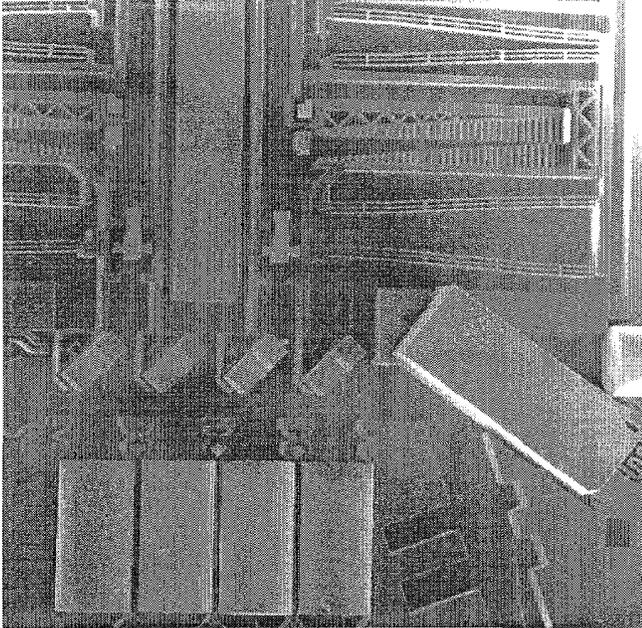
Figure 7. *SEM view of a portion of a partially assembled optical switch showing the actuators, the inserted mirrors, and the cylindrical GRIN lenses held in place with integral alignment springs. Additional polarization detection optics are present for use in a data storage system.*

In order to achieve the highest optical efficiency, separately fabricated mirrors are inserted into holders in the linear actuators. These mirrors have high reflectivity with minimal scattering loss, which is an improvement over approaches that use the etched sidewall as the mirror surface **[10]**. Switching losses of less than 1 dB at a wavelength of 650 nm (and less than 0.5 dB at 1550 nm) have been demonstrated using single mode optical fibers.
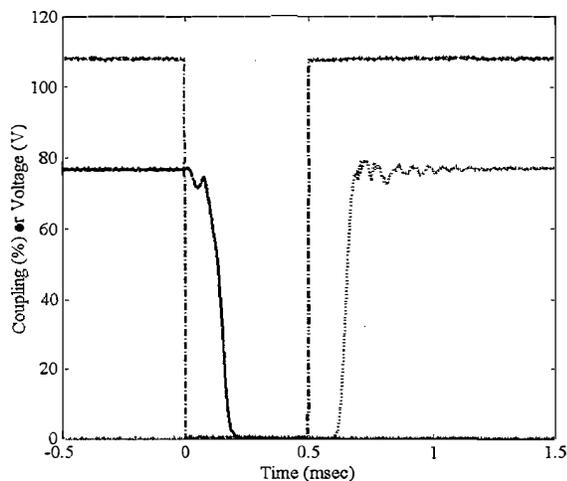
Figure 8. *Measured drive voltage (dash-dot line) and coupling efficiency as light is switched from one fiber (solid line) to another (dotted line).*

The coupling efficiency and switching speed were measured using a high-speed power meter and a digitizing oscilloscope. Since the resonant frequency of the actuators is typically 1 kHz, the switching time is slightly greater than 0.5 ms, as shown in Figure 8. Note that the actuator begins to change state when the drive voltage is removed, and the new state is locked in after the voltage is reapplied. The previously selected actuator and the newly selected actuator change state simultaneously, due to the same voltage signal.

SUMMARY

Through substantial improvements in design, comb-drive actuators consuming only 1 mm^2 have been demonstrated with stable, static deflections of 200 um and switching speeds of less than 1 ms. The simple, high-yield, fabrication process has produced consistent performance that closely matches FEA and numerical models of the electrostatic actuators. These actuators have been utilized in a compact optical design that features high-speed, low-loss switching.

ACKNOWLEDGEMENTS

REFERENCES

[1] J.H. Jerman, et. al., "Electrostatic Microactuator and Method for Use Thereof", U.S. Patent 5,998,906 (1999).

[2] R. Legtenberg, et. al., "Comb-Drive Actuators for Large Displacements", *J. Micromech. Microeng.*, 6 (1996), p. 320-329.

[3] C. Marxer, et. al., "An Electrostatic Actuator with Large Dynamic Range and Linear Displacement-Voltage Behavior for a Miniature Spectrometer", *Transducers '99*, p. 786-789.

[4] J. Mohr, et. al., "Micro Optical Switching by Electrostatic Linear Actuators with Large Displacements", *Transducers '93*, p. 120-123.

[5] C. S.-B. Lee, et. al., "Multiple Depth, Single Crystal Silicon MicroActuators for Large Displacement Fabricated by Deep Reactive Ion Etching", *1998 Solid-State Sensor and Actuator Workshop*, Hilton Head, SC, p. 45-50.

[6] W. Tang, et. al., "Laterally Driven Polysilicon Resonant Microstructures", *Sensors and Actuators*, 20, p. 25-32, 1989.

[7] W. Ye, et. al., "Optimal Shape Design of an Electrostatic Comb Drive in Microelectromechanical Systems", *J. Microelectromechanical Systems*, 7 (1998), p. 16-26.

[8] J. Grade, "Large-Deflection, High-Speed, Electrostatic Actuators for Optical Switching Applications", Ph.D. dissertation, Stanford University, 1999.

[9] E. Klaassen, et. Al., "Silicon Fusion Bonding and Deep Reactive Ion Etching; A New Technology for Microstructures", *Sensors and Actuators*, 52 (1996), p. 132-139.

[10] C. Marxer, et. al., "Vertical Mirrors Fabricated by Reactive Ion Etching for Fiber Optical Switching Applications", *10th Int. Workshop on Micro Electro Mech. Sys.*, p. 49-54.

MEMS-CONTROLLABLE MICROLENS ARRAY FOR BEAM STEERING AND PRECISION ALIGNMENT IN OPTICAL INTERCONNECT SYSTEMS

Adisorn Tuantranont, Victor M. Bright, Jianglong Zhang, Wenge Zhang, John Neff, and Y. C. Lee

NSF Center for Advanced Manufacturing and Packaging of Microwave, Optical, and Digital Electronics (CAMPmode), Department of Mechanical Engineering, University of Colorado, Boulder, CO, USA 80309-0427
http://mems.colorado.edu, Tel: (303) 735-1734, Fax: (303) 492-3498,email:Adisorn.Tuantranont@colorado.edu

ABSTRACT

Novel, 2-dimensional MEMS-controllable microlens array has been integrated with a Vertical-Cavity-Surface-Emitting-Laser (VCSEL) array using a flip-chip assembly. The MEMS/VCSEL hybrid system is demonstrated for efficient and reliable smart pixel arrays for board-to-board or chip-to-chip optical interconnects in digital systems. By translating polymer microlens, fabricated on a MEMS X-Y movable plate, using electro-thermal actuators, a beam steering angle of 70 mrad is achieved. VCSEL beam steering was successfully demonstrated in our MEMS/VCSEL hybrid system to collimate and steer laser beam for a precision alignment in a 2-dimensional free-space optical interconnect.

INTRODUCTION

Free-space optical interconnection is attractive for several applications including telecommunication switching networks and fine-grained parallel processing /computers. Currently optical interconnects are used for data communication over long and medium distances in wide area networks (WANs). However over short distances inside computer systems or clusters of electronic networks, board-to-board or chip-to-chip multi-GHz bandwidth optical interconnects are required to replace electrical interconnects in the near future. This replacement is necessary due to problems with signal integrity and crosstalk at high frequency in electrical interconnects. The advantages of free-space holographic interconnects include direct interconnects between boards, arbitrary interconnection patterns, multiple fan-outs, channel isolation and increased bandwidth, thus avoiding the interconnection bandwidth bottleneck of systems with strictly in-plane electronic interconnects. Furthermore, the cost of free-space interconnects

MEMS-Microlens Array

VCSEL Array   **Hologram Array**   **Detector Array**

Figure 1. MEMS-controllable microlens array for enhanced alignment in a 2-dimensianal board-to-board optical interconnect.

has been reduced substantially with the advancement of vertical-cavity surface-emitting lasers (VCSELs). However alignment challenges are inevitable in optical interconnect systems. Mechanical vibrations or thermal expansion can cause misalignment in the systems. In this paper, we propose an enhanced alignment approach for free-space interconnects without a demand for tighter assembly tolerances using a MEMS-controllable microlens array. Figure 1 shows the use of MEMS-controllable microlens array to enhance alignment for efficient and reliable smart pixel arrays on board-to-board optical interconnect in digital systems. The microlens collimates an incident VCSEL beam and the microlens actuators steer the beam to a predefined position on a hologram array by laterally translating the microlens. Several groups have recently described efforts to steer laser beam using decentered microlens techniques [1-2]. These approaches require a light-weight and small travel of the microlens components to achieve agile beam steering. Despite the number of papers on decentered microlens beam steering, all of them seem to require a large space for array of actuators such as comb-drive actuators or scratch drive actuators (SDA) to laterally translate microlens. Moreover, none of them has operated at a VCSEL wavelength of 830 nm because the MEMS devices need to be fabricated on a silicon substrate, which is not transparent at this wavelength. The device presented in our approach contains a dense array of individually controllable microlenses, allowing for operation with a VCSEL array. In this paper, the design, fabrication, MEMS/VCSEL integration, and beam steering results are presented.

MEMS-CONTROLLABLE MICROLENS

The 2-dimensional MEMS-controllable microlens array was fabricated through the commercially available surface micromachining technology (Multi-User MEMS Process-MUMPs) [3]. Photo-bisbenzocyclobutene (BCB) electronic resin, a photo-sensitive polymer, was used in our own process to fabricate micro semi-spherical polymer lenses on MEMS. Each 52.5 μm focal length microlens was fabricated directly on a MEMS X-Y movable plate with a 30 μm circular aperture as shown in Figure 2. The microlens was modeled as the thin paraxial lens. The focal length of the microlens is set by controlling the initial thickness of the photo-BCB layer and the diameter of the lens, which are defined by photolithography. The photo-BCB thickness required for a given radius of curvature for the lens was calculated by equating the volume of spherical portion representing the final lens shape to the volume of a cylinder representing the photo-BCB column before reflow [4]. Transmission of microlens is more than 95% at operating wavelength (830 nm). The lens profile is in 5% error compared to an ideal spherical lens profile. Since only small fraction of lens surface is used to refract light, the effects of spherical aberration and steering-angle-dependent off-axis wavefront aberration are minimized.
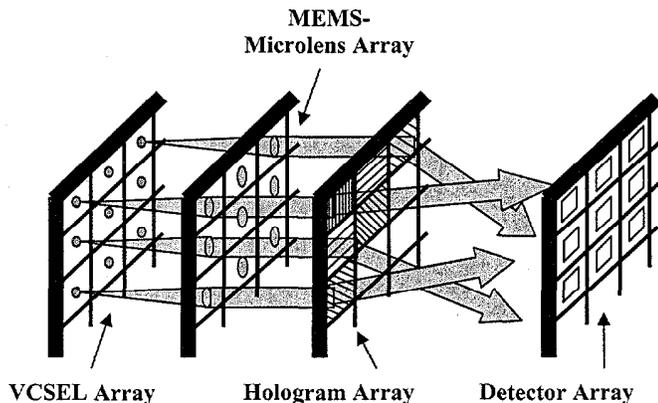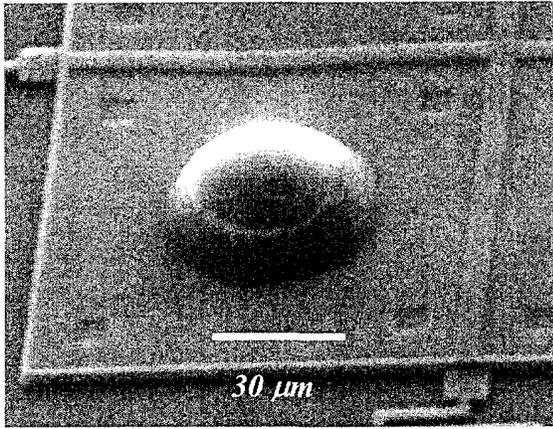
Figure 2. Polymer microlens fabricated on a polysilicon MEMS movable plate.

Two arrays of electro-thermal actuators are coupled with the 80 μm ×80 μm polysilicon movable plate to translate the microlens in X-Y plane. The microlens array is located at a focal plane in front of the VCSEL array and collimates the incident VCSEL beams. By the concept of decentered microlens, if the microlens is decentered with respect to the beam axis, then the beam will propagate through the off-axis point of microlens. The microlens still collimates the beam, but the beam is directed to a nonzero field angle as shown in Figure 3.
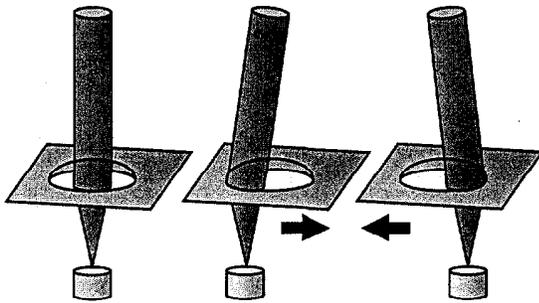
Figure 3. The decentered microlens concept of MEMS-controllable microlens array.

The steering angle depends on the lateral displacement Δd of lens with respect to the beam axis and the focal length f of the lens as shown in Figure 4. In this scheme, large steering angles are possible by large lens displacement on the order of lens radius.

$$\theta = \arctan \frac{\Delta d}{f}$$
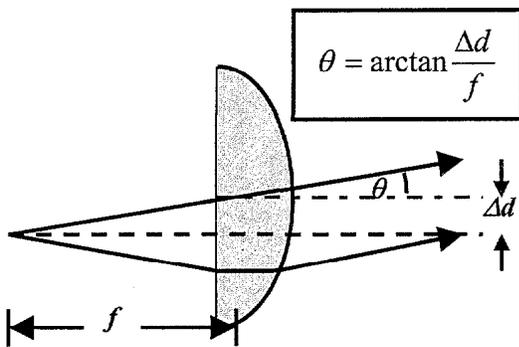
Figure 4. The steering angle of a decentered microlens.

A simulation of the MEMS-controllable microlens was conducted with ABAQUS, which is a commercially available finite element analysis tool. The model consists of an electro-thermal analysis to obtain the temperature distribution resulting from a simulated input power to electro-thermal actuators. This is then coupled to mechanical analysis in which the temperature distribution is used to determine deflections resulting from thermal expansion mismatch in hot/cold arms of actuators. The finite element simulation of the temperature distribution in actuators and movement of the plate is shown in Figure 5.
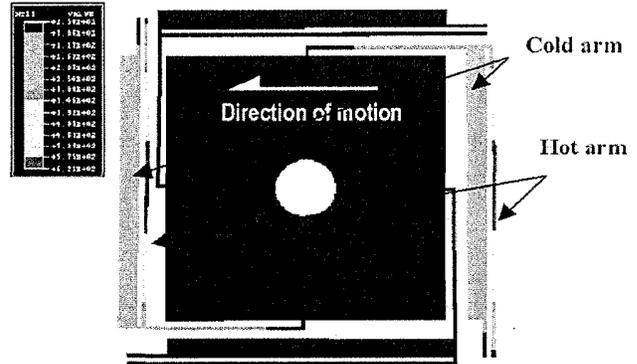
Figure 5. FEM simulation of thermal distribution and movement in electro-thermal actuators.

'U'-shaped lateral electro-thermal actuators or 'heatuators' used to drive central plate in a lateral motion consist of a single-material connected narrow arm and wider arm. Current is passed through the polysilicon actuator, and the higher current density in the narrower 'hot' arm causes it to heat and expand more than the wider 'cold' arm. The arms are joined at the free end, which constrains the actuator tip to move laterally in an arcing motion towards the cold arm side [5]. The tips of the actuators are coupled to the central plate by long thin flexures on opposite sides of the plate. Left actuator in Figure 5 pulls the plate while the right actuator pushes the plate at the same time. By push/pull mechanism of actuators, the motion of the plate is linearly proportional to the consumed power. The mechanical crosstalk in X and Y directions is minimized by using long thin flexures.

MEMS FLIP CHIP TRANSFER /INTEGRATION

The procedures for flip-chip transfer of MEMS to a non-silicon substrate (in this case quartz substrate) need two post-processing steps [6]. First, the MEMS structures are fabricated at a commercial foundry as shown in Figure 6(a), where a MEMS device is attached to the silicon substrate only by the oxide encasing it. The target substrate is patterned with gold interconnects and bonding pads designed to receive the MEMS structures, as shown in Figure 6(b). Then, in the second step shown in Figure 6(c), the two chips are bonded together by thermosonic bonding. Finally, the bonded structure is subjected to a hydrofluoric acid (HF) rinse, as shown in Figure 6(d), to free the MEMS structure and remove the silicon substrate.

The MEMS-controllable microlens array was designed and patterned in the first releasable polysilicon layer of the MUMPs process. The photo-BCB microlenses are patterned and fabricated on polysilicon plates by our own process. The plates and actuators are built upside-down on top of sacrificial oxide so that the bonding pads fabricated on MEMS are matched with bonding pads

(a) MEMS structure with microlens on silicon substrate.

(b) Matching gold bonding pads and interconnects.

(c) Flip chip and thermosonic bonding.

(d) Release of MEMS structure from silicon substrate.
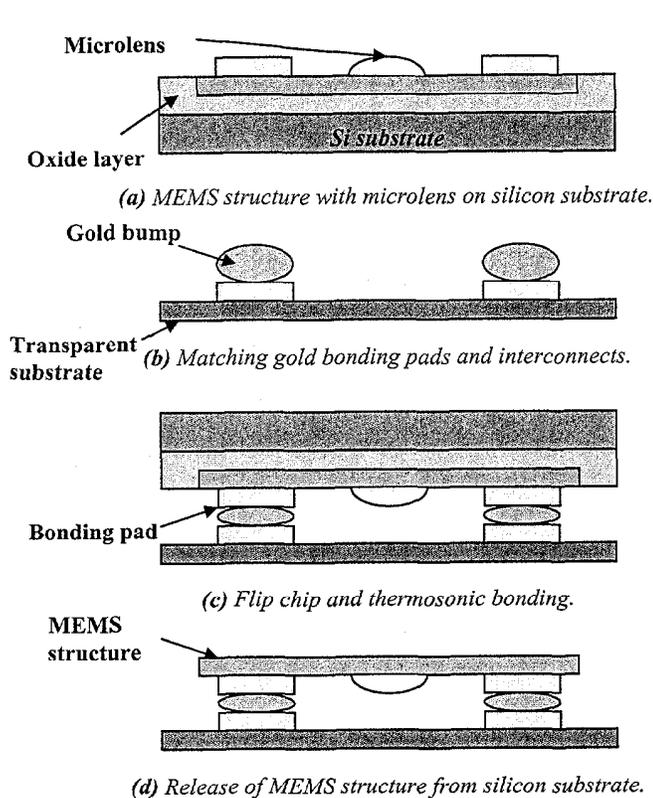
Figure 6. Flip-chip assembly process of MEMS to a non-silicon substrate.

on quartz target substrate. The quartz receiving substrate was patterned with gold wires and bonding pads for flip-chip bonding and wire bonding. A layer of photoresist is used to protect the quartz and gold interconnects when the assembly is released in HF. The layer is patterned to remove the photoresist over the bonding pads to expose the pads for flip-chip bonding. Gold bumps of 70 μm height are placed on pads and used to bond and electrically connect the MEMS structure and the wiring substrate. The MEMS-controllable microlens array is then flipped and bonded with the wiring quartz substrate using thermosonic bonding. The final step in process is to release the bonded device in a HF bath and remove the silicon substrate. The final device after flip-chip transfer to a quartz substrate is shown in Figure 7.

VCSEL ARRAY

The VCSEL array consists of an 8×8 array of top-emitting VCSELs on a 250 μm pitch. Molecular Beam Epitaxy (MBE) was used to grow the VCSEL structure on a GaAs substrate. The VCSEL structure consists of two distributed Bragg reflector (DBF) mirrors surrounding a single GaAs quantum well, and is ion implanted for current confinement in the active region. The VCSEL operates at approximately 830 nm with a bandwidth of 0.2 nm and a 2 nm wavelength variation across the array. A Guassian beam with a (e^{-2}) half-angle of 0.135 radians (2 μm in waist) is emitted from the top surface [7]. The threshold current is 1.5 mA. The VCSEL array is flip-chip bonded on a glass substrate using thermosonic bonding. The bonding process uses gold or gold-tin alloy, which is applied to the contacts. The package containing the VCSEL is soldered to a carrier board. A commercially available microlens array, which has a focal length of 560 μm and a pitch of
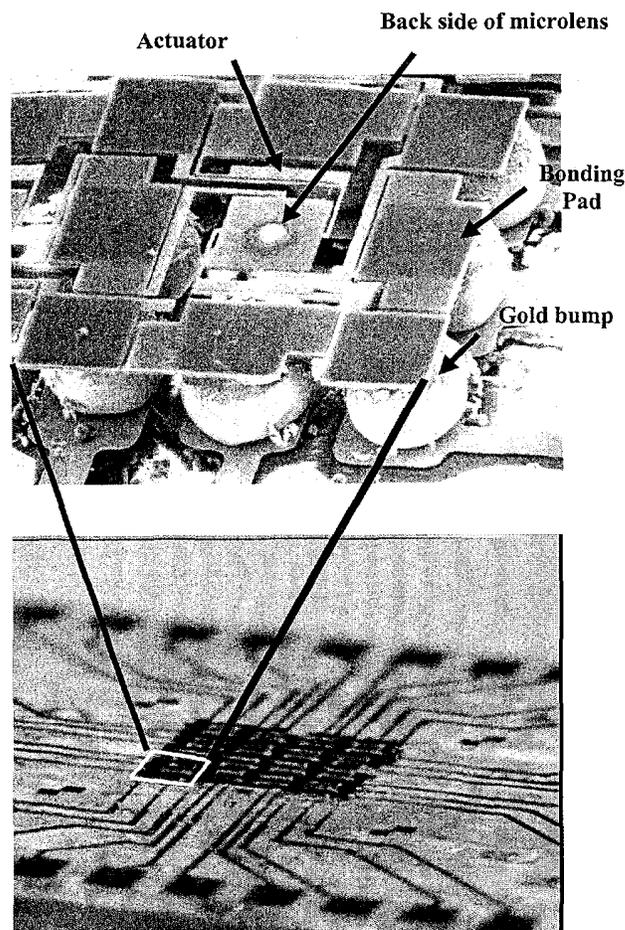
Figure 7. SEMs of the flip-chip assembled MEMS-controllable microlens 4×4 array on a quartz substrate.

250 μm is aligned with VCSEL array and attached to the opposite side of the glass substrate using a UV-curable epoxy. The resultant collimated beams are 160 μm in diameter. Figure 8 shows a packaged VCSEL array.
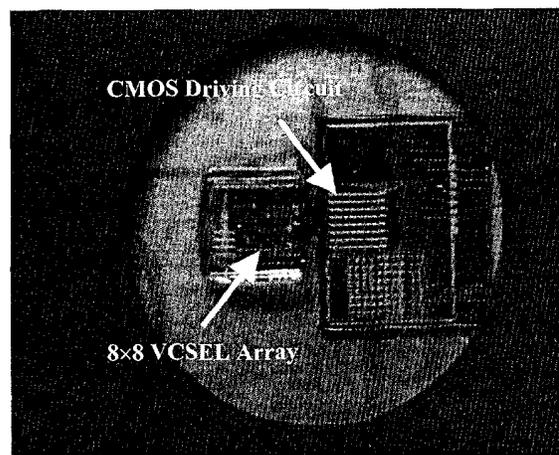
Figure 8. Flip-chip assembled 8×8 VCSEL array on a glass substrate.

103

BEAM STEERING RESULTS

The MEMS-controllable microlens array chip was actively aligned with the VCSEL array chip. First, a decentered keplerian telescope is used to achieve a collimated beam steering. The controllable microlens is placed at its focal length f_M in front of the focal plane of the lens f_L. The microlens then collimates and steers beam to the disired directions by translating in X-Y plane. Consequently, the telescope optics, which consist of positive lens ($f_1 = 100$ mm) and negative lens ($f_2 = -50$ mm) are used to magnify the steering angle by the ratio of $f_1 / f_2 = 2$). CCD camera was used to detect the deflection of the beams in the far field plane. Figure 9 shows the keplerian telescope and telescope lenses in the beam steering experiment setup. The 2-dimensional beam steering is successfully demonstrated as shown in Figure 10.

Figure 11. The maximum beam steering angle of 70 milliradians.

and the maximum beam steering of 70 mrad is achieved. VCSEL beam steering was successfully demonstrated in our MEMS/VCSEL hybrid system to collimate and steer laser beam for a precision alignment in a 2-dimensional free-space optical interconnect system.
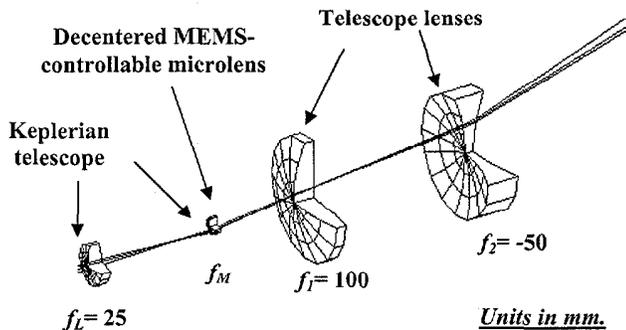
Figure 9. Optical experimental setup for beam steering demonstration.

The maximum beam steering of 70 mrad (≈ 4 degrees) is implemented by translating the microlens by one-half of the lenslet diameter as shown in Figure 11, which is adequate for enhancing optical alignment in free-space optical interconnects. Measurement verifies insertion loss of < 2 dB. The system was demonstrated up to a rate of 1 kHz of operation frequency, result of the low inertia and small travel of the microlens components.
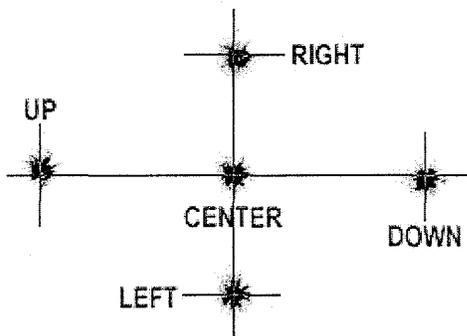
Figure 10. The 2-dimensional beam steering demonstration.

CONCLUSIONS

Novel, 2-dimensional MEMS-controllable microlens array has been integrated with a Vertical-Cavity-Surface-Emitting-Laser (VCSEL) array. Flip-chip transfer of MEMS structure on transparent substrate is described in detail. By translating polymer microlens fabricated on MEMS X-Y movable plate using electro-thermal actuators, the 2-dimensional beam steering is demonstrated

REFERENCES

1. H. Toshiyoshi, J. G. Su, J. Lacosse, and M. C. Wu, "Micromechanical lens scanners for fiber optics switches", *Proc. MOEMS'99*, pp. 165-170 (1999).

2. R. Goring, D. Doring, P. Bucker, B. Gotz, P. Chreiber, P. Dannberg, E. B. Kley, and M. Cumme, "Microoptical concepts for miniaturised scanners and switches from design to realisation", *Proc. MOEMS'99*, pp. 76-87 (1999).

3. D. A. Koester, R. Mahadevan, A. Shishkoff, and K. W. Markus, "Multi-User MEMS Processes (MUMPs): Design handbook, Rev. 4," *Cronos Integrated Microsystems*, 3021 Cornwallis Road, Research Triangle Park, NC 27709 (1999).

4. A. Tuantranont, V. M. Bright, W. Zhang, J. Zhang, and Y. C. Lee, "Self-aligned assembly of microlens arrays with micromirrors," *SPIE* Vol. 3878, pp. 90-101 (1999).

5. J. H. Comtois and V. M. Bright, "Surface micromachined polysilicon thermal actuator arrays and applications", *Technical Digest of the 1996 Solid-State Sensor and Actuator Workshop*, Hilton Head Isl., SC, pp. 174-177.

6. W. Zhang, K. F. Harsh, M. A. Michalicek, V. M. Bright, and Y. C. Lee, "Flip-chip assembly for RF and optical MEMS", *Proc. ASME InterPack'99*, Vol. EEP-26-1, pp. 349-354 (1999).

7. H. J. Zhou, J. Neff, Y. Chen, V. Fedor, Y. C. Lee, C. C. Mou, W. Berseth, T. Mclaren, and E. Tang, "Demonstration of a massively parallel bi-directional crosspoint switch with optical control", *SPIE* Vol. 3005, pp. 266-272 (1997).

THE IMPACT OF MEMS ON THE CHEMICAL AND PHARMACEUTICAL INDUSTRIES

Klavs F. Jensen
Departments Chemical Engineering and Materials Science and Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139

ABSTRACT

Microfabrication techniques are increasingly used in different fields of chemistry to realize structures with capabilities exceeding those of conventional macroscopic systems. In addition to already demonstrated analysis applications (μTAS), MEMS microreactor systems are expected to have a number of advantages for chemical synthesis, chemical kinetics studies, and process development. Current efforts in microfabrication and application of microchemical systems are reviewed. Chemical processing advantages from increased heat and mass transfer in small dimensions are demonstrated with model gas, liquid, and multiphase reaction systems. These case studies also serve to illustrate benefits of integrating sensors for flow, temperature, and chemical composition with microfluidic reaction and control components. The role of process simulations in the development of microreaction technology is included along with approaches to scale up via replication of microchemical devices.

INTRODUCTION

Microreactors are sub-millimeter-scale, integrated chemical systems built by microfabrication methods generally used for MEMS [1], LIGA [2, 3], mechanical micromachining [2, 3], and soft lithography methods [4, 5]. Microchemical systems have feature sizes in the micron to hundreds of micron range and may be integrated with sensors and actuators. The reduction in size and integration of multiple functions creates structures with capabilities that exceed those of conventional macroscopic systems, adding new functionality while potentially making low cost mass production possible.

Miniaturization of chemical analytic devices in "micro-total-analysis-systems" (μTAS)[6] represents a natural extension of MEMS technology to chemistry and biology with obvious applications in combinatorial chemistry, high throughput screening, and portable analytical measurement devices. The merging of μTAS techniques with microreaction technology promises to yield a wide range of novel devices for high throughput screening, as well as on-line monitoring of production systems.

Microreaction technology is also expected to have a number of advantages for chemical production [3, 7]. The high heat and mass transfer rates possible in microfluidic systems allow reactions to be performed under more aggressive conditions with higher yields than can be achieved with conventional reactors. More importantly, new reaction pathways deemed too difficult in conventional microscopic equipment, e.g., direct fluorination of aromatic compounds [8], could be pursued. Even if a microreactor fails, the small quantity of chemicals released accidentally could be easily contained. Moreover, the presence of integrated sensor and control units could allow the failed reactor to be isolated and replaced while other parallel units continued production. These inherent safety characteristics suggest that production scale systems of multiple microreactors should enable distributed point-of-use synthesis of chemicals with storage and shipping limitations, such as highly reactive and toxic intermediates (e.g., cyanides, peroxides, azides). As a demonstration of these concepts, DuPont has synthesized a number of potentially hazardous chemicals, including isocyanates, in a microreactor formed by bonding silicon wafers patterned to form channels, preheaters, and catalytic reactor sections [7].

Scale-up to production by replication of microreactor units used in the laboratory would eliminate costly redesign and pilot plant experiments, thereby shortening the development time from laboratory to commercial production. This approach would be particularly advantageous for the fine chemical and pharmaceutical industries, where production amounts are often small–less than a few metric tons per year. The strategy would also allow for scheduled, gradual investment in new chemical production facilities without committing to a large production facility from the outset.

Microchemical systems for combinatorial synthesis and screening of small molecules and systems for nucleic acid synthesis and detection have already revolutionized drug discovery in pharmaceutical companies. Rapid screening of catalysts and product synthesis pathways could lead to analogous productivity increases in chemical industry laboratories. Experimentation at the conventional bench scale is limited by high costs of reagents and safety concerns that the small volumes and inherent safety characteristics of microreactors could effectively eliminate. The introduction of new chemicals is also limited by the risk and high capital costs of scaling from laboratory to production plant. The ability to scale-up by replication allows a gradual scale up with input from customers and no reactor redesign.

Chemical laboratories already use small reactors for testing process chemistries; for example, catalyst testing often uses small tubular reactors formed by filling powdered catalysts into glass tubes. However, the overall procedure is faced with bench top analytical equipment, large panels of complex fluid handling manifolds, and containment vessels. Chemical detection is the rate-limiting step in most methods since detailed product information must be obtained using sequential screening. With the continual advances in μTAS and microfabrication techniques, these macroscopic test systems could eventually be replaced by PC-card sized microchemical systems consisting of integrated microfluidic, sensor, control, and reaction components. Such systems would clearly require less space, utilities, and produce less waste. They would enable high-throughput screening of process chemistries under controlled conditions, which is often difficult in conventional macroscopic systems. Moreover, the small dimensions imply laminar flow, making it feasible to fully characterize heat and mass transfer and extract chemical kinetic parameters from sensor data. The high heat and mass transfer rates possible in microfluidic systems also allow reactions to be performed under more uniform temperature conditions. By using the control of fluid contacting patterns afforded by MEMS

fabrication techniques, it should become possible to design reactors that achieve specific research and development objectives.

MICROREACTOR DESIGN AND FABRICATION

Current microreactor designs depend on the target applications, which can be broadly classified in order of increasing complexity as homogeneous gas-phase reactions, heterogeneous catalytic gas-phase reactions, liquid-liquid reactions (both miscible and immiscible systems), heterogeneous liquid phase reactions, and multiphase (gas-liquid-solid) reactions. Fluid flow combined with mass and heat transfer (transport phenomena) are easiest to handle in the case of two gases reacting without the solid catalyst, but such systems are rare. Most chemical systems of commercial interest require the presence of a catalyst–either as a thin film on a solid surface or dispersed on a high surface area porous support. Mixing is often a critical issue in liquid phase reactions and the laminar flows characteristics of microreactors imply that the turbulent mixing typically realized in macroscopic equipment has to be replaced by effective mixing by diffusion across thin liquid layers. The contact pattern for reaction components becomes especially important in the case of multiphase reactions, such as the hydrogenation of an unsaturated hydrocarbon over a solid catalyst. Efficient mixing and access to catalytic sites can lead to orders of magnitude enhancement in performance, which suggests that the precise control offered by microfabrication could have particular opportunities in multiphase reactions. Besides this application, microreaction technology is most likely to be successful for fast reactions, systems requiring precise control of reagent contacting, and for novel reaction pathways that are difficult to realize on a macroscopic scale.

Current microreactor fabrication methods fall into four main categories:

    (i)   MEMS based,
    (ii) LIGA derived,
    (iii) patterned laminated structures, and
    (iv) soft lithography based polymer devices.

MEMS based devices are based on classical bulk machining and deep reactive ion etching (RIE) + bonding approaches. Recent work at MIT serves as examples in the following sections. Forschnungszentrum Karlsruhe and the Institute for Micro technology Mains (IMM) have been the primary drivers of LIGA derived microreactors, micromixers, and micro-heat-exchangers. These devices typically consist of multiple layers of multichannels fabricated in metals. Coating of the interior surface of these channels with a catalytic active material turns mixers and heat exchangers into reactors. The devices have been demonstrated for a number of model gas-surface catalytic reactions, liquid phase, and gas-liquid reactions, as well as selected commercial demonstrations. The latter includes a partial oxidation reaction of a vitamin intermediate, an addition reaction for fine chemicals production, and polymer synthesis. Similar devices have been fabricated using microelectro discharge machining (μEDM) techniques. Reactors have also been fabricated in polymers by microinjection molding, and embossing techniques.

Pacific Northwest National Laboratories (PNNL) has pioneered microfabrication by lamination of multichannel sheets. Each sheet of metal, ceramic, or plastic is patterned by a variety of techniques including photolithography, □EDM, etching (ion or chemical), and laser ablation. The sheets are then laminated to form the final microchannel reactor. Devices have been demonstrated for energy related applications, specifically fuel evaporation and reforming [9, 10].

Lamination of patterned ceramic tapes can also be used to create microreactor components. Laminating multiple layers machined with desired patterns again makes three-dimensional device structures. In the "green state," ceramic tapes are soft, pliable, and easily machinable. A well-developed thick film technology exists for deposition of various metals and electrical components on the tapes in the pre-fired state and the formation of three-dimensional interconnects.

"Soft lithography" refers to a collection of techniques pioneered by Whitesides and his coworkers [4, 5]. The common feature in these methods is the use of a pattern transfer element that is a transparent elastomer, usually poly(dimethylsiloxane) PDMS, having a pattern embossed on its surface. PDMS is suitable for aqueous systems, but it swells in many organic solvents and has limited temperature stability. Therefore, PDMS based microreaction components have seen limited biological applications, including fluid mixers and electrophoresis. However, when combined with electroplating of metals and molding of ceramics, the rapid prototyping advantages of soft lithography have considerable applications in microreaction technology. Examples of this approach are included in the following discussion of MEMS based microreactors.

GAS-PHASE REACTORS

A common microreactor design used for gas-phase heterogeneous reactions is a microchannel device that can be integrated with a heat exchange layer for highly exothermic reactions [9-11]. These devices have the advantage of high productivity per unit volume, but they suffer from lack of sensing and active control within the microchannel assembly–similar to conventional ceramic monolith reactors. Thin walled microreactors offer the opportunity for the integration of heaters, as well as flow and temperature sensors on the external side of the reaction region (see Figure 1). The micron-thick wall provides good thermal contact with the catalyst in the interior and the low thermal mass of the thin wall has the further advantage of fast (~10 ms) thermal response times.

The catalyst can be deposited on the interior wall by a variety of techniques including thin film deposition by physical or chemical vapor deposition and liquid preparation techniques. Thin film deposition techniques of the catalyst are typically limited to metals and simple oxides. Moreover, the resulting film has a low surface area and must be roughened by thermal activation. Wet preparation techniques allow the creation of catalyst systems similar to those used in conventional reactors, but surface tension effects cause the catalyst solution to collect in channel corners. Aerosol sprays or ink jet printing allows placement of the catalyst on a desired region of the membrane.

Oxidation and partial oxidation reactions of a number of compounds, including hydrogen, ammonia, and ethane, have been successfully carried out in the flammable regime [13]. The ability to control the catalyst temperature, and therefore the selectivity, of partial oxidation reactions is a critical factor determining reactor performance. During an exothermic reaction, the heat is supplied to the catalyst wall by both the heaters and the reaction. The primary heat loss mechanism is through conduction in the thin top wall to the reactor bulk silicon side walls. Therefore, by varying the thickness and thermal conductivity of the top wall (by using SOI wafers with varying silicon thickness), the ability of the reactor to dissipate the heat of the reaction can be controlled (see Figures 2a and b). This feature provides access to a wider temperature window of operating conditions and milder oxidation conditions than can be achieved in conventional systems.
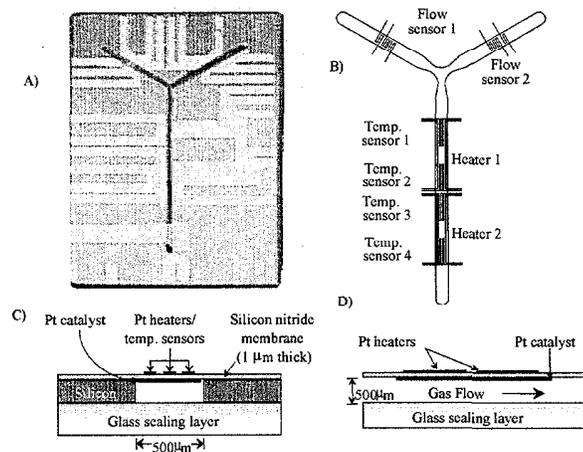
Figure 1 *Gas-phase microreactor. (A) Photograph of device. (B) Top view schematic. (C) Side view across channel. (D) Side view along flow direction [12].*
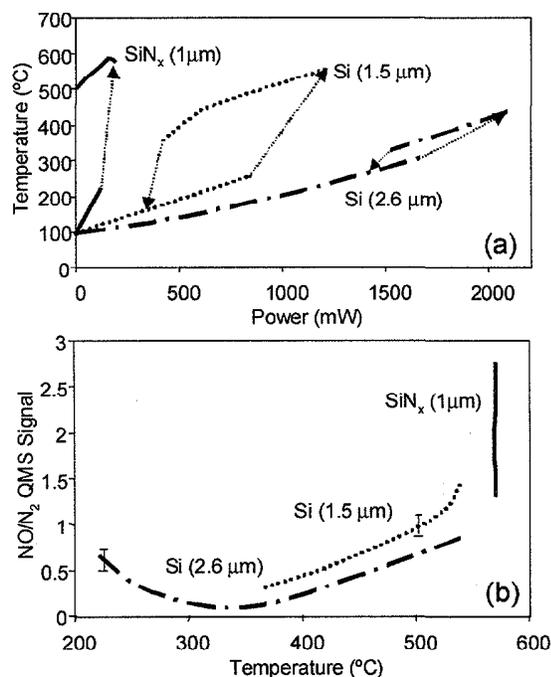
Figure 2 *(a) Ignition/extinction behavior with varying input power for ammonia oxidation over platinum in microreactors with different wall materials and thickness. (b) Corresponding variation in selectivity towards nitrogen and nitrogen monoxide.*

The use of a permeable membrane allows the integration of separation with chemical reaction, as in macroscopic membrane reactors [14]. For example, the integration of a submicron-thick palladium membrane makes a highly efficient hydrogen purification device and provides the potential for conducting hydrogenation and dehydrogenation reactions [15]. This approach also allows the use of other membrane materials for highly selective gas separation. Alternatively, the holes in the support structure can be used to disperse gas into a liquid stream or can act as a particle filter.

LIQIUD-PHASE REACTORS

Mixing is a critical issue in the design of liquid phase microreactors. The small dimensions in microfluidic devices imply small Reynolds numbers and laminar flow so that mixing occurs by diffusion. This characteristic becomes both a challenge and an advantage for liquid phase reaction systems. The slow mixing of co-flowing streams offers additional opportunities for phase transfer reactions and separation devices [16, 17], and it can be exploited in novel fabrication schemes [18].

Mixing is generally accomplished by the repeated lamination of the two streams to be mixed in order to increase the contact area and reduce diffusion lengths [19]. Ultimately, the design of a micromixing unit is a trade-off between mixing speed, pressure drop, volume flow, feasibility of microfabrication, and integration with chemical detection devices.
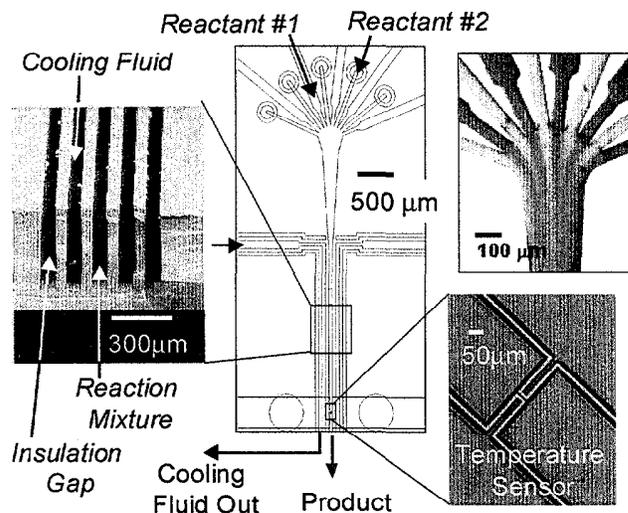
Figure 3 *Liquid phase reactor with lamination of fluid streams and integrated with heat exchangers and temperature sensors in downstream reaction zone. Upper right hand insert shows mixing of acid base mixture in gray tone image of indicator color.*

Figure 3 shows an example of a microfabricated liquid phase reactor that integrates laminar mixing, hydrodynamic focusing, rapid heat transfer, and temperature sensing. The integrated reactor was fabricated using SIO wafers. DRIE [20] was used to form channels 50-400 μm wide and ~500 μm deep. An anodically bonded Pyrex wafer capped the wafer. Model studies with acid-base reactions show that the reactor achieves complete mixing in ~10 ms depending on the fluid properties. The relatively high transfer coefficient [1500 W/(m^2·C°)] realized in the microfabricated heat exchanger provided excellent cooling of even highly exothermic reactions such as hydrolysis of propionyl acid chloride.

MULTIPHASE REACTORS

Gas-liquid-solid reactions are ubiquitous throughout the chemical industry and provide unique opportunities for microfabrication. The high surface-to-volume ratios attainable in microfabricated structures, leading to improved thermal management and fast mass transfer, suggest that microfabricated multiphase systems could have performance advantages relative to conventional macroscopic systems. We have demonstrated these

characteristics for single and multichannel, multiphase packed bed reactors (Figure 4), again fabricated using deep RIE technology and bonding techniques [21, 22]. In addition to the microfluidics, thin-film resistive heaters and temperature sensors have been integrated with the device. Each reaction chamber holds porous catalyst particles in place by a microfabricated filter as reactants are fed continuously in a co-current fashion. The gas and liquid reactant streams are brought into contact by a series of interleaved, high-aspect ratio inlet channels (25 μm wide × 300 μm deep) designed to increase gas-liquid mass transfer. The catalyst restrainer is formed within the reaction channel by etching a series of posts 40 μm wide separated by 25 μm gaps.
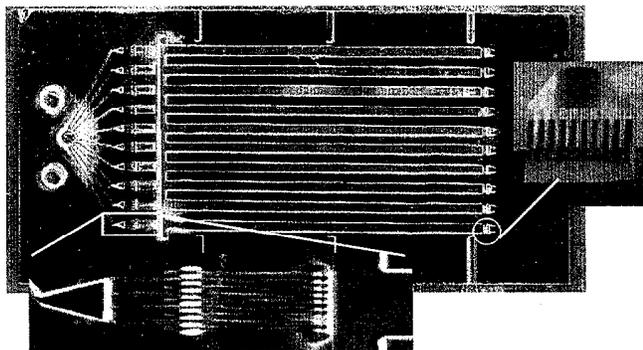
Figure 4 *Multichannel multiphase packed bed reactor* [22].

Model reaction studies (including hydrogenation of cyclohexene and oxidation of benzyl alcohol) over catalyst particles have been used to characterize the mass transfer and reaction behavior of the multiphase microreactor. As in the case of macroscopic reactors, a bubble flow regime is observed at low gas flow rates; pulsing flow occurs at high gas flow rates, and a steady co-current flow is found at intermediate conditions. These results demonstrate that basic multiphase flow concepts extend to microfluidic systems.

Microfabrication also provides unique opportunities for creating catalyst support systems that avoid packing variations associated with distribution of catalyst particle sizes. For example, posts can be formed by DRIE and rendered catalytic by wash coats and surface chemistry treatments. Similar structures have been developed for chromatography and concentration of DNA [23].

Selected multiphase systems have been investigated. Oxidation reactions, exemplified by the oxidation of benzaldehyde, have been demonstrated in a single-channel reactor system. In this case, the inherent safety of a small reaction volume (4 μl) allows the reaction of pure oxygen and organic solvents at elevated temperatures. The hydrogenation of cyclohexene has been used to characterize the mass transfer coefficient of the multiple channel reactor using traditional reaction engineering analysis. Values of the mass transfer coefficient for the multiple channel microreactor have been determined to be two orders of magnitude larger than those reported for typical macroscopic reactors. Therefore, for reactions that operate in a mass-transfer limited regime, microreaction devices could be considerably smaller in volume and still maintain high productivity.

INTEGRATION OF CHEMICAL SENSORS

The determination of chemical products and their amounts is often a time-consuming challenge in chemical research and development. In order for microreaction technology to be successful it must include chemical sensors. Having a small reactor interfaced to large bench top analytical equipment often implies that external, fluid chamber volume is much larger than the volume of the microreactor. This discrepancy raises the potential that the observed reactor volume is the reactor-analytical system interface, instead of the microreactor. The problem is avoided only if (i) the reaction reaches full conversion in the microreactor, (ii) the reactants and products are separated in the microreactor, or (iii) the reaction is quenched in the microreactor.

In the case of liquid phase reactions that can be monitored by color changes or variations in the UV spectrum, the reactor can readily be interfaced to an optical fiber light source and detector to enable on-line monitoring. However, UV-visible spectroscopy has limited chemical resolution, *i.e.*, features are broad and common for many compounds.

Infrared (IR) detection is a broadly applicable detection method that can be integrated with liquid phase microreactors by using the IR transparency of silicon. Thickness, doping of the wafer, as well as reflections from the substrate surface, adversely affect transmission. Although the 40% transmission obtained using uncoated samples is sufficient for most applications, one can apply an anti-reflective (AR) coating to increase silicon transmission [24].

The short path lengths and optical densities usually imply that IR absorption spectroscopy is not practicable in gas-phase microreactors. As an alternative optical method for gas-phase detection, photoacoustic spectroscopy (PAS) offers several advantages. In photoacoustic spectroscopy, incident light is modulated at an acoustic frequency. If the optical wavelength couples to an energy transition in the gas, the gas absorbs the light resulting in a periodic gas expansion. PAS applies to many chemical compounds and its sensitivity scales inversely with dimensions. The recent detection of propane in a microfabricated photoacoustic cell illustrates the potential of the spectroscopy for technique in microscale chemical analysis [25].

COMBINING MEMS AND SOFT LITHOGRAPHY

In many cases, silicon functions well as the substrate for a microchemical reactor, but different materials (*e.g.* polymers) that have desirable material properties may be preferable for some application [26]. The photosensitive epoxy SU-8 [27] offers opportunities for producing hybrid microchemical devices in silicon and SU-8, quartz and SU-8, and entirely from SU-8. In all cases silicon, quartz, or pyrex function as the substrate on which the device is formed. In some cases, the device is used as is, and in other cases it is released from the substrate. SU-8 can be used to create sealed microfluidic channels, to bond materials and to planarize structures. The ability to planarize allows patterned metal structures (for electrochemical detection/synthesis or for heating) to be incorporated into reaction chambers. Figure 5 shows an electrochemical reactor in which the electrodes and fluidic interconnects were formed on a silicon substrate and the fluidic channels were formed using an SU-8 bonding approach. The flexibility creates additional opportunities for liquid-phase microreactors. At low temperatures, for example, it becomes feasible to fabricate composite quartz/SU-8 devices that allow for *in situ* UV detection through a quartz window.

SIMULATION OF MICROREACTOR SYSTEMS

An in-depth understanding of the operating characteristics of microreactors is necessary to evaluate the benefits and disadvantages associated with new microreactor designs. Several

commercial software tools are available to quantify different aspects of microfabricated devices, but they are not universally applicable [28, 29]. Simulations serve not only as a design tool, but also as a means to interpret experimental data [30]. Models have been used to evaluate design changes such as changing the top wall material and thickness, as described above, heater design, [31], and flow meter design [12] in the gas-phase reactor. Computational fluid dynamic software also provides insight into liquid-phase mixing. In particular, the use of simulation allows inclusion of three-dimensional wall effects often neglected in liquid layer mixers. If all physical parameters are known, simulations should accurately reflect experimental observations since the basic transport equations for laminar flow are well established. Thus, the computational approach to new designs can avoids costly, iterative experimental design process where components are fabricated, tested, and then redesigned to improve performance.
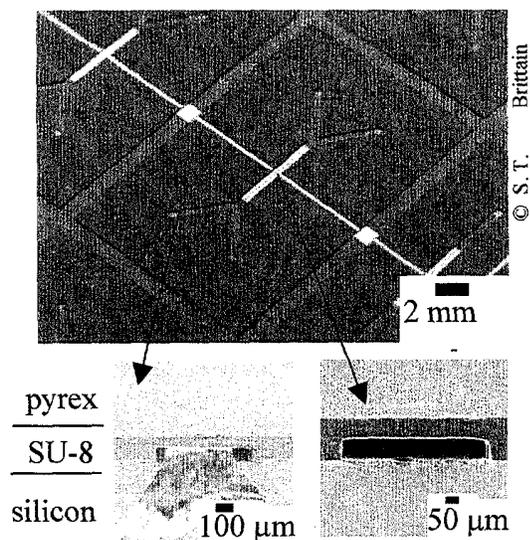
Figure 5. Photograph of electrochemical microreactor and SEM images showing cross-sections through the devices.

SCALE-UP

In the past, the microreaction community has focused on the design of individual microreactors. No significant efforts have been made to build parallel arrays of reactors as originally envisioned by early papers in the field. The concept of scale-up by replication of many units (scale-out) at first glance may appear to be simple, but the strategy presents new challenges that have not been addressed in previous chemical system scale-up. Particularly, the areas of reactor monitoring and control become increasingly complex as the parallel array size grows to a large number of reactors.

To address issues in microreactor scale-up, MIT and DuPont have begun construction of a multiple reactor test station for gas-phase systems [32]. The fabrication of this system serves as a test case for the development of more complex microreactor based systems. The approach addresses the difficulties in developing electrical and fluidic interfaces to microreactors that must be interchanged frequently, have heated outlet lines, and require millisecond real-time control algorithms. The multiple microreactor test station (Figure 6) not only includes microfabricated reactors, but also integrates other MEMS components for fluidic control. This system contains all the

components of a conventional catalyst test station (feed gas manifold, reactor feed manifold, reactors, and control circuitry), but with a dramatically smaller footprint. It represents the first demonstration of a parallel microreactor array that integrates reactors with fluidic distribution and control components in a scalable fashion.

CONCLUSIONS

The examples represent shown here a small fraction of the many designs for microreactors being pursued or envisioned by different research groups. In developing microreaction technology, it will be essential to focus on systems where microfabrication can provide unique process advantages. Such advantages could be derived from increased mass and heat transfer, leading to improved yield and safety for an existing process. The real value of the miniaturization effort, however, would be in exploring new reaction pathways and finding economical and environmentally benign solutions to chemical manufacturing. It will be important to exploit characteristics resulting from the small dimensions beyond the high transport rates, specifically forces associated with high surface area-to-volume ratios. In order for microreactors to move beyond the laboratory into chemical production, they must also be integrated with sensors and actuators, either on the same chip, or through hybrid integration schemes. The integration of chemical systems with sensors in µTAS is already rapidly expanding the field, and cross-fertilization with microreactors for chemical synthesis will ultimately result in integrated chemical processors. The packaging of multiple reactors presents significant challenges in fluid handling, local reactor monitoring, and control. Thus, the realization of microreaction technology will require multidisciplinary research efforts in collaboration with users of the technology.
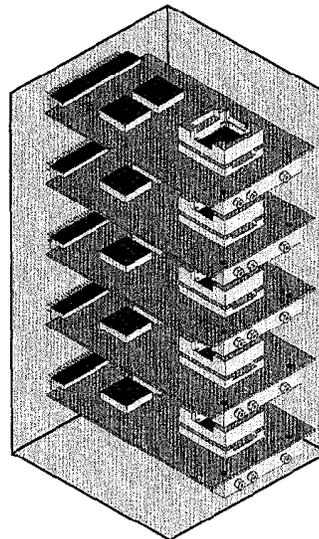
Figure 6. Conceptual drawing of packaged, parallel microreactor system [32].

ACKNOWLEDGEMENTS

REFERENCES

1. K. D. Wise, "Special issue on Integrated Sensors, Microactuators, and Microsystems (MEMS)", *Proceedings of the IEEE*, *86*, 1531 (1998).
2. W. Ehrfeld, C. Grärtner, K. Golbig, V. Hessel, R. Konrad, H. Löwe, T. Richter, and T. Schulz, "Fabrication Of Components and Systems for Chemical and Biological Microreactors", in *Microreaction Technology*, W. Ehrfeld, Editor Springer, Berlin (1998), 72-90.
3. W. Ehrfeld, V. Hessel, and H. Lehr, "Microreactors for Chemical Synthesis and Biotechnology - Current Developments and Future Applications", in *Microsystem Technology in Chemistry and Life Science* (1998), 233-252.
4. Y. N. Xia and G. M. Whitesides, "Soft Lithography", *Ann.Rev. Mat. Sci*, *28*, 153 (1998).
5. Y. N. Xia and G. M. Whitesides, "Soft Lithography", *Ang. Chem-Int. Ed. 37*, 551 (1998).
6. A. van den Berg and D. J. Harrison, "Micro Total Analysis Systems'98", *Micro Total Analysis Systems µTAS '98*, eds. Kluwer Academic Publishers, Dordrecht (1998).
7. J. J. Lerou, *et al.*, "Microfabricated Minichemical Systems: Technical Feasibility", in *Microsystem Technology for Chemical and Biological Microreactors: Papers of the Workshop on Microsystem Technology, Mainz, 20-21 February, 1995*, DECHEMA, Frankfurt (1996), 51-69.
8. R. D. Chambers and R. C. H. Spink, "Microreactors for Elemental Fluorine", *Chem. Commun.*, 883 (1999).
9. A. Y. Tonkovich, J. L. Zilka, M. J. LaMont, Y. Wang, and R. S. Wegeng, "Microchannel Reactors for Fuel Processing Applications. I. Water Gas Shift Reactor", *Chem. Eng. Sci.*, *54*, 2947 (1999).
10. T. A. Ameel, R. O. Warrington, R. S. Wegeng, and M. K. Drost, "Miniaturization Technologies Applied to Energy Systems", *Energy Conv. Management*, *38*, 969 (1997).
11. G. Wießmeier and D. Hönicke, "Microfabricated Components for Heterogeneously Catalysed Reactions", *J. Micromech. Microeng.*, *6*, 285-289 (1996).
12. D. J. Quiram, K. F. Jensen, M. A. Schmidt, P. L. Mills, J. F. Ryley, and M. D. Wetzel. "Integrated Microchemical Systems: Opportunities for Process Design" in *Foundations of Computer Aided Process Design*. 1999. Colorado, CACHE, Colorado (1999),
13. R. Srinivasan, I.-M. Hsing, P. E. Berger, K. F. Jensen, S. L. Firebaugh, M. A. Schmidt, M. P. Harold, J. J. Lerou, and J. F. Ryley, "Micromachined Reactors for Catalytic Partial Oxidation Reactions", *AIChE Journal*, *43*, 3059 (1997).
14. R. Govind and N. Itoh, "Membrane Reactor Technology", *AIChE Symposium Series*, 268, AIChE, New York, (1989).
15. A. Franz, K. F. Jensen, and M. A. Schmidt. "Palladium Based Micromembranes for Hydrogen Separation and Hydrogenation/Dehydrogenation Reactions" in *12th International Conference on MicroElectro-Mechanical Systems*. Orlando, Florida, IEEE (1999), 382-385.
16. J. P. Brody, P. Yager, R. E. Goldstein, and R. H. Austin, "Biotechnology at Low Reynolds Numbers", *Biophysical Journal*, *71*, 3430 (1996).
17. J. R. Burns and C. Ramshaw, "Development of a Microreactor for Chemical Production", *Trans IChemE*, *77*, 206 (1999).
18. P. J. A. Kenis, R. F. Ismagilov, and G. M. Whitesides, "Microfabrication Inside Capillaries Using Multiphase Laminar Flow Patterning", *Science*, *285*, 83 (1999).

19. W. Ehrfeld, K. Golbig, V. Hessel, H. Lowe, and T. Richter, "Characterization Of Mixing In Micromixers By a Test Reaction: Single Mixing Units and Mixer Arrays", *Ind. Eng. Chem. Research, 38,* 1075 (1999).
20. A. A. Ayon, R. Braff, C. C. Lin, H. H. Sawin, and M. A. Schmidt, "Characterization of a Time Multiplexed Inductively Coupled Plasma Etcher", *J.Electrochem. Soc.*, *146*, 339 (1999).
21. M. W. Losey, M. A. Schmidt, and K. F. Jensen. "A Micro Packed-Bed Reactor for Chemical Synthesis", in *3rd International Conference on Microreaction Technology*. 1999. Frankfurt, Springer, Berlin (1999).
22. M. W. Losey, S. Isogai, M. A. Schmidt, and K. F. Jensen. "Microfabricated Devices for Multiphase Catalytic Processes", in *4h International Conference on Microreaction Technology*. 2000. Atlanta, GA, AIChE, New York (2000), 416-422.
23. L. A. Christel, K. Petersen, W. McMillan, and M. A. Nortrup. "Nucleic Acid Concentration and PCR for Diagnostic Applications", in *µTAS '98 Workshop*. 1998. Banff, Canada, Kluwer Academic Publishers, Dordrecht (1998), 277-280.
24. T. Floyd, M. A. Schmidt, and K. F. Jensen. "Towards Integration of Chemical Detection for Liquid Phase Microchannel Reactors", in *4th International Conference on Microreaction Technology*. 2000. Atlanta, GA, AIChE, New York (2000), 416-422.
25. S. L. Firebaugh, K. F. Jensen, and M. A. Schmidt. "Miniaturization And Integration of Photoacoustic Detection With a Microfabricated Chemical Reactor System", in *µTAS 2000 Workshop*. 2000. Twente, The Netherlands, Kluwer Academic Publishers, Dordrecht, The Netherlands (2000).
26. J. J. Jackmann, T. Floyd, M. A. Schmidt, and K. F. Jensen. "Development of Methods for On-Line Chemical Detection With Liquid-Phase Microchemical Reactors Using Conventional and Unconventional Techniques" in *µTAS 2000 Workshop*. 2000. Twente, The Netherlands, Kluwer Academic Publishers, Dordrecht, The Netherlands (2000).
27. M. O. Heuschkel, L. Guerin, B. Buisson, D. Bertrand, and P. Renaud, "Buried Microchannels In Photopolymer For Delivering of Solutions To Neurons In a Network," *Sensors and Actuators B-Chemical*, *48*, 356 (1998).
28. *MEMCAD*, Microcosm Technologies, Inc., 5511 Capital Center Dr., Suite 104, Raleigh, NC 27606.
29. *CFD-ACE*, CFD Research Corporation, 215 Wynn Dr., Huntsville, AL 35805.
30. I.-M. Hsing, R. Srinivasan, M. P. Harold, K. F. Jensen, and M. A. Schmidt, "Simulation of Micromachined Chemical Reactors for Heterogeneous Partial Oxidation Reactions," *Chem. Eng. Sci.*, *55*, 1 (2000).
31. D. J. Quiram, I.-M. Hsing, A. J. Franz, R. Srinivasan, K. F. Jensen, and M. A. Schmidt. "Characterization of Microchemical Systems Using Simulations", in *2nd International Conference on Microreaction Technology*. New Orleans, LA, AIChE, New York (1998) 205-10.
32. D. J. Quiram, J. F. R. Patrick L. Mills , M. D. Wetzel, J. M. Ashmead, T. M. Delaney, D. J. Kraus, J. S. McCracken, K. F. Jensen, and M. A. Schmidt. "Package Level Integration of Silicon Microfabricated Reactors to Form a Miniature Reactor Test System", in *Technical Digest of the 2000 Solid-State Sensor and Actuator Workshop*. 2000. Hilton Head Isl., SC, Transducer Research Foundation (2000).

Integrated Elastomer Fluidic Lab-on-a-chip – Surface Patterning and DNA Diagnostics

Hou-Pu Chou[*], Marc A. Unger[+], Axel Scherer[*+], and Stephen R. Quake[+]
Departments of Electrical Engineering[*] and Applied Physics[+], Caltech,
Pasadena, CA 91125, USA

ABSTRACT

We recently developed a method of multilayer fabrication for elastomeric devices, which we used to fabricate monolithic active valves and pumps. Here we describe efforts to use these pumps and valves in an integrated DNA diagnostic chip and show results of a key component, surface patterning, with two different kind of surface chemistries by using similar elastomeric channel devices. Flow control, reagent metering, in-line mixing and loop circulations are also demonstrated.

INTRODUCTION

Several techniques are currently being developed towards the goal of an integrated fluidic lab-on-a-chip[1-6]. Among these, monolithic microvalves and micropumps made from silicone elastomer[5] have great potential because of their simplicity, robustness, easy fabrication and low cost. Here, we describe some new results and extensions to our previous work as part of our ultimate goal of creating an integrated DNA diagnostic chip.

We are interested in using microfabricated chips to measure gene expression and detect the presence of pathogenic DNA. DNA expression arrays have proved to be a useful tool in studying gene expression in a variety of organisms, including yeast, worm, mouse and human[7]. The sensitivity of such arrays is limited in part by the diffusion of target DNA to the probes that are anchored on the surface[8]. A better approach is to use microfluidic devices in order to pump solutions of target DNA over a set of anchored probes in order to ensure that all of the target DNA is exposed to each of the probes. This would provide increased sensitivity as well as decreasing the amount of time needed for hybridization. Chips with high sensitivity would also be useful for measuring single cell gene expression. This higher sensitivity may eliminate the need for PCR in many cases of pathogen detection and therefore make it possible to do multiple disease diagnosis with one integrated lab-on-a-chip.

Making such chips requires a number of important technological advances in the current state of the art of microfluidics. First, one needs to be able to fabricate microfluidic devices in a way that is compatible with the delicate surface chemistry required to anchor or synthesize DNA probes on a chip. Second, one must be able to effect the desired patterning or surface chemistry. Third, one must be able to manipulate small amounts of material and perform the necessary biochemical reactions on chip. Finally, one needs to be able to pump the targets over the probes.

The first two issues we have addressed by using "soft lithography", which in part refers to the notion of fabricating channel systems out of soft elastomeric materials using replication molding techniques[9]. We have found that elastomeric devices provide a number of important advantages over conventional micromachining, such as ease of fabrication, room temperature sealing of devices to glass substrates, good optical properties, and low materials cost, especially compared to single crystal silicon:

~$.05/cm³ vs. ~$2.50/cm³. Microfluidic networks fabricated in such a manner can easily be sealed to substrates with delicate surface chemistry. Another aspect of soft lithography we have taken advantage of is the ability to chemically pattern surfaces using fluid flow[10,11]. The third and final issues we have addressed by developing a multilayer fabrication process in silicone elastomer. This multilayer process allows easy fabrication of devices with moving parts, including microfluidic valves and pumps[5].

Here we describe results in which we chemically pattern surfaces with biotin/avidin and DNA using fluidic networks in a way that is compatible with further fluidic processing. We also show how pumps can be incorporated into an integrated device to both meter reagents and pump fluid in a closed loop. Rotary flow in a closed loop cannot be achieved by electrophoresis or electro-osmotic flow used in most common lab-on-a-chips because of the existence of two electric polarities.

DEVICE DESIGN AND FABRICATION

Multilayer soft lithography is a process recently developed in our lab in order to make 3-D monolithic elastomer devices with a combination of air and fluid channels[5]. When an air channel passes above another fluid channel, the thin membrane between these two channels becomes a valve. By applying air pressure in the air channel, the membrane collapses and stops the fluid flow. Releasing the pressure then re-opens this valve. Three valves in series become a peristaltic pump when an appropriate on/off air pressures are applied in a sequence. For example, 100, 110, 010, 011, 001, 101 pumps water to the right, where 1 and 0 represent closed or open valves respectively. A schematic diagram is shown in Figure 1. A brief description of the fabrication process follows.
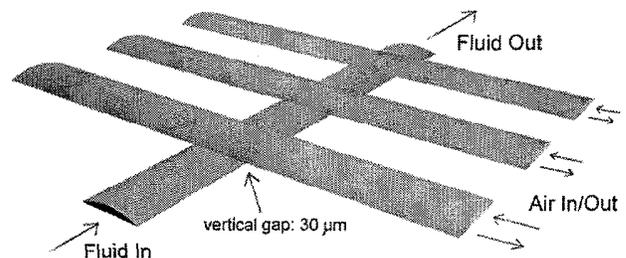
Figure 1. Schematic diagram of a peristaltic pump with three valves on top of a fluid channel.

Air and fluid mother molds were fabricated on silicon wafers by photolithography. Photoresist (Shipley SJR5740) was spun onto the silicon substrate at spin rates corresponding to the desired channel heights. After photolithography, intrusive channels made of photoresist were formed. Fluid channel molds were baked on a hot plate of 200°C for 30 minutes so that the photoresist could reflow and form a rounded shape, which is important for complete valve closure[5]. A one minute trimethylchlorosilane (TMCS)
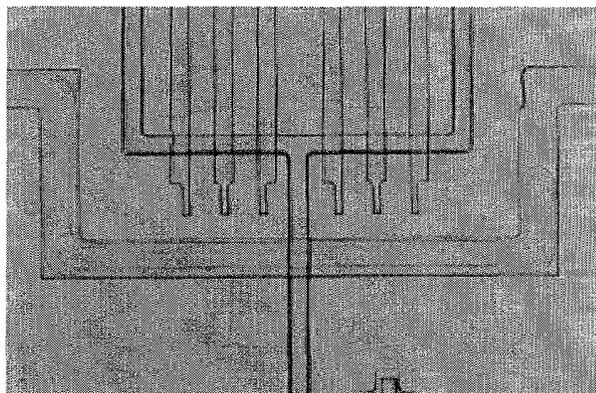
(a).
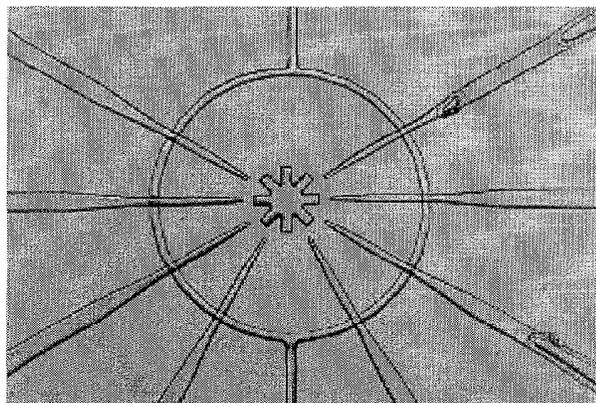
(b).

Figure 2. Snapshots of a two-layer DNA diagnosis chip. (a) shows the input mixing T-shaped fluid channel and six corresponding control microvalves. The wide 100-μm air channel beneath is used to close the inlet when the peristaltic pump at the loop starts operating for hybridization. (b) shows the center fluid loop for DNA hybridization. Any three of the finger air channels form a peristaltic pump. Fluid channel width: 50μm.

vapor treatment was applied to these molds before each RTV replication process to prevent adhesion of cured RTV to the photoresist. With this protective coating, molds can be reused many times.

30:1 GE-RTV 615A:615B was spun on a fluid channel mold at 2,000 RPM, which covers the photoresist channel and leave a thin membrane on top of it. At the same time, 3:1 GE-RTV 615A:615B was poured onto an air channel mold. After baking both in an oven of 80°C for 20 minutes, the block of 3:1 RTV with air channels at the bottom was peeled off from the second mold. Air supply through-holes were punched. Aligned to the fluid pattern under a microscope, it was then pressed against the thin 30:1 RTV on the first mold. A post-bake of an hour at 80°C made the two silicone pieces chemically bond to each other. After peeling it off from the mold and punching the fluid through-holes, the monolithic RTV device could seal hermetically to a glass cover slip. In principle this glass cover slip could be chemically patterned in advance, as we describe in the next section, to make an active diagnostic chip. If high-resolution transparency photomasks were used (minimum feature size: ~10μm), the whole process from the design to the final products would be accomplished in just a day.

Snapshots of a final assembled device are shown in Figure 2. The central loop is the key component, at which the DNA hybridization probes are laid down along the ring on top of the glass substrate. DNA samples and later fluorescent intercalating dyes can enter from the two branches of the top T-channel. On/off states of each microvalves are controlled by external pneumatic valves (Lee LHDA1211111H) which either apply 100-kPa air pressure to the microvalves or vent them to the atmosphere. A maximum cycling frequency of 75 Hz has been demonstrated with complete opening and closing of the valves[4].

SURFACE PATTERNING

We developed two independent methods of surface patterning. The first method allows patterning of the protein streptavidin, a common biochemical "glue" that binds biotin with nearly covalent strength. With the streptavidin surfaces, one can selectively anchor biotin-labeled reagents, including proteins and nucleic acids. The second method allows direct attachment of amine-modified DNA molecules to the surface using a commercially available surface chemistry from the company Surmodics.

In the first method, we derivatized half of the surface of a cover slip (VWR #1) with biotin.[12] We then attached channels and flowed avidin-fluorescein conjugate down them, after which we flushed the channels with water, removed the channels, and flushed the entire cover slip. The avidin molecules bound to the derivatized part of the surface with high affinity in the regions defined by the channels, forming fluorescent stripes as shown on Figure 3 (a). The regions which were not derivatized with biotin function as a control and showed a much lower level of avidin binding.

We were also able to perform successive patterning steps to the surface. In this case we patterned the surface with non-fluorescent streptavidin by bonding an elastomeric device with channels to the substrate and flowing the streptavidin down the channels. As before, the streptavidin bound selectively to the surfaces that were exposed to the channels, and not to the surface covered by the elastomer. Since streptavidin is a tetramer, each molecule has at least two exposed groups free to bind more biotin. We demonstrated this by then removing the elastomeric channels and re-bonding in an orientation that was rotated by 90 degrees. We flowed biotin-fluorescein conjugates down the channels and then washed with water. From the images in figure 3 (b), it is clear that the fluorescent biotin binds selectively to the regions that are derivatized with streptavidin.

We also prepared surfaces patterned with DNA using commercially available silanized slides (Surmodics 3D-Link). DNA samples were prepared by PCR of a 2kpb region of lambda phage DNA using amino-terminated primers. The DNA was attached *in situ* by flowing it through an elastomeric channel replica made from the air channel mold of the diagnosis chip, i.e. the finger pattern in Figure 2 (b). After overnight incubation, the elastomeric device was peeled off from the slide, and washing and immobilization steps were followed according to the 3D-Link protocol. To show that the DNA was attached to the surface, we aligned and attached a diagnostic RTV device (as shown in Figure 2) to the same slide. Then we flowed the DNA intercalating dye PicoGreen (Molecular Probes P-7581) through the bottom fluid channel, of which the central ring intersected with every DNA finger pattern on the slide. The intersection of the channels fluoresced, as shown in Figure 3 (c).
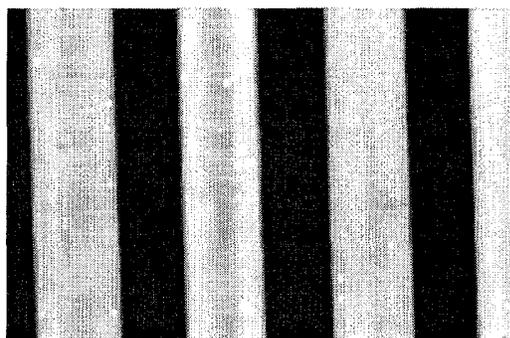
(a).
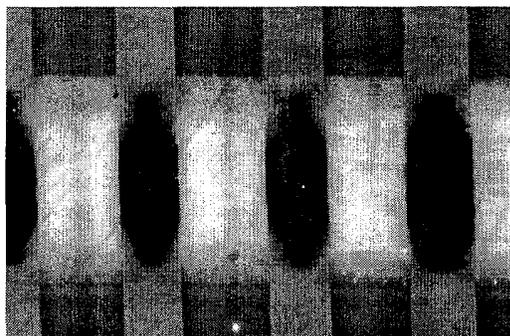
(b).
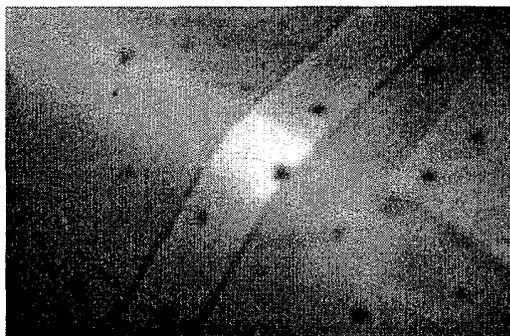
(c).

Figure 3. Images of chemically patterned cover slips. (a) A line pattern obtained by flowing avidin-fluorescein conjugates vertically on a biotinylated cover slip. (b) A checkerboard pattern obtained by flowing streptavidin horizontally (200µm) and biotin-fluorescein conjugates vertically (100µm). (c) DNA patterned on a Surmodics slide lights up when fluorescent dye PicoGreen was flowing in the central ring loop. Top-left to bottom-right was DNA with slightly auto-fluorescence. Top-right to bottom-left was part of the central ring of a diagnosis chip.

RESULTS

We have demonstrated two methods of chemical patterning of surfaces using soft lithography, and shown that the chemical patterns can also be interfaced with elastomeric microfluidics. These methods are the starting point for proof of principle of DNA diagnostics, and provide the foundation to demonstrate sensitive and selective DNA detection on chip. It is a straightforward extension to use protein-binding assays with these fluidic systems

as well. One issue that remains to be addressed is the strength of adhesion – the elastomer adheres to derivatized surfaces, but not with the same strength as to clean glass.

Our prototype DNA diagnostic chip has a junction for mixing and metering reagents, which then leads into a fluidic loop. The probe molecules will be anchored in the loop so that the targets can circulate around. The loop has peristaltic pumps around it to control circulation. The fluidic connections into and out of the loop are controlled by input and output valves, respectively.

We have demonstrated mixing and metering in the first part of the chip. A solution containing fluorescent dye can be mixed with an aqueous solution – the flows can be alternated with two different valve-firing schemes. If the valves are opened and closed in synchrony, the fluid mixing is controlled by diffusion. One observes two segregated flows when the valves are open, which quickly mix by diffusion when the valves are closed. If the valves are opened and closed alternately, then slugs of fluid are injected into the stream.

We have also pumped fluid within the closed loop using the peristaltic pumps on the perimeter. The channels were loaded with fluorescent beads (2.5 µm in diameter). The beads could be visualized as the fluid circulated around the loop and clearly showed rotary motion with no net flux into or out of the loop. Thus reagents can be repetitively exposed to diagnostic probes anchored on the surface, and their binding will not be limited by diffusion. All target DNA in a sample should be eventually captured by their corresponding probes after several passages. Such a device can also be used to rapidly mix viscous liquids, since the parabolic flow profile of the fluid will tend to "wrap" the two fluids around each other.

CONCLUSION

Here we have demonstrated elements of an active DNA diagnosis chip using multilayer soft lithography. DNA fragments can be patterned on a glass substrate with appropriate surface chemistry combination and by using elastomeric fluidic channels. A monolithic elastomeric diagnosis device can then be aligned and attached to the derivatized surface. DNA samples can be brought in from the top, circulated several times in the hybridization loop and expelled from the bottom outlet. Fluorescent dye used to determine the hybridization can be brought in from another input channel. Fast mixing can also be done with the active pumping mechanism. Laminar flow and diffusion limitation in this low-Reynolds number regime can be overcome easily by this active pumping agitation.

With the power of multilayer soft lithography, many complicated functions can easily be designed and fabricated as we have shown here. We anticipate that more active lab-on-a-chip devices will be developed rapidly. The problem of buffer depletion due to electrolysis in electro-osmotic flow control does not exist in these devices. Rotary mixing, rapid cell sorting, precise chemical metering and time-controlled reactions are all possible.

REFERENCES

1. A. T. Woolley, K. Q. Lao, A. N. Glazer and R. A. Mathies, "Capillary Electrophoresis Chips with Integrated Electrochemical Detection," *Analytical Chemistry*, 70 (4), pp. 684 (1998).
2. N. H. Chiem and D. J. Harrison, "Microchip Systems for Immunoassay: an Integrated Immunoreactor with Electrophoretic Separation for Serum Theophylline Determination," *Clinical Chemistry*, 44 (3), 591 (1998).

3. M. U. Kopp, A. J. de Mello and A. Manz, "Chemical Amplification: Continuous-flow PCR on a Chip," *Science*, 280 (5366), 1046 (1998).

4. L. C. Waters, S. C. Jacobson, N. Kroutchinina, J. Khandurina, R. S. Foote, and J. M. Ramsey, "Microchip devices for cell lysis, multiplex PCR amplification, and electrophoretic sizing," *Analytical Chemistry.*, 70 (1), 158 (1998).

5. M. A. Unger, H. P. Chou, T. Thorsen, A. Scherer and S. R. Quake, "Monolithic Microfabricated Valves and Pumps Using Multi-layer Soft Lithography," *Science* in press, (2000).

6. A. Y. Fu, C. Spence, A. Scherer, F. H. Arnold and S. R. Quake, "A Microfabricated Fluorescence-Activated Cell Sorter," *Nature Biotechnology*, 17, 1109 (1999).

7. G. Ramsay, "DNA chips: State-of-the art," *Nature Biotechnology*, 16, 40 (1998).

8. Diffusion length $l = \sqrt{Dt}$, where D is the diffusion constant and t is the time. Let D be 10^{-7} cm^2/s for a typical 1 kbp DNA and t be an hour. Then, the diffusion length l would be 0.19 mm. So, if only passive diffusion is used, each hybridization spot can only cover an area of about 0.4 mm in diameter. Even after a day, 24 hours, only target samples in an area of ~2 mm in diameter can possibly reach a specific probe to give a positive signal. This is an extremely slow process for big molecules like DNA.

9. Y. N. Xia and G. M. Whitesides, "Soft Lithography", *Angew. Chem. Int. Ed.*, 37, 550 (1998).

10. R. S. Kane, S. Takayama, E. Ostuni, D. E. Ingber, and G. M. Whiteside, "Patterning Proteins and Cells Using Soft Lithography," *Biomaterials*, 20 (23-24), 2363 (1999).

11. E. Delamarche, A. Bernard, H. Schmid, B. Michel, and H. Biebuyck, "Patterned Delivery of Immunoglobulins to Surfaces Using Microfluidic Networks," *Science*, 276, 779 (1997).

12. Marc A. Unger, Emil Kartalov, and Stephen R. Quake, manuscript in preparation.

INVESTIGATION OF A MINIATURIZED CAPILLARY ISOELECTRIC FOCUSING (cIEF) SYSTEM USING A FULL-FIELD DETECTION APPROACH

Amy E. Herr, Joshua I. Molho
Juan G. Santiago, and Thomas W. Kenny
Mechanical Engineering Department, Stanford University
Stanford, CA 94305-4021

David A. Borkholder, Gregory J. Kintz,
Phillip Belgrader, M. Allen Northrup
Cepheid
Sunnyvale, California

ABSTRACT

Sample handling in miniaturized bioanalytical instruments typically depends on pressure or electrokinetic effects. This is particularly true of post-separation detection processes where a sample is moved past a single point detector. In contrast to such methods, the present study investigates a separation system that takes advantage of a steady-state, stationary separation technique known on the macroscale as capillary isoelectric focusing (cIEF). Due to the stationary nature of the final separation, sample analysis is conducted entirely without mobilization. This work explores the merits and limitations associated with miniaturizing this technique. Analyses were conducted on fluorescent peptide and protein samples in miniaturized separation columns (capillaries and glass micromachined channels between 8 and 14 mm long). In the interest of developing overall system portability, blue light emitting diodes (LED's) were used for sample excitation. Detection was accomplished using a charge-coupled device (CCD). This approach yielded complete multi-sample separations in less than 1 minute.

INTRODUCTION

Miniaturization of biochemical analysis systems has been investigated for several years, as researchers attempt to enhance the performance of specific separation systems. Performance, cost, and overall system throughput has been predicted to improve with shrinking channel dimensions[1]. Systems developed using micro electro-mechanical systems technologies have enabled further miniaturization of the complete analysis system.

Capillary isoelectric focusing (cIEF) relies upon equilibrium between the diffusion and electrophoresis of a species within pH and voltage gradients to accomplish separations (see Fig. 1). In cIEF, individual species are driven to a stationary location within the separation column that corresponds to the isoelectric point (pI, the pH at which the net charge on the sample goes to zero) of the respective species. This equilibrium behavior results in a high density of tightly-focused sample species spatially distributed within the separation column[2].

Traditional cIEF relies on a dual-stage separation scheme. The first stage consists of sample focusing, while the second stage relies on a chemical, hydrodynamic, or electroosmotic mobilization scheme. The necessity of this second stage arises from efforts to perform cIEF in instruments originally designed for electrophoresis. Such instruments make use of a single point detector at the end of the separation column. Thus, sample mobilization, in the case of cIEF, does not take advantage of the stationary, steady state characteristics of the final separation.
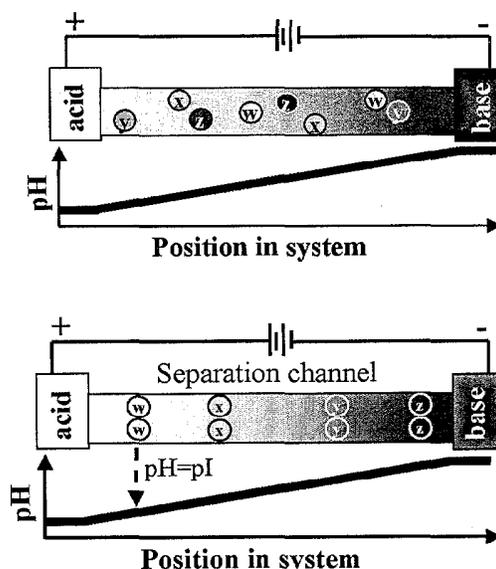
Figure 1. Schematic of cIEF separation. (top) *A pH gradient stabilized with a solution of ampholytes (represented by the white/black color gradient) has been established in the separation column. The protein samples, represented by the circles w-z, are distributed throughout. **(bottom)** An applied axial electric field causes the proteins to electrophoretically migrate to the column position where the local pH equals the pI of the protein.*

Several studies [3,4] have used small separation column lengths for cIEF, but have sacrificed overall system portability by relying on bulky excitation sources. Significant reductions in column length (from a typical length of 30 cm to that of 4-5 cm) have permitted the entire separation column to be imaged via a full-field detector such as a CCD[5]. This allows for the collection of detailed spatial and temporal information. This method of detection also significantly shortens the analysis time[6] and virtually eliminates mobilization-related dispersion, as the second-stage of the traditional dual-stage approach is entirely omitted. This work explores some of the limits in the miniaturization of cIEF and investigates the performance of a miniaturized prototype system.

THEORY

cIEF Model

In cIEF, a mixture of protein samples is subjected to a strong electric field in a medium designed to have a linear pH gradient. At the molecular level, samples electrophoretically migrate along the channel until they reach the point in the pH gradient that is equal to their isoelectric pH (pI), where the protein no longer has a net charge. At the pI, the electrophoretic force on the protein due to the applied electric field is zero. Once the separation is complete and all samples have reached their respective isoelectric points, the separation is simply a stationary spatial distribution.

After the protein samples have "focused" at their isoelectric point locations, a steady-state equilibrium is achieved between electrophoresis and diffusion. In this final stage, the distribution of the protein band is determined by the opposing forces of diffusion (which tends drives proteins away from their pI locations) and an electrophoretic force that restores proteins toward their pI locations. This balance between electrophoresis and sample diffusion can be described by:

$$C\mu E = D\frac{dC}{dx} \qquad (1)$$

with C as the sample concentration; μ, the electrophoretic mobility; E, the applied field strength; D, the diffusion coefficient of the species; and dC/dx, the concentration gradient. For simplicity, assume that the pH gradient, d(pH)/dx, and the mobility slope, $d\mu/d(pH)$, are constant within the small interval associated with a focused protein sample, this differential equation yields a modified Gaussian solution for the concentration distribution at steady state given by:

$$C(x) = C_o \exp[\frac{-x^2}{2\sigma^2}] \qquad (2)$$

where σ in Eq. 2 is:

$$\sigma = \pm\sqrt{\frac{D}{E}\frac{dx/d(pH)}{-d\mu/d(pH)}} \qquad (3)$$

As the values of the free parameters (i.e., d(pH)/dx, E) increase, the width of the distribution decreases.

Resolving Power

Resolving power for a particular cIEF system, be it macro- or microscale, is simply the minimum allowable proximity of two focused neighboring zones. The definition commonly used relates the inflection points (0.61 of the concentration maximum) of two identical neighboring Gaussian curves to the concentration minimum between the two concentration maxima. Svensson[7] proposed that when the curves are 3σ apart, the aforementioned concentration minimum is as deep as the inflection points on the Gaussian curves making this a resolvable sample separation (Fig. 2).

Using this criterion, the minimum pH separation between adjacent zones can be described as:

$$\Delta(pH)_{min} = (\Delta L)_{min}\frac{d(pH)}{dx} = 3\sigma\frac{d(pH)}{dx} \qquad (4)$$

To determine the resolving power[8], set ΔpH equal to ΔpI and replace σ with Eq. (3):
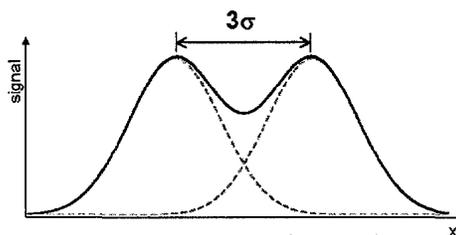
Figure 2. Minimum resolution of equivalent neighboring Gaussian peaks. *The solid line shows the sum of the two peaks, the dashed lines show the individual distributions.*

$$\Delta(pI)_{min} = 3\sqrt{\frac{D}{E}\frac{(dpH/dx)}{(-d\mu/d(pH))}} \qquad (5)$$

Substituting in the definition of the applied electric field strength, $E = \phi/L$, in terms of applied potential, ϕ, and the length of the separation channel, L, and assuming that ϕ is held constant, the minimum resolvable pI difference is given by:

$$\Delta(pI)_{min} = 3\sqrt{\frac{D}{(-d\mu/d(pH))}\frac{(dpH)}{\phi}} \qquad (6)$$

Thus the minimum resolvable isoelectric point difference between adjacent focused bands is approximately independent of L and dependent only upon the individual protein characteristics, the total pH difference across the separation channel, and the applied potential. This channel length independence hypothesizes significant column miniaturization with minimal loss in the separation system performance. The limit to such miniaturization is probably determined by a maximum current density which, in turn, determines the maximum allowable temperatures and temperature gradients in the system. This effect is discussed below.

Temperature Considerations

Holding the applied potential constant while shrinking the length of the separation column (as in the discussion above) will increase the field strength and, thus, lead to increased Joule heating within the separation column. Such heating effects can cause otherwise plug-like sample bands to experience enhanced diffusion near the channel centerline because of local changes in viscosity. Such sample dispersion can limit the resolving power of miniaturized cIEF separation, as neighboring samples zones 'merge' into one another.

Researchers have investigated the effects of temperature gradients on other separation schemes, namely capillary electrophoresis[9]. This investigation is concerned with sample dispersion as a function of temperature. Assuming steady-state operation, only radial temperature variation, and an isotropic thermal conductivity, the heat conduction equation is given by:

$$\frac{1}{r}(\frac{d}{dr})(r\frac{dT}{dr}) = -\frac{\dot{q}}{k} \qquad (7)$$

where r is the radial direction, T is the temperature, $\dot{q}$ is the heat generation rate, and k is the thermal conductivity. Integrating this equation and applying a constant surface temperature boundary condition, results in a solution for the temperature distribution:

$$T(r) = T_o + \frac{\dot{q}(r_o)^2}{4k}[1-(\frac{r}{r_o})^2] \qquad (8)$$

116

where r_o is the inner radius of the separation channel, in this example 50 μm. Note that, in some cases, the assumption of fixed wall temperature may not be valid but serves here to demonstrate the effects of temperature-gradient-driven dispersion. This analysis estimates the temperature dependence of the viscosity[10] to be:

$$\eta_{r=o} \approx \eta_o exp[a + b(\frac{T_{r=r_o}}{T_{r=o}}) + c(\frac{T_{r=r_o}}{T_{r=o}})^2] \qquad (9)$$

where T_o = 273K, η_0 = 0.001792 kg/(m sec), and a = -1.94, b = -4.80, and c = 6.74. Making use of the Nernst-Einstein equation, $D=\mu RT$, relates the mobility, $\mu=\varepsilon\zeta E/\eta$, and the temperature to the diffusivity of the sample. Where ε is the permitivity of the buffer, ζ is the zeta-potential, and E is the applied field strength. This describes how the diffusivity of the sample in a miniaturized cIEF system compares to that in a benchmark macroscale system (see Table 1).

Due to the exponential nature of the temperature dependence of the viscosity, channels less than 1 mm in length experience a rapid increase in the centerline diffusivity. So, as the separation channel is shortened below 1 mm, this analysis predicts significant sample blending. The choice of channel length investigated in this study (8 mm) avoids this significant centerline dispersion, yet allows for rapid separations and full-field imaging.

Miniature cIEF: Separation Channel Length (cm)	Comparison with Macroscale cIEF: Increase in Diffusivity (% Increase in $D_{centerline}$)
10.0	< 0.1
0.5	4
0.1	23

Table 1. Joule heating and diffusivity. Comparison of centerline diffusivity for a miniaturized system (channel lengths given in the first column) to that of a macroscale system (channel length is 20 cm). Both cases are for a 50 μm channel radius. The applied potential is held constant and all sample parameters are the same for each case.

Other limitations to the miniaturization of the cIEF system exist. Among these are the absolute temperature rise with respect to the surrounding and the performance of the detection system. The former limitation is concerned with the maximum temperature in the system and is a strong function of the size of the cross-sectional geometry of the separation column and the thermal properties of the capillary walls and sample holder. In turn, the maximum allowable temperature is a function of the type of protein that is used. The later consideration has significant impact on the design for portability of the miniaturized system. These limitaions will be explored in detail in future work.

EXPERIMENTAL APPARATUS

No work at present addresses the need for portability of the entire cIEF system: separation channel, excitation source, and detection scheme. Previous works utilize a CCD detector and a miniaturized separation channel, but rely on lasers for fluorescence excitation. For example, Pawliszyn[3] investigated a non-mobilized system composed of an array detector, 4-5 cm long separation channel, and an Argon ion laser. The lasers employed are bulky and require cooling. This work investigates the performance of an excitation system based on a compact (3 mm in diameter), low-power (~240 mW), inexpensive (less than $5 each) LED source (Chicago Miniature; Hackensack, New Jersey).

The prototype non-mobilized system consists of an array detector, an 8 mm separation column (Polymicro Technologies, Tempe, Arizona) permanently mounted between fluid reservoirs, and an LED array excitation source (see Fig. 3). UV-curable refractive index matching epoxy (Norland Optical Adhesives, New Brunswick, New Jersey) was used to affix the column to the Delrin (DuPont; Wilmington, Delaware) cartridge. These short columns, compared to tens of centimeters in traditional cIEF, can be imaged and simultaneously monitored until the optimum separation is achieved.

The fused-silica channels were cleaned and prepared prior to each separation. The channels were rinsed for 5 min. with a 1 M NaOH solution, followed by a 5 min. 20 mM NaOH solution rinse. To suppress electroosmotic flow (EOF), a neutral EOF-suppressing polymer, in this case a 0.4% solution of methylcellulose, was then rinsed through the channel for 5 min. and allowed to equilibrate in the channel for approximately 15 min. Proper EOF suppression is critical[11] for a non-mobilized system.

Proposed sample detection techniques involving cIEF suggest that fluorescence detection increases detection limits over more traditional schemes based on UV detection, as UV absorbing ampholytes interfere with sample detection[12]. In the present study, fluorescently-tagged peptides were prepared and diluted to the appropriate concentrations with ampholytes (Bio-Rad CE-IEF Ampholytes 3/10)[4]. The sample mixture was then injected into the prepared capillary cartridge. The reservoirs in the capillary cartridge were filled with basic catholyte (20mM NaOH) or acidic anolyte (10mM H_3PO_4). The cartridge was placed in the miniaturized system, platinum electrodes were inserted into each reservoir, and a potential (0 kV-2 kV) was applied. A CCD camera (Hamamatsu S7034-0907) with PC interface was used to acquire real-time spatial intensity distribution information (Fig. 4).

RESULTS AND DISCUSSION

Hydrodynamic Mobilization

This full-field detection approach was used to extract information about the dispersion effects introduced through a pressure-driven sample mobilization. To accomplish this, samples were allowed to separate and were then mobilized by a slight pressure-driven flow (mobilization velocity of ~ 0.3 mm/sec) (see Fig. 5). The electric field was maintained during sample mobilization, so as to minimize sample dispersion due to diffusion (this approach is often applied in macroscale systems).

By the time the sample was mobilized out of the short separation channel, the overall width of the sample had increased significantly. An estimate of the number of plates, N, obtainable in the hydrodynamically-mobilized system (see Fig. 6) was based upon a skewed peak model developed by Foley and Dorsey[13]. While the hydrodynamic mobilizaiton velocity used in this study is on the same order of that traditionally used in macroscale systems (i.e., u ~ 0.3 mm/sec vs. u ~ 0.1 mm/sec), the loss in system efficiency is likely to be exaggerated as compared to macroscale cIEF. This arises from the fact that this investigation was conducted in free solution; whereas macroscale cIEF is often performed in a gel matrix. The gel acts to resist the pressure-driven flow, hence a direct comparison of efficiency between the two schemes is not appropriate. The overall dispersive trends, however, are illustrutive and indicate the existance of enhanced sample band dispersion due to mobilization. The efficiency values do indicate directly that such hydrodynamic mobilization in a miniature cIEF free solution system utilizing a single-point detector would significantly reduce the system performance.
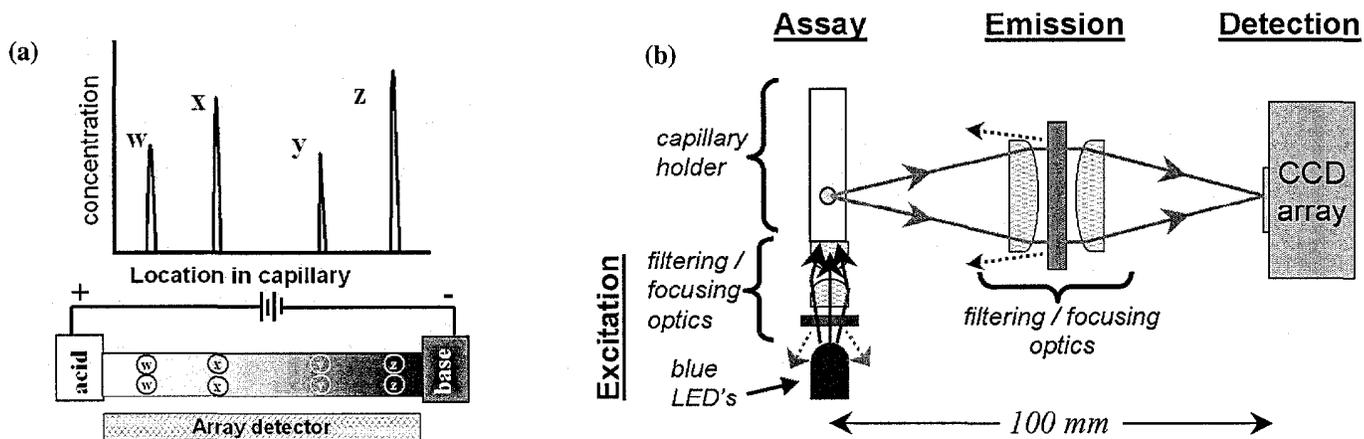
Figure 3. Schematic of protein-fluorescence array detection. *An axial electric field has been applied, resulting in the formation of an axial pH gradient within a background field of ampholytes (represented by the white/black color gradient). Proteins initially distributed homogeneously in the column migrate to their respective pI location (represented by circles w-z). (a) Concentration/intensity vs. spatial location information is obtained from the detection system. (b) A series of blue LEDs provide excitation for fluorescently-tagged proteins in a cIEF assay. The emission is then focused onto a CCD array that is then vertically binned.*
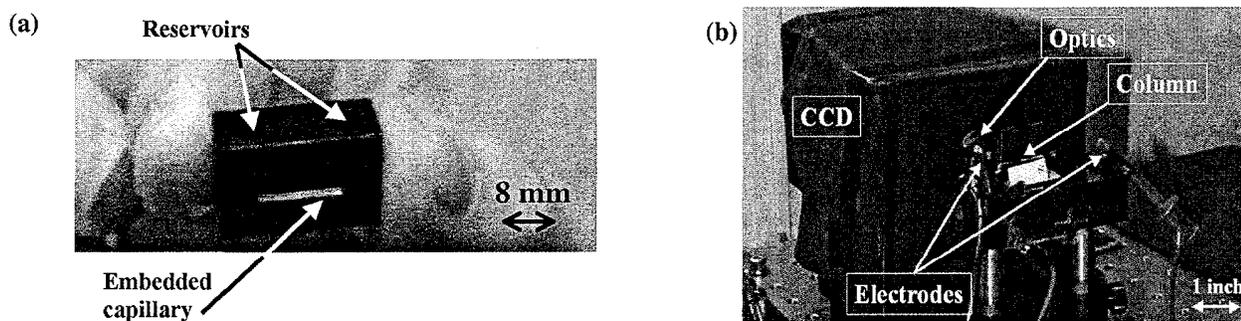
Figure 4. Miniaturized IEF System. *(a) Capillary cartridge and (b) the hardware mount for the excitation source and optics, the cartridge mount, the emission filtering/focusing optics, and the CCD detection component of the diagnostic are shown.*
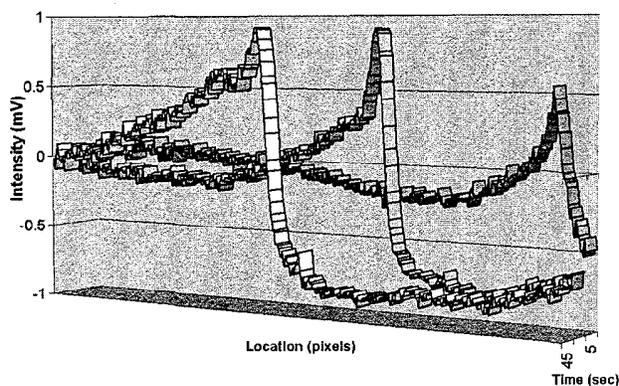
Figure 5. Effect of hydrodynamic mobilization of a single protein. *Hydrodynamic mobilization (u ~ 0.3 mm/sec) of 100 nM protein sample. The protein is focused and pressure-driven flow has been applied for (from back to front) 5 sec. ($\Delta t_p = 5s$), 25 sec., and 45 sec. The sample moves out the end of the column where a single point detector would potentially be located. Capillary I.D. 75μm, O.D. 360μm, and length 8 mm*
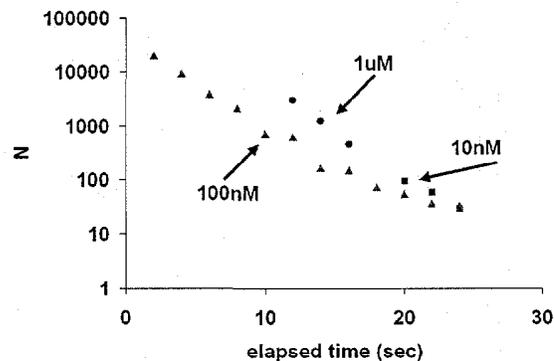
Figure 6. Effects of hydrodynamic mobilization on the number of plates (N) attainable. *Protein samples were focused and subsequently mobilized through the separation column using pressure driven flow. The width of a given peak was measured and N was calculated. This initial data shows a degradation in the system performance as the sample moves past an imaginary single point detector. The real-time, array detector approach allows us to image such dynamic processes.*

Multi-Sample Separation

The short separation column (8 mm) in the prototype cIEF system allows the entire column to be imaged; thus, enabling the separation to be monitored continuously and captured when it has reached steady state. Figure 7 illustrates this approach, showing a complete separation of two peptides taking place in less than one minute, compared to tens of minutes for traditional cIEF.

After steady-state is achieved, the separation remains stable exhibiting the profile dipicted in Fig. 7 for tens of minutes. The two sample peaks are separated by 5σ (using the σ of the wider peak, Peak 1) and are, thus, quite easily resolved. This quality agrees well with the previous resolving power analysis, as peaks separated by more than 3σ should be readily resolvable. Also, consideration of the initial temperature gradient analysis suggests that there is limited Joule heating (less than 3% more than would be predicted for a 200 cm long separation channel) in this system.
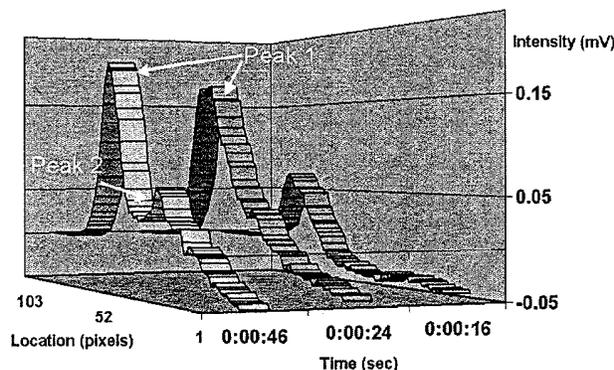
Figure 7. Array detection of a cIEF separation of two fluorescently tagged protiens. The separation column was filled with a homogenous mixture of ampholytes (BioLyte CE-IEF Ampholytes 3/10) and 500 nM sample. One reservoir was filled with an acid (pH 2); the other was filled with a base (pH 10). Upon application of the electric field, (16 sec) the samples start to focus into one peak. (24 sec) A small second peak is moving right and separates from the major peak at approximately 46 sec. Two peaks are then clearly resolved. The estimated pH separation was calculated to be $\Delta pH = 0.65$ pH units. $E = \sim 500$ V/cm, capillary I.D. $75\mu m$, O.D. $360\mu m$, and length 8 mm.

CONCLUSIONS

The short (8 mm) separation columns used in this miniaturized cIEF system allow for real-time, full-field imaging of the complete separation process with a CCD. This point is of particular importance as traditional cIEF relies on a two-stage sample analysis protocol: the sample separation and mobilization of the sample past a point detector. During any mobilization scheme, diffusion and distortion of the separated sample bands may occur, resulting in degraded system performance. Thus, the full-field imaging system allows for less degradation in the final system resolution, as mobilization is unnecessary.

This prototype miniaturized cIEF system allows for the study of various system parameters (e.g., the applied field strength, axial pH gradient, and assay characteristics) on the overall separation efficiency. This system consisted of a short separation column permanently mounted between fluid reservoirs with LED-based fluorescence excitation and CCD detection. Initial work with this prototype system investigated fundamental principles of cIEF, including the resolving power of a cIEF system and temperature

gradient considerations limiting the miniaturization of the separation component of the system. This work successfully demonstrates rapid multi-component sample separations in less than 1 minute using the miniature full field imaging system.

REFERENCES

1. A. Manz et al., "Planar Chips Technology for Miniaturization of Separation Systems: A Developing Perspective in Chemical Monitoring", *Advances in Chromatography*, **33** (1993), 2-66.
2. P.G. Righetti, *Immobilized pH Gradients: Theory and Methodology* (Amsterdam: Elsevier, 1990).
3. X.-Z. Wu, J. Wu, and J. Pawliszyn, "Fluorescence imaging detection for capillary isoelectric focusing", *Electrophoresis*, **16** (1995), 1474-1478.
4. K.A. Cruickshank, J. Olvera, and U.R. Muller, "Simultaneous Multiple Analyte Detection Using Fluorescent Peptides and Capillary Isoelectric Focusing", *Journal of Chromatography A*, **817** (1998), 41-47.
5. J. Wu and J. Pawliszyn, "Absorption Spectra and Multicapillary Imaging Detection for Capillary Isoelectric Focusing Using a Charge Coupled Device Camera", *Analyst*, **120** (1995), 1567-1571.
6. Q. Mao and J. Pawliszyn, "Capillary isoelectric focusing with whole column imaging detection for analysis of proteins and peptides", *Journal of Biochemical and Biophysical Methods*, **39**, no. 1-2, (1999), 93-110.
7. H. Svensson, *J. Chromatogr.*, **25** (1966), 266-273.
8. H. Rilbe, "Theoretical Aspects of Steady-State Focusing," in *Isoelectric Focusing*, ed. N. Catsimpoolas (New York: Academic Press, 1976), 13-52.
9. E. Grushka, R. M. McCormick, and J.J. Kirkland, "Effect of Temperature Gradients on the Efficiency of Capillary Zone Electrophoresis Separations", *Anal. Chem.*, **61** (1989), 241-246.
10. F.M. White, *Fluid Mechanics* (San Francisco: McGraw-Hill, Inc, 1994).
11. M. Minarik et al., "Dispersion Effects Accompanying Pressurized Zone Mobilization in Capillary Isoelectric Focusing of Proteins", *Journal of Chromatography A*, **738** (1996), 123-128.
12. G.F. Verbeck and S.C. Beale, "Isoelectric Point Analysis of Proteins and Peptides by Capillary Isoelectric Focusing with Two-Wavelength Laser-Induced Fluorescence Detection", *J. Microcolumn Separations*, **715** (1999), 708-715.
13. J. P. Foley and J. G. Dorsey, "Equations for Calculation of Chromatographic Figures of Merit for Ideal and Skewed Peaks", *Analytical Chemistry*, **55**, no. 4, (1983), 730-737.

A MEMS RADIO-FREQUENCY ION MOBILITY SPECTROMETER FOR CHEMICAL AGENT DETECTION

Raanan A. Miller, Gary A. Eiceman*, Erkinjon G. Nazarov*, Thomas A. King

Charles Stark Draper Laboratory, MS 37, 555 Technology Square, Cambridge, MA 02139

*Department of Chemistry, New Mexico State University, Las Cruces, NM 88003

ABSTRACT

A first-of-a-kind MEMS Radio-Frequency Ion Mobility Spectrometer (RF-IMS) with a miniature drift tube (0.6 cm^3) has been fabricated and tested. The spectrometer has detection limits in the parts-per-billion and the ability to identify chemicals not resolved in conventional time-of-flight ion mobility spectrometry. Spectrometer operation with a miniature 10.6eV (λ=116.5 nm) UV photo-discharge lamp and a 100 μCi radioactive ionization source has been demonstrated. The RF-IMS has been interfaced to a mass spectrometer and RF-IMS spectral peaks have been confirmed. The RF-IMS/MS configuration illustrates another use for the RF-IMS as a pre-filter for atmospheric pressure chemical ionization mass spectrometry applications.

INTRODUCTION

Ion Mobility Spectrometry is the leading technology for on-site detection of chemical warfare agents, explosives, and illicit drugs [1]. It is also the technology chosen by NASA for air quality monitoring aboard the international space station [2]. IMS is inherently suitable for field operation as it utilizes a relatively simple fluidic system and operates at atmospheric pressure. It is highly sensitive, portable, has low power consumption and moderate cost.

Commercially available IMS systems are based on time-of-flight (TOF), i.e., they measure the time it takes ions to travel from a shutter-gate to a detector through an inert atmosphere (1 to 760 Torr). The drift time is dependent on the mobility of the ion (i.e., its size, mass and charge) and is characteristic of the ion species detected [1]. The central component of the IMS is the drift tube, which ultimately determines the instrument's performance [3]. The drift tube consists of an ionization region, a shutter, a drift region and a detector. Current spectrometers use conventionally machined drift tubes (minimum size about 40 cm^3) for ion identification, but future applications will require drift tube designs and fabrication technologies which enable even smaller, lower cost, lower power and more reproducible spectrometers. MEMS technology offers a natural path to realizing such spectrometers. Yet to-date, attempts to fabricate miniaturized MEMS IMS drift tubes have lead to spectrometers with poor sensitivity or resolution [4,5].

In this work, a recently developed IMS-based technique, RF-IMS, has been adopted. This method uses the non-linear mobility dependence of ions on high strength RF electric fields for ion filtering [6]. The RF-IMS technique scales down well, allowing miniaturization of IMS drift tubes through MEMS, while preserving sensitivity and resolution [7]. In this paper, we report on the fabrication and characterization of a novel MEMS RF-IMS.

RF-IMS OPERATION PRINCIPLE

Conventional time-of-flight ion mobility spectrometers (TOF-IMS) operate in the low field regime where the applied field strength is less than 1000 V/cm and the mobility is essentially constant. However, Mason and McDaniel [8] found that the mobility of an ion is field dependent and can change significantly as the field strength increases. Different ion species will have particular mobility dependencies on electric field, see Fig. 1. The RF-IMS utilizes the change in mobility between the high field and low field conditions to identify the different ion species.
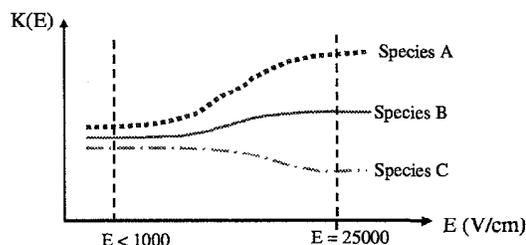
Figure 1. *Mobility dependence on electric field for three hypothetical ions.*

Fig. 2 illustrates the operation of the RF-IMS. As a gas sample is introduced into the spectrometer it is ionized and the ions are transported through an ion filter towards a detector by a carrier gas. The ion filter is a tunable filter which selects the ion species allowed to pass through the filter by adjusting the RF and compensation electric fields applied between the ion filter electrodes. The filtering mechanism is governed by the interaction between the ion and the net applied field which alternates between high and low electric field strengths.
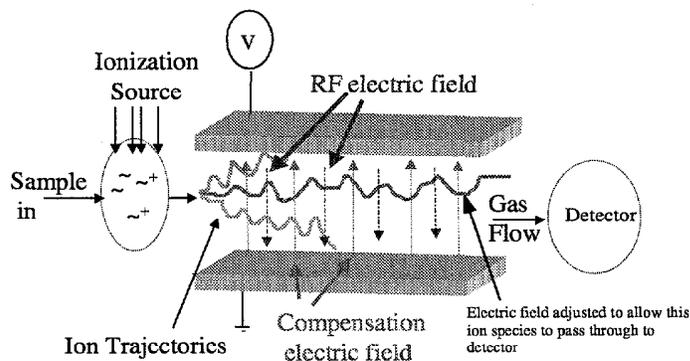
Figure 2. *Schematic showing ions filtered by the RF and compensation fields applied between the filter.*

The conventional IMS operates with short pulses of ions (usually produced by a shutter-gate) which allow only about 1% of the ions formed in the ionization region to reach the detector [3]. Thus, the ion current is limited by the fixed number of ions contained in the ion pulse. Miniaturization of the conventional IMS therefore, can significantly decrease the number of ions available for detection. The pulse width can not be increased however, since resolution in the TOF-IMS is dictated by the width of the ion pulse. The wider the pulse, the poorer the resolution.

A significant advantage of the RF-IMS approach is that it does not require ion pulses for operation and the resolution is not dictated by the width of the ion pulse. The ions are introduced continuously into the ion filter and almost 100% of the "tuned"

ions are passed through the filter, maintaining the high sensitivity of the device. The RF-IMS approach also avoids the complexity of generating short, spatially well confined, ion pulses. In fact, the RF-IMS approach actually benefits from miniaturization since the electric fields required to filter the ions are on the order of 10,000 V/cm. By keeping the dimensions small, about 500 microns, the voltages required for ion filtering are easily realizable.

Further insight into the operation principle of the RF-IMS can be gained by considering a single ion, transported by a gas stream between two parallel electrodes as shown in Fig. 3a. The ion experiences a force from the carrier gas flow (z-directed) which transports it through the filter. A transverse force (y-directed) produced by combining the RF electric field and a DC compensation field also acts on the ion. These fields are generated by applying voltages to the parallel plate electrodes. A (simplified) RF electric field waveform, Fig. 4a, with maximum field strength $|E_{max}| > 10,000$ V/cm and minimum field strength $|E_{min}| << |E_{max}|$ is used here to illustrate the ion filtering principle. The asymmetric RF waveform is designed such that the time average electric field is zero and

$$|E_{max}| \, t_1 = |E_{min}| \, t_2 = \beta \qquad (1)$$

where t_1 is the portion of the period where the high field is applied, t_2 is the time the low field is applied, and β is a constant corresponding to the area under-the-curve in the high field and low field portions of the waveform. The RF field causes the ions to oscillate in a direction transverse to the carrier gas flow as they are transported down the channel by the carrier gas, Fig. 3a. The ion velocity in the y-direction is given by [2]

$$V_y = KE \qquad (2)$$

Here K is the coefficient of ion mobility for the ion species and E is the electric field intensity, in this case entirely in the y-direction. At a constant density N of carrier gas the dependence of the mobility on the electric field intensity can be represented by the following expression [7,8]

$$K(E) = K_o[1 + \alpha_2 E^2 + \alpha_4 E^4 + ..] \qquad (3)$$

where α_2 and α_4 are coefficients of a series expansion, and K_o is the mobility coefficient in a vanishingly small field [6]. As the electric field strength increases (above 5000V/cm) the second and higher order terms in the series become significant and the mobility coefficient can change substantially (10 – 15%) from its low field value, Fig. 1 [7]. For illustration here, it is asumed the mobility increases with increased field strength therefore K_1 (at E_{max}) > K_2 (at E_{min}). The ion displacement from its initial position in the y-direction is the ion velocity in the y-direction, V_y, multiplied by the length of time, Δt, the field is applied

$$\Delta y_{RF} = V_y \, \Delta t \qquad (4)$$

In one period of the applied RF field the ion moves in both the positive and negative y-directions. By substituting equation (2) into equation (4) the average displacement of the ion over one period can be written as

$$\Delta y_{RF} = K_1 |E_{max}| \, t_1 - K_2 |E_{min}| \, t_2 \qquad (5)$$

Using equation (1) this expression can be re-written as

$$\Delta y_{RF} = \beta \, (K_1 - K_2) \qquad (6)$$

Since β is a constant determined by the applied RF field, the y-displacement of the ion per period depends on the change in mobility of the ion between the high and low field conditions. Assuming the carrier gas only transports the ion in the z-direction,

the average displacement of the ion in the y-direction as a function of time can be expressed as

$$y(t) = \frac{\Delta y_{RF}}{(t_1 + t_2)} t \qquad (7)$$

see Fig. 3b. When a low strength DC field, Fig. 4b, is applied to the ion filter electrodes ($|E_c| < |E_{min}| << |E_{max}|$) in a direction opposite to the average RF-induced (y-directed) motion of the ion, the trajectory of the ion species can be "straightened". This allows the ions of particular species to pass unhindered between the ion filter electrodes, Fig. 3c, while ions of all other species get deflected into the filter electrodes. The DC voltage that "tunes" the filter and produces a field which compensates for the RF-induced motion is characteristic of the ion species and is called the compensation voltage. A complete spectrum for the ions in the gas sample can be obtained by ramping or sweeping the DC compensation voltage applied to the filter.
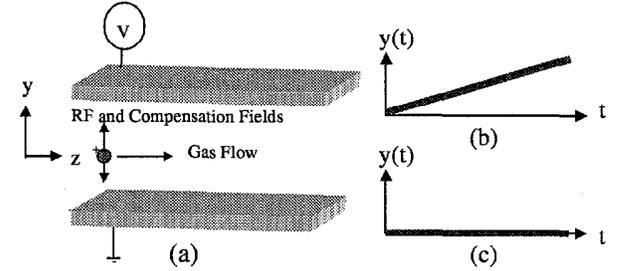
Figure 3(a) *Schematic of a single positive ion located between the ion filter electrodes. **(b)** Average y-displacement of ion from initial position with RF field applied but no compensation field applied. **(c)** Average y-displacement with both RF and compensation fields applied.*
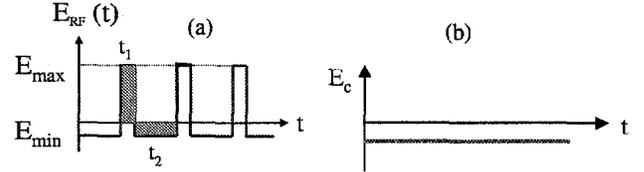
Figure 4(a) *RF waveform applied to filter electrodes.**(b)** Compensation voltage applied to cancel out displacement produced by RF-field.*

DRIFT TUBE AND RF CIRCUIT DESIGN AND FABRICATION

The fabricated MEMS RF-IMS drift tube, shown in Fig. 5, contains an ionization region, tunable ion filter electrodes, and a detector. The drift tube is fabricated from two Pyrex wafers and a heavily boron doped silicon wafer. Metal electrodes are fabricated on the top and bottom Pyrex wafers forming the ion filter, deflector and detector electrodes. The metal electrodes (1200 angstrom gold on 400 angstrom titanium) are formed on the Pyrex wafers by a sputter deposition and metal lift-off process. The Pyrex wafers are diced into rectangular pieces once the metal electrodes have been formed. Holes are drilled into the top Pyrex pieces to provide a means for the ionization sources to interact with the sample gas. The silicon wafer thickness defines the gap between the top and bottom electrodes of the ion filter. In the present method, the silicon wafer is diced into strips that are 25 mm long, 2 mm wide, and 0.5 mm thick. The silicon strips are aligned with the metal electrodes and anodically bonded (~ 1000V, 350°C) to the Pyrex. Metal electrodes also make contact to these silicon strips allowing control of their potential.
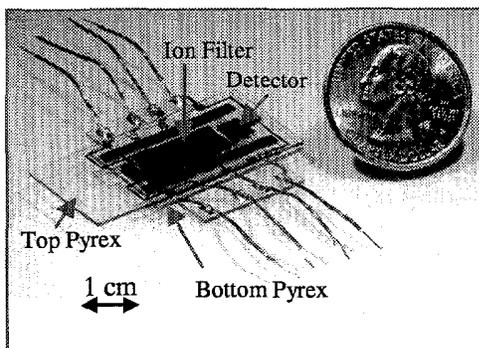
Figure 5. *Photograph of a micromachined RF-IMS drift tube.*

The RF field required to filter the ions is produced by a soft-switched semi-resonant circuit that incorporates a flyback transformer to generate the high voltage pulses [9]. The circuit, shown in Fig. 6, provides a peak-to-peak RF voltage of 1400 volts at a frequency of 2 MHz with a duty cycle of about 30%. In the practical implementation, the RF voltage is superimposed onto the compensation voltage.
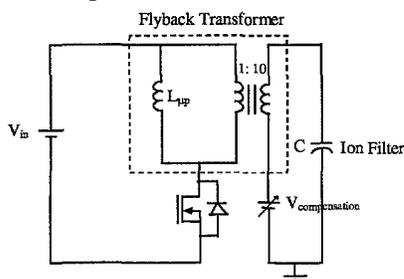
Figure 6. *RF waveform generator circuit uses a soft-switched semi-resonant design in order to minimize power consumption.*

RESULTS AND DISCUSSION

A schematic of the experimental setup used to characterize the spectrometer is shown in Fig. 7. Sample vapors were prepared in a vapor generator using Teflon permeation cells. All supply lines following the vapor generator were heated to approximately 50°C. For spectrometer characterization; the vapor generator was interfaced to the RF-IMS via a Teflon test fixture.

Either a miniature 10.6 eV (λ=116.5 nm) UV photo-discharge lamp or a 100 μCi radioactive americium source served as the ionization source for the RF-IMS. The source was positioned above the ionization region of the drift tube as shown in Fig. 7. The bottom ion filter electrode was grounded while the high voltage RF field and compensation voltages were applied to the upper filter electrode. A polynomial waveform synthesizer was used to sweep the compensation voltage and a Tektronics storage oscilloscope was used to record the signal from the detector amplifier and produce a spectrum for the sample gas.

Figure 8 shows the response of the RF-IMS to different concentrations of toluene vapor introduced into the spectrometer. In this experiment a UV lamp was used as the ionization source. The characteristic compensation voltage for toluene was found to be –7 volts. The peak intensity and area-under-peak depend linearly on concentration for the concentration range measured. Concentrations of toluene as low as 100 parts-per-billion have been detected. The spectrometer was also characterized with other chemicals such as acetone and benzene and each chemical exhibited its own characteristic compensation voltage peak.

The small ionization volume of the MEMS RF-IMS allows the use of low activity ionization sources such as the 100 μCi
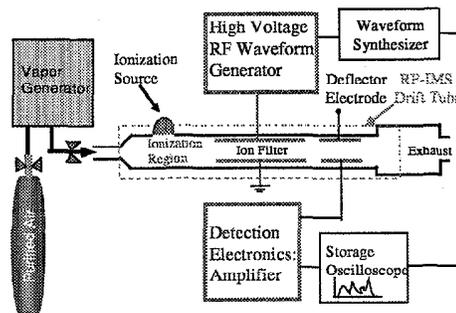
Figure 7. *RF-IMS characterization setup.*
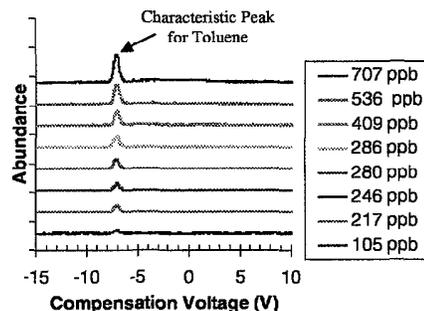
Figure 8. *MEMS RF-IMS spectra for several different toluene vapor concentrations.*

americium source. The americium source is the same material and activity as employed in household smoke detectors. Adopting the americium source for ionization will enable the RF-IMS to be used in many civilian applications where low power is important, while avoiding the extensive regulations typical of traditional TOF-IMS which use radioactive Ni^{63}. Fig. 9 shows the spectra obtained using the americium source with dimethyl methylphosphonate (DMMP- a chemical warfare agent simulant) at concentrations ranging from 300 to 90 parts-per-billion. Detection of lower concentrations were limited by the minimum vapor concentrations which the vapor generator system could deliver. The characteristic peak for DMMP is found to occur at a positive compensation voltage in contrast to toluene, acetone, and benzene which have negative compensation voltages. The small peak at about –4.5 volts, Fig. 9, is believed to be due to an impurity present in the carrier gas which is ionized by the radioactive source. The concentration of this impurity appears to be constant throughout this experiment.
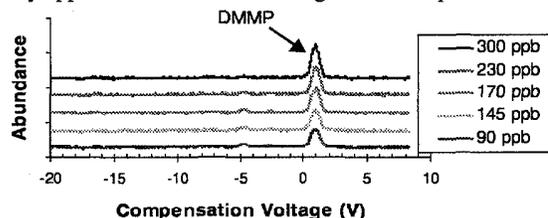
Figure 9. *MEMS RF-IMS spectra for DMMP generated using a radioactive americium ionization source.*

Because the RF-IMS operates in a different mobility regime compared to the traditional TOF-IMS, chemicals which are not possible to resolve in a conventional TOF-IMS can be resolved using the RF- IMS. For example, Fig. 10a shows spectra for m- and p-xylene isomers obtained with a conventional TOF-IMS. These molecules have the same mass and very similar effective cross-sectional areas. As a result, the spectra for these two ions are virtually identical. In contrast, Fig. 10b shows the spectra obtained for these same chemicals in the MEMS RF-IMS. The peaks in this case are clearly resolved.
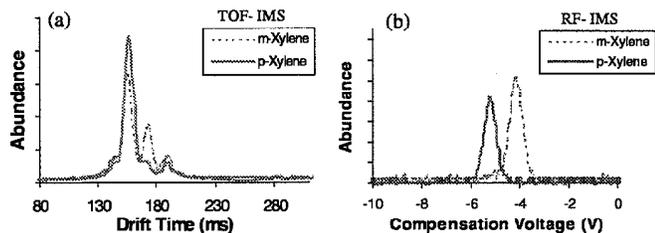
Figure 10(a) *Conventional TOF-IMS spectra for m-xylene and p-xylene.* ***(b)*** *Spectra for m-xylene and p-xylene obtained with the MEMS RF-IMS.*

The response of the MEMS RF-IMS to mixtures has also been investigated. Fig. 11a shows the resultant spectra for a mixture of benzene and acetone at trace concentrations ionized with a UV source. The acetone and benzene peaks are well resolved with the acetone peak occurring at a compensation voltage of –3.5 volts and the benzene peak at about –10.5 volts. Fig. 11b shows spectra for benzene and acetone measured independently in a conventional TOF-IMS. It is evident that a mixture of these two chemicals would be poorly resolved in a traditional TOF-IMS.
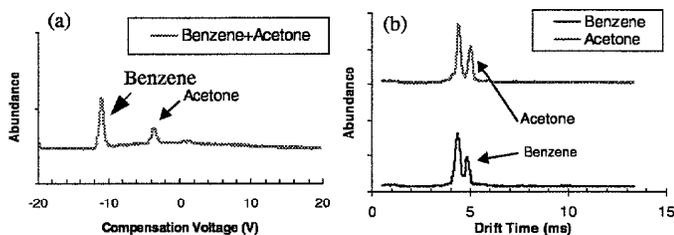
Figure 11(a) *RF-IMS spectra for a benzene acetone mixture.* ***(b)*** *Acetone and benzene spectra measured independently in a conventional TOF-IMS.*

Another exciting application for the RF-IMS, in addition to its use as a miniature detector, is as an atmospheric pressure chemical ionization (APCI) pre-filter for a mass spectrometer (MS). This pre-filtering capability should provide for enhanced discrimination of complex mixtures. The planar design and small size of the RF-IMS make it easy to interface to a mass spectrometer inlet. The RF-IMS also allows control of the ions and their environment during injection into the mass spectrometer inlet. The design of the RF-IMS enables the selective injection of only ions into the mass spectrometer, while ejecting neutrals separately out the RF-IMS side, Fig. 12. Both positive and negative ion species can be injected into the mass spectrometer simply by changing the polarity of the deflector electrode. The RF-IMS was interfaced to a TAGA 6000 tandem mass spectrometer as show in Fig. 12. A hole (0.7 mm diameter) was made in the detector electrode to facilitate the ion injection into the mass spectrometer. In addition to acting as a pre-filter for a mass spectrometer, this RF-IMS/MS configuration enables confirmation of the identities of ion species present in the RF-IMS spectral peaks.
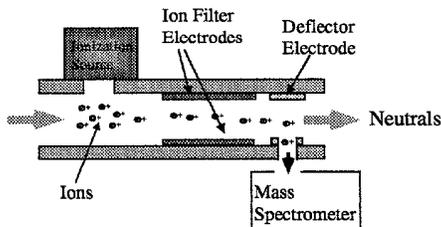
Figure 12. *Schematic showing the use of the RF-IMS as a pre-filter to a mass spectrometer.*

Figure 13a shows the measured mass spectra with the RF-IMS ion filter off. In this case all ions that are produced in the ionization region are unperturbed by the filter and can make it into the mass spectrometer. The ion species measured are acetone monomers, acetone dimers and some impurities from the carrier gas. With the RF-IMS filter on, and tuned to the characteristic compensation voltage for acetone –3.5 volts, the dominant MS peak is the acetone monomer peak all other peaks are suppressed, Fig. 13b, indicating that the RF-IMS is indeed working as a filter.
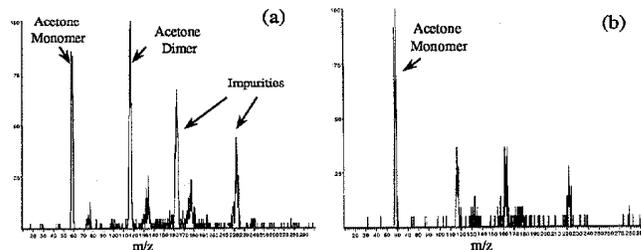
Figure 13(a) *RF-IMS/MS spectra with no fields applied to the ion filter (full scale = 4.8×10⁴ cps).* ***(b)*** *Spectra with the filter on and a compensation voltage of –3.5V applied (full scale = 2.2×10⁴ cps).*

CONCLUSIONS

A MEMS RF-IMS drift tube has been successfully fabricated and characterized. High spectrometer sensitivity and ability to resolve chemicals not separated in conventional TOF-IMS has been demonstrated. The MEMS RF-IMS enables the realization of miniature, low cost, high sensitivity, high reliability chemical detectors. The spectrometer has also been demonstrated as a pre-filter to a mass spectrometer. The RF-IMS/MS combination will allow better resolution of complex mixtures.

ACKNOWLEDGMENTS

REFERENCES

1. G. A. Eiceman, Z. Karpas, Ion Mobility Spectrometry, CRC Press, 1994.
2. C. S. Reese, S. J. Taraszewsky, "Near-Real-Time Analysis of Toxicologically Important Compounds Using The Volatile Organic Analyzer for the International Space Station", *Proceedings of the 8th International Conference on IMS*, August 1999, Buxton, U.K.
3. Z. Karpas, "Ion Mobility in Forensic Science", Encyclopedia of Analytical Chemistry, Wiley, New York, 2000.
4. J. Xu, W. B. Whitten, and J.M. Ramsey, "Studies of Miniature Ion Mobility Spectrometer", *7th International Conference on Ion Mobility Spectrometry*, Hilton Head Island, SC, 1998.
5. M. A. Butler, " Ion Trapping and Ion Mobility Spectroscopy", *7th International Conference on Ion Mobility Spectrometry*, Hilton Head Island, SC, 1998.
6. I. A. Buryakov, E. V. Krylov, E. G. Nazarov, *Int. J. Mass Spectrometry and Ion Processes 128*, 143- 148, 1993.
7. U. Kh Rasulev, E.G. Nazarov, V.V. Palitsin" Surface Ionization Gas-Analyzing Devices with Separation of Ions by Mobility," *4th International Workshop on IMS*, August 6-9,1995, Cambridge,U.K
8. E. A. Mason and E. W. McDaniel, Transport Properties of Ions in Gases, Wiley, New York, 1988.
9. E.V. Krylov, *Instruments and Experimental Techniques*, 40 (5), 628-631, 1997.

VISCOUS DRAG ON A LATERAL MICRO-RESONATOR: FAST 3-D FLUID SIMULATION AND MEASURED DATA

Wenjing Ye
School of Mechanical Engineering, Georgia Institute of Technology
Atlanta, GA 30339-0405

Xin Wang, Werner Hemmert, Dennis Freeman and Jacob White
EECS, Massachusetts Institute of Technology
Cambridge, MA 02139

ABSTRACT

Computing drag forces on geometrically complicated 3-D micromachined structures, such as an entire laterally moving comb resonator, is a challenging problem. The standard semi-analytic approach, based on assuming a regular flow patterns such as Couette or one-dimensional Stokes flow, results in a good rough estimate for the drag force but can be insufficiently accurate to correctly predict geometric sensitivities. Designers have not tried to improve the accuracy using numerical simulation, because traditional finite-element based 3-D fluid simulation is too computationally expensive to use for such complicated geometries. However, the recently developed FastStokes solver, based on the precorrected-FFT accelerated iterative methods, has made fluid analysis of an entire comb tractable. Herein, we compare measurements to simulation to show that drag force can be accurately predicted using 3-D simulation. In addition, detailed examination of the results from 3-D simulation indicates that there is significant contribution to drag due to sidewall forces, and therefore structure levitation will not increase quality factor as dramatically as one would expect from Couette flow analysis alone.

INTRODUCTION

Analysis of the dynamic behavior of micromachined devices packaged in the air or fluid requires the determination of damping effects on the devices due to viscous forces from the surrounding fluid. Because of the small spatial scale of the devices, the fluid usually satisfies Stokes' equation. Therefore, computing drag forces involves solving Stokes' equation on geometrically complicated 3-D micromachined structures, such as an entire laterally moving comb resonator, and this is a computationally challenging problem. Instead, semi-analytic approaches are often used, based on assuming a regular flow patterns such as Couette or one-dimensional Stokes flow [1] [2]. These approximations result in a good rough estimate for the drag force but can be insufficiently accurate to correctly predict geometric sensitivities. Designers have not tried to improve the accuracy using numerical simulation, because traditional finite-element based 3-D fluid simulation is too computationally expensive to

use for such complicated geometries. However, the recently developed FastStokes solver [3] [4], based on the precorrected-FFT accelerated iterative methods [5], has made fluid analysis of an entire comb tractable. In this paper, we compare measurements to simulation to show that drag force can be accurately predicted using 3-D simulation. In next section, a short description of Fast-Stokes is given. It is followed by the description of the experimental measurements of the resonant frequency and the quality factor of a micro-resonator. In section 3, numerical simulations of the drag force and the quality factor of the device are presented. Results are compared with the measured data, and good agreement is demonstrated. In addition, the computational results are examined in detail to show that sidewall forces are significant. Finally, in sections 4 and 5, we give conclusions and acknowledgements.

FASTSTOKES

FastStokes is a 3-D fast solver for both steady and unsteady Stokes' equations. The program combines a boundary element method for discretizing an integral formulation of Stokes' equation with precorrected-FFT accerated technique for solving the discretized equations.

The equivalent indirect boundary integral equations for unsteady Stokes' flow in the frequency domain are [6]

$$u_j(\mathbf{x_0}) = -\frac{1}{8\pi\mu} \int_S f_i(\mathbf{x}) G_{ij}(\hat{\mathbf{x}}) ds \qquad (1)$$

where S is the surface on the object in an infinite fluid, and the Greens function is given by

$$G_{ij}(\hat{\mathbf{x}}) = \frac{\delta_{ij}}{r} A(R) + \frac{\hat{x}_i \hat{x}_j}{r^3} B(R) \qquad (2)$$

$$A = 2e^{-R}(1 + \frac{1}{R} + \frac{1}{R^2}) - \frac{2}{R^2}$$

$$B = -2e^{-R}(1 + \frac{3}{R} + \frac{3}{R^2}) + \frac{6}{R^2}$$

$$R = \lambda r$$

In (1), $u_j(\mathbf{x_0})$ is the jth component of the velocity vector at the source point $\mathbf{x_0}$, $\mathbf{f}$ is the Stokeslet density function and the real part of $\mathbf{f}$ corresponds to the real part of the surface traction if there is a single moving object that rigidly translates in the fluid, ρ and μ are the density

and the viscosity of the fluid respectively, $\hat{x}_i$ is the ith component of the relative position vector between the source point and the field point, i.e. $\hat{x}_i = x_{0i} - x_i$, r is the length of the relative position vector ($r = |\mathbf{x_0} - \mathbf{x}|$), λ is the frequency parameter which is defined as $\lambda^2 = i\omega \frac{\mu d^2}{\rho}$, ω is the frequency of the fluid and d is the characteristic size of the object.

A piece-wise constant collocation scheme is used to solve (1) [7]. The surface of the object (device) is discretized into n small panels (see, for example, Figure 3), and the components of the traction forces are assumed to be constant on each panel. A system of equations for the panel unknowns is then derived by insisting the integral equations are satisfied at each panel centroid. The result is a system which relates the known velocity vector $\mathbf{u}$ to the vector of traction forces $\mathbf{f}$, as in

$$\left\{ \begin{array}{c} \mathbf{u}^1 \\ \mathbf{u}^2 \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{u}^n \end{array} \right\} = P(\omega) \left\{ \begin{array}{c} \mathbf{f}^1 \\ \mathbf{f}^2 \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{f}^n \end{array} \right\} \tag{3}$$

where $\mathbf{u}^i$ and $\mathbf{f}^i$ are the velocity vector and vector traction force for the ith panel, P is a $3n \times 3n$ matrix whose elements are given by

$$P_{kl}^{ij}(\omega) = \int_{\Delta j} \left[\frac{\delta_{kl} A(R)}{|\mathbf{x}^i - \mathbf{x}|} + \frac{(x_k^i - x_k)(x_l^i - x_l)}{|\mathbf{x}^i - \mathbf{x}|^3} B(R) \right] ds$$

$$k, l = 1, 2, 3. \tag{4}$$

Here, $\mathbf{x}^i$ denotes the centroid of the ith panel and Δj denotes the surface of the jth panel.

The linear system in (3) could be solved with Gaussian elimination, but then the cost of system solution would grow like the cube of the number of unknowns. For the comb example in Figure 3, there are more than 50,000 unknowns, and Gaussian elimination would require more than 10,000 gigaflops to solve. In addition, storing the matrix grows like the square of the number of unknowns, and for the comb example would require more that 20 gigabytes.

Instead, (3) is solved using a combination of an iterative method, complex GMRES [8], and a precorrected-FFT technique [5]. This combination of techniques results in an algorithm whose solution time and memory grows nearly linearly with the number of panels. For example, the 50,000 unknown problem in Figure 3 can be solved in 15 minutes on a workstation, and requires only one gigabyte of memory.

Experimental Measurement

The test structure was a surface-micromachined lateral resonator (Figure 1) fabricated using the MUMPS

Table 1: Resonator dimensions

finger gap	$2.88\mu m$
finger length	$39.96\mu m$
finger overlap	$19.44\mu m$
beam length	$151\mu m$
beam width	$1.1\mu m$
center plate	$54.9\mu m \times 19.26\mu m$
thickness	$1.96\mu m$
substrate gap	$2\mu m$
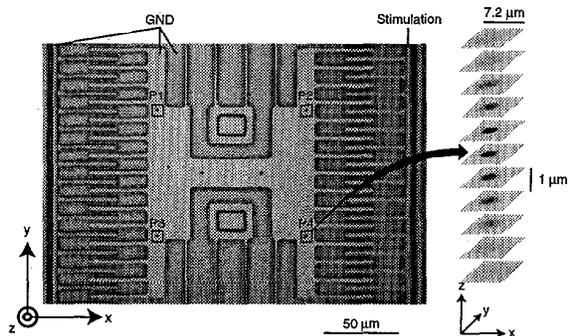
Figure 1: Two- and three-dimensional images of a lateral resonator. The left panel shows the central portion of a lateral resonator imaged near its best plane of focus. The resonator was excited to move by applying an electrical stimulus to the right comb drive while holding the left comb drive and shuttle at ground. Four regions of interest enclosing anti-stiction dimples are indicated by black boxes labeled P1, P2, P3, and P4. A 3D image of an anti-stiction dimple (right panel) is obtained by stacking a sequence of images from multiple planes of focus (2 μm spacing). The in-plane and out-of-plane scales differ, as shown by the scale bars, for clarity.

process at MCNC (now Cronos Integrated Microsystems Inc., Research Triangle Park, NC). Dimensions are shown in Table (1). The resonator was set into motion in air (atmospheric pressure) using an electrical stimulus to one comb drive. The magnitude and angle of the resulting motions were measured using computer microvision [9] and results are shown in Figure 2. Results were well fit by a second order system response with a resonant frequency of 19.2 kHz and a quality Q of 27.

Numerical Simulation

A second-order spring-mass-dashpot model described in equation (5) was generated to capture the dynamic behavior of the resonator.

$$m_{eff}\ddot{x} + c\dot{x} + kx = F_{actuator}, \tag{5}$$
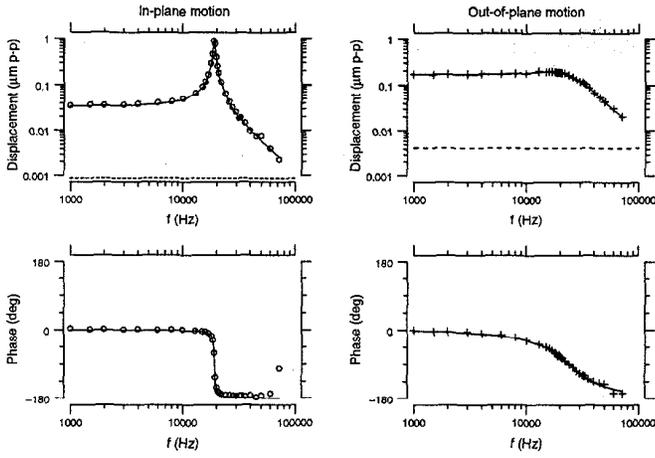
Figure 2: Frequency response for in-plane (left) and out-of-plane (right) motion of the anti-stiction dimple. Sinusoidal voltages (10 Vpp AC plus 50 V DC) were used to stimulate motions of the resonator. Motions were measured by analyzing a sequence of 3D images (see Figure 1) obtained using stroboscopic illumination at 8 evenly spaced times during the stimulus period. Circles indicate the magnitude and phase of motions measured at 32 frequencies form 1 to 72 kHz. Best-fitting second-order resonance curves are shown by the solid lines.

where x is the displacement of the beams and $F_{actuator}$ is the force given by the comb fingers due to the electric field.

In equation (5), the effective mass m_{eff} was determined using Rayleigh's method [10]. It was defined as

$$m_{eff} = m_p + \frac{1}{4}m_c + \frac{12}{35}m_b, \tag{6}$$

where m_p, m_c and m_b are the masses of the center plate, moving comb fingers and the beams respectively. The stiffness of the beams k was computed from the effective mass and measured resonant frequency:

$$f = \frac{1}{2\pi}\sqrt{\frac{k}{m_{eff}}}. \tag{7}$$

For the studied resonator, the effective mass and the stiffness of the spring are $3.5 \times 10^{-11}kg$ and $0.511\frac{N}{m}$.

The dashpot damping coefficient was computed by calculating the viscous drag force and then dividing by the velocity. The drag force was computed at the resonant frequency $f = 19200$ HZ. Two models were used: the Couette flow model (including damping due to both structure-substrate interaction and finger-finger interaction) as in [1], and then the 3-D Stokes model, Fast-Stokes. The surface mesh used for FastStokes is shown in Figure (3). A convergence study was performed to ensure the accuracy of the simulation results. Table (2) shows the simulated drag force on the resonator for different discretizations. The table shows that using 23,424
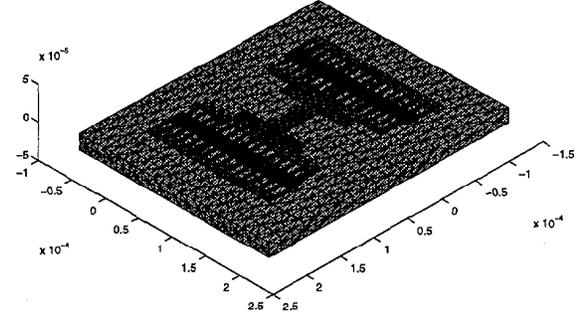
Figure 3: The meshed resonator with substrate, total 16575 panels

Table 2: Viscous drag force on the resonator

Number of Panels	FastStokes	Couette
10048	165.04(nN)	95.02
16576	168.34(nN)	
23424	169.45(nN)	

panels results in a drag force that is within one percent of the exact solution to the Stokes' equation.

The quality factor Q of the resonator was determined from

$$Q = \sqrt{\frac{km_{eff}}{c}}. \tag{8}$$

Results are shown in Table (3), together with the measured Q. As the table shows, the Couette flow model overpredicts the Q by a factor of nearly two, whereas the FastStoke's based result is within 10 percent. In Figure (4), we show the drag force distribution on the comb. As can be seen from the figure, there are significant contributions to the drag force on the sides of the comb. These contributions are mostly caused by the pressure gradients existing in the flow field and the edge effects due to the finite size of the device. Based on an infinite-plate assumption, both Couette-model and 1-D Stokes-model fail to capture this component of the drag. Integration of the various components of the simulated drag force shows that side forces account for 33 percent of the drag, top force contributes (Stokes-type damping

Table 3: Q factor of the resonator

	FastStokes	Couette	Measurement
Q factor	25.16	44.59	27
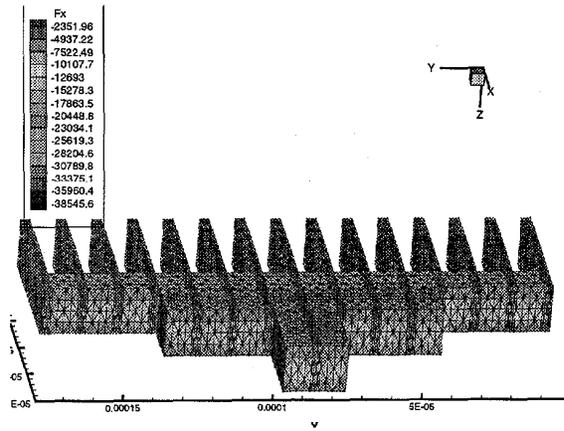
Figure 4: Drag force distribution on the resonator, bottom (substrate-side) view

in the ambient fluid) another 12 percent, and the bottom force (Couette-type damping) is 55 percent. The implication is that levitating the comb will *not* increase the Q as much as a Couette flow model predicts.

CONCLUSIONS

In this paper, we compare measurements to simulation to show that viscous drag force can be accurately predicted using 3-D simulation. A lateral microresonator ocsillating in the air was studied. The drag force was simulated both using FastStokes, a precorrected-FFT accelerated unsteady Stokes solver and a simple Couette flow model. Quality factor of the device was then determined from the simulated results and compared with the experimental data. A good agreement was obtained between the 3-D results with the experimental data (error is within 10 percent), whereas the Couette model overpredicts the Q by a factor of nearly two. In addition, the detailed examination of the results from 3-D simulation indicates that there is significant contribution to drag due to sidewall forces, and therefore structure levitation will not increase quality factor as dramatically as one would expect from Couette flow analysis alone.

ACKNOWLEDGMENTS

REFERENCES

[1] W. C. Tang, T.-C. H. Nguyen, M. W. Judy and R. T. Howe, "Electrostatic-comb drive of lateral polysil-icon resonators," *Sensors and Actuators*, vol. A21 - 23, pp. 328-381, 1990.

[2] Y.-H. Cho, B.-M. Kwak, A. P. Pisano and R. T. Howe, "Viscous energy dissipation in laterally oscillating planar microstructures: a theoretical and experimental study." *Proceedings. IEEE. Micro Electro Mechanical Systems. An Investigation of Micro Structures, Sensors, Actuators, Machines and Systems (Cat. No.93CH3265-6).* IEEE. pp.93-8. New York, NY, USA, 1993.

[3] N. R. Aluru and J. White, "A Fast Integral Equation Technique for Analysis of Microflow Sensors Based on Drag Force Calculations," *Proc. MSM*, pp. 283-286, Santa Clara, April 1998.

[4] W. Ye, J. Kanapka, X. Wang, J. White, "Efficiency and Accuracy Improvements for FastStokes, A Precorrected-FFT Accelerated 3-D Stokes Solver", *Proc. of MSM*, pp. 502-505, Puerto Rico, 1999.

[5] J. R. Phillips and J. K. White, "A Precorrected-FFT method for Electrostatic Analysis of Complicated 3-D Structures," *IEEE Trans. on Computer-Aided Design*, October 1997, Vol. 16, No. 10, pp. 1059-1072.

[6] C. Pozrikidis, *Boundary integral and singularity methods for linearized viscous flow*, Cambridge University Press, Cambridge, 1992.

[7] P. K. Banerjee, *The Boundary Element Methods in Engineering*, McGraw-Hill Book Company, England, 1981.

[8] Y. Saad and M. Schultz, 'GMRES: A generalized minimal residual algorithm for solving symmetric linear systems,' *SIAM J. Sci. Statis. Comput.*, 7, 856-869, 1986.

[9] D. M. Freeman, A. J. Aranyosi, M. J. Gordon, and S. S. Hong, "Multidimensional motion analysis of MEMS using computer microvision," 1998 Solid-State Sensor and Actuator Workshop, June 8-11, pp. 150-155. 1998.

[10] W. C. Tang, M. G. Lim and R. T. Howe, "Electrostatic comb drive levitation and contral method," *Journal of MEMS*, vol. 1, No. 4, pp. 170 - 177, 1992.

PREDICTIVE DESIGN OF REVERSE INJECTION MECHANISM FOR ELECTROKINETIC DNA SAMPLE INJECTION

Manish Deshpande, Ken B. Greiner, John West and John R. Gilbert
Microcosm Technologies
215 First Street, Cambridge MA 02142, USA

Luc Bousse and Abdel Minalla
Caliper Technologies Corp.
605 Fairchild Drive, Mountain View, CA 94043

ABSTRACT

A novel injector designed by CAD analyses is presented in this paper. The injection mechanism uses a pull-back inserted between the pinch and the switch stages of the typical cross injector to create a band that is more symmetric and narrower than the typical injector. The resulting separation efficiency is greatly enhanced as a result of the more optimal band shape. The injector was designed using a validated CAD tool, and verified subsequently by experiments. The resulting separation efficiency achieved was significantly greater than current injectors.

INTRODUCTION

Electrokinetic microfluidic microsystems are powerful analytical tools for many applications, such as nucleic acid analysis, enzyme assays, and immunoassays [1-4]. Such systems have gained considerable importance as components in micron-scale integrated chemical/biochemical analysis or synthesis systems, also referred to as lab-on-a-chip. The basic "unit process" operations in these systems are sample injection, mixing, chemical reaction or modification, separation, and detection. Assembling a system of many "unit process" nodes requires one or more transport mechanisms to move sample and reagents through the "wires" of the system. Many of these systems rely on electrokinetic physics as their transport mechanism, although pressure and pneumatic applications have also been demonstrated. Complicated relationships exist between the microchannel geometries, the conditions under which the devices operate, and the behavior of the multi-component fluids transported in these channels. In the past researchers have been forced to use costly trial and error methods to understand and design such microfluidic systems.

CAD tools can be a valuable aid in the design of microfluidic systems. Numerical analyses provide significant insight into the fluid mechanics in these systems. They allow the extraction of material and flow properties that are generally not well documented, or that vary from application to application or from one manufacturing technology to another. Furthermore such tools help the designer to explore a much larger space of designs than is easily available from experiment, and do so in a quantitative way that enables the extraction of key parameters for improved or optimal operation of common microchemical system components.

Simulations of electrokinetic flows have been reported in the literature for both electrophoretic [5,6] and electroosmotic [7,8] flows. In [6] and [8], simulations of pinched injection have been reported in channel intersections, demonstrating the application of electrical fields to position the species plug in the intersection, prior to switching. These analyses are 2D steady-state analyses with fixed field boundary conditions. Three-dimensional simulations of pinched injection have also been reported by these authors in previous work [9]. In all these cases, the analyses were aimed at demonstrating numerical capabilities and an understanding of the fundamental physics prevalent in the device. Agreement with experiment reported in these analyses is generally good.

The capabilities of CAD tools can only be fully realized when they are applied in predictive design of a device that improves or optimizes its performance. In this paper, we present the predictive use of simulation to design an improved injector, followed by experimental confirmation of that design and the use of the improved injector to achieve the fastest separations of DNA oligos currently achievable.

Microfluidic networks have a distinct advantage over conventional electrophoresis system in the area of rapid separations. Microfluidic systems have demonstrated separations in the order of seconds in comparison to several minutes in conventional systems [3,4]. A critical parameter in achieving rapid separations is the ability to inject a narrow plug into the separation column. Typically, microfluidic system achieve this by using pinched injection in a cross injection, as shown in Figure 2. The drive potential is supplemented
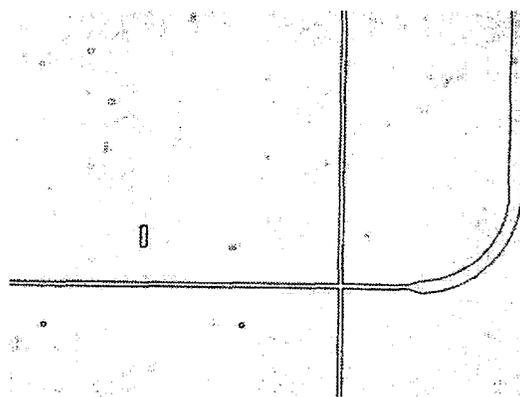

Figure 1. Photograph of a NS95 microchip used for the reverse injection experiments.
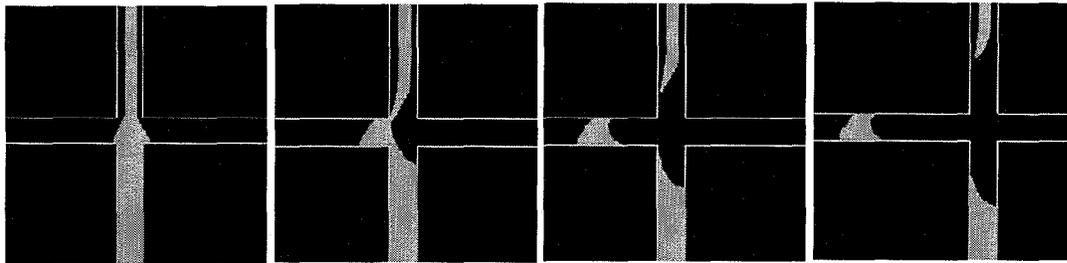
Figure 2. Typical Electrokinetic injection process showing the pinching and the switching process.

by fields in the transverse channels to shape the plug in the intersection prior to switching into the separation column as seen in the figure. The plug thus injected can be of the order of the channel dimension or smaller. As shown in Figure 2 the plug is broadened in a trapezoidal shape. This broadening limits the possible resolution of the following separation column, causing the separation column to be longer than necessary. Designing injection to achieve a plug that is symmetric and narrower than the channel geometries is highly desirable.

Here, we will use CAD to focus on this aspect of the injection in an attempt to improve it. We will begin with a short description of the numerical tools followed by details of the simulation. Following that we will present experimental results that corroborate our findings, resulting in an improved injector, that demonstrates faster separations at lower voltages than conventional injectors.

NUMERICAL FORMULATION

The basic equations describing the fluid motion are the Navier-Stokes equations with appropriate electromigratory flux terms to represent the effect of the applied electric field on the charged species. The basis for electrophoresis is the differential migration of the charged species ions relative to the carrier molecules under the application of the external field. The differential migration is primarily an effect of the difference in the net charge between the solvent and solute ions, although frictional effects may also have some relevance. The migration velocity of the charged species can be expressed in terms of the applied field strength as $V_{ep} = \mu_{ep} E$, where μ_{ep} is the electrophoretic mobility of the ion in the carrier species. It is important to note that in most cases the carrier does not move under electrophoresis.

The motion of a charged species in the electric field can be determined by incorporating an electrokinetic transport mechanism in the species equation. The transport of the species is through the combined effect of the electroosmotic motion of the carrier fluid and the electrophoretic transport of the species under the effect of the applied electric field.

The numerical analyses presented above are derived under the following assumptions –

- Neutral Carrier: The carrier fluid is assumed to be electroneutral everywhere, except within the double layer.
- Dilute Sample: The carrier fluid is assumed to be predominant in calculating the physical properties of the fluid.

- Uncoupled Transport: Individual *sample* species do not affect each other as to their diffusion or mobilities.
- No Chemical Reactions: The charged sample species are assumed to be fully ionized in the mixture, and do not react with each other.

The above assumptions allow the density of the mixture to be assumed constant, reducing the problem to the incompressible form. The momentum and species equations are decoupled and can be solved separately.

The modeling of electrokinetic effects is incorporated into the FlumeCAD system. FlumeCAD is an integrated design environment consisting of 3D design, modeling and simulation software tools, which enable the creation and analysis of complex microfluidic devices. Inherent in the design flow implemented in FlumeCAD is the ability to translate from a layout and process view of the device to a solid model and to continue to a 3D device model allowing simulations that characterize the various physical phenomena present in the device. The numerical solution uses a three-dimensional finite element based engine as the back-end solver for the analyses.
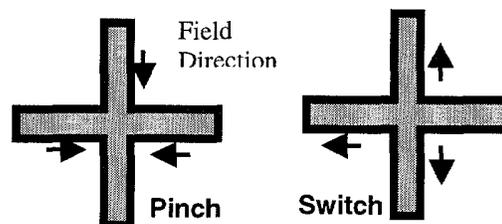

Figure 3. Schematic of the typical two phase injection
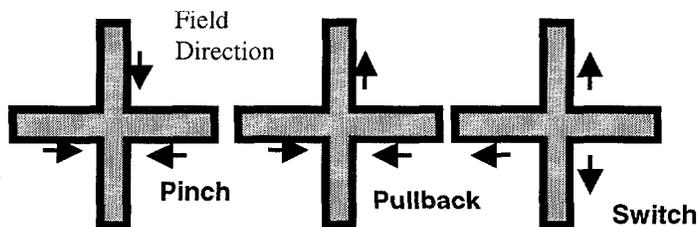
*Figure 4. Schematic of Reverse Injection including pull-back.*
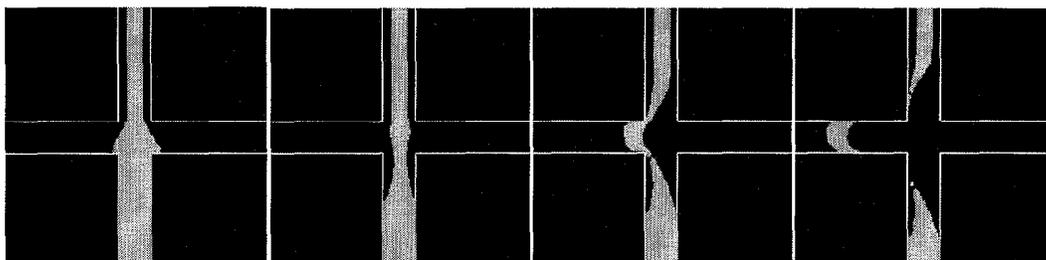
Figure 4. *Reverse Injection process showing pull-back and subsequent switching into separation column.*
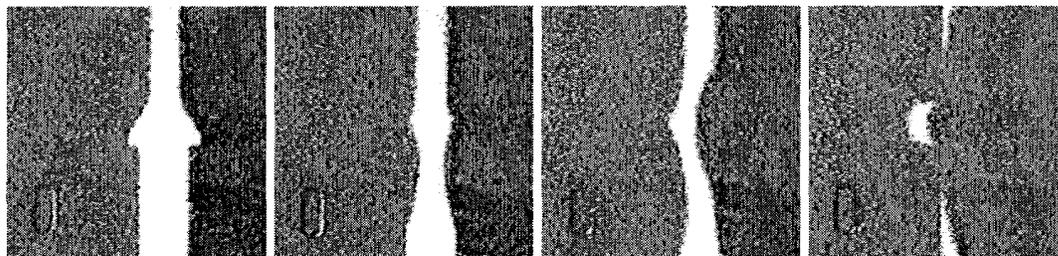
Figure 5. *Experimental verification of reverse injection predicted by numerical simulations above.*

DESIGN ANALYSIS

The design analyses were conducted on a cross injector. The numerical tools were first validated by comparing against experiment for the classic two-phase pinched injection. A schematic of this injection mechanism is shown in Figure 3 – corresponding simulation results are shown in Figure 2. The injection consists of a "pinch" – fields in the transverse channels supplementing the drive current – followed by a "switch" – sweeping the sample into the separation column along with current towards both top and bottom to separate the sample plug and prevent leakage. The simulations presented here were compared against corresponding experimental observations in [9] and show good qualitative agreement, indicating that the simulation tools capture all the relevant physics in the problem quite well.

As Figure 2 shows the switched plug in the separation column is trapezoidal in shape. The shape in the separation column is a consequence of the shape of the pinch in the intersection between the two channels. Since the sample is positioned in the intersection by a pinch, the shape in the intersection is trapezoidal, with the base of the trapezoid being wider than the characteristic dimension of the intersection. As the trapezoidal bands move downstream the individual fragments separate out based on their electrophoretic mobilities, and broaden based on their diffusivities. The effective band width in the separation column is determined by the width of the trapezoid base, which limits the separation efficiency.

A better injection mechanism is required to create a better sample plug in the intersection prior to separation. This can be achieved by alterations in the geometry or in the switching field sequence. One such example, showing a six-port switch, where the pinch and switch are carried out in different intersections, was presented in [10] and results in a narrower band in the separation column. A second approach termed "reverse injection", presented here, uses a three-phase switching sequence. The schematic of this sequence is shown in Figure 4. A "pull-back" phase is added between the pinch and the switch. In the creating of the pinch, the sample is positioned in the intersection by the pinching fields. Downstream of the intersection the band is significantly narrower than the width of the intersection, and is dependent on the strength of the pinching field. Adding a pull-back phase momentarily reverses the field in the drive channel, while maintaining the pinching fields as seen in the schematic. The pull-back causes the downstream band to retract into the separation column creating a significantly narrower band at the entrance of the separation column. Additionally the downstream band is also symmetric since it is not modified by the effect of the pinch. The usual switching step then follows to drive the sample into the separation column.

Simulations showing the effect of the "pull-back" are presented in Figure 5. The effect on the band width and shape in the separation column is dramatic – the band is now both narrower and straighter than the conventional injector case in Figure 2. The results indicate that we can achieve DNA bands that are symmetric and narrower than the underlying channel geometry – consequently the resulting separation efficiency is expected to improve significantly as a result of the band shape.

EXPERIMENTAL VERIFICATION

Following the CAD analysis, experiments were conducted to verify the predictions presented in the previous section. The separations were carried out in soda-lime glass chips with channels 10 μm deep and 30 μm wide, filled with a 6% polyacrylamide-based separation matrix. The buffer was 100 mM TAPS-Tris. For imaging
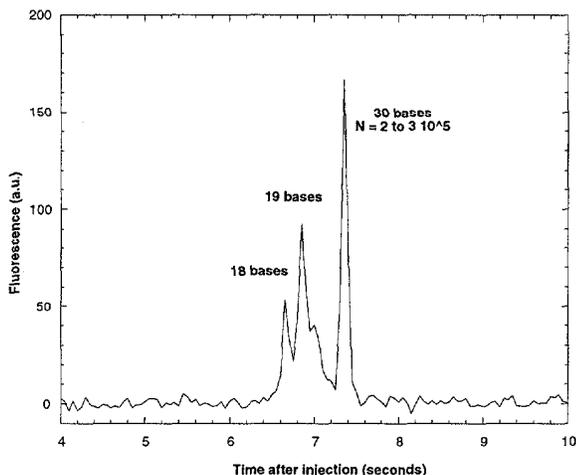
Figure 6. *Ultra-high speed separations using reverse injection process.*

studies, the samples were single-stranded DNA oligonucleotides at a concentration of 20 µM, fluorescently labeled at fluorescein wavelengths. Images were recorded on a Hammamatsu intensified CCD camera. For separations, similar oligonucleotides were used, but labeled at red wavelengths with Cy-5 dye, and at a concentration of 100 nM. Detection was carried out using a red diode laser, on a fluorescent microscope system as described earlier [3]. Experimental images corresponding to the reverse injection case are shown in Figure 6. The experimental results agree very well with the numerical predictions from Figure 5. The details of the sample behaviour including the shape in the intersection prior to switching, and the shape of the subsequently injected band, are correctly captured by the simulations. Separations from this injector are presented in Figure 6. The narrow band created by reverse injection enabled the demonstration of 7 second DNA separations in a 0.5cm long on-chip column, using only 150V (300V/cm). These are the fastest separations accomplished to date and are a result of the improved injector design emerging out of the CAD analyses.

CONCLUSIONS

The application of CAD in the design of microfluidic chip components is demonstrated in this paper, using the cross injector. In typical separations, the efficiency is limited by the width of the band in the separation column, which in turn is limited by the width of the band positioned in the intersection. Using a validated CAD tool, we have improved the injection characteristics by inserting a pull-back between the pinch and the switch. This injection, called "reverse injection" retracts the narrower and symmetric band downstream of the intersection back into the intersection, and subsequently into the separation column. The injection characteristics have been verified by experiment and have yielded faster separations using lower fields than have been accomplished hitherto.

The primary conclusion of this work is the successful application of CAD to predict and improve the performance of microfluidic components. CAD also has the advantage of being

able to explore a much wider parameter space to design these improved components, and hence should serve as an important component of the design process.

REFERENCES

[1] D.J. Harrison, K. Fluri, K Seiler, Z. Fan, C.S. Effenhauser and A. Manz, "Micromachining a miniaturized capillary electrophoresis-based chemical analysis system on a chip," Science, **261** (1993) 895-897.

[2] S.C. Jacobson, R. Hergenroder, L.B. Koutny, R.J. Warmack and J.M. Ramsey, "Effects of Injection Schemes and Column Geometry on the Performance of Microchip Electrophoresis Devices," *Anal. Chem.,* **66**(1994), 1107.

[3] Luc Bousse, Bob Dubrow and Kathi Ulfelder, "High Performance DNA Separations in Microchip Electrophoresis Systems," u-TAS '98, 271.

[4] A.T. Woolley and R.A. Mathies, "Ultra High Speed DNA Sequencing using capillary electrophoresis chips," *Anal. Chem.,* **65** (1995), 3676.

[5] X. C. Qui, L. Hu, J. H. Masliyah, and D. J. Harrison, "Understanding fluid mechanics within electrokinetically pumped microfluidic chips", *1997 International Conference on Solid-State Sensors and Actuators,* Chicago, IL (1997).

[6] P.M. St. John, et al, "Metrology and Simulation of Chemical Transport in Microchannels", *Solid State Sensors and Actuators Workshop,* Hilton Head, SC (1998).

[7] N. A. Patankar and H. H. Hu, "Numerical Simulation of Electroosmotic Flow", *Anal. Chem.,* **70**, (1998).1870-81.

[8] Sergey V. Ermakov, S.C. Jacobson, and J.M. Ramsey, "Computer Simulations for Microchip Electrophoresis," u-TAS '98, 149.

[9] Bousse, L., Minalla, A., Deshpande, M., Greiner, K., and Gilbert, J., "Optimization of Sample Injection Components in Electrokinetic Microfluidic Systems," MEMS '99, Orlando, FL (1999).

[10] Deshpande, M., Greiner, K., West, J., Gilbert, J., Bousse, L., and Minalla, A., "Novel Designs for Electrokinetic Injection in µTAS," Proc. µTAS 2000, Enschede, Netherlands (2000), in print.

[11] Bousse, L., "Electrokinetic Microfluidic Systems," SPIE Microfluidic Devices and Systems II, Santa Clara, CA (1999).

A LOW DISPERSION TURN FOR MINIATURIZED ELECTROPHORESIS

Joshua I. Molho, Amy E. Herr, Bruce P. Mosier,
Juan G. Santiago, and Thomas W. Kenny
Mechanical Engineering Department, Stanford University
Stanford, CA 94305-4021

Reid A. Brennen and Gary B. Gordon
Agilent Laboratories, Agilent Technologies
Palo Alto, CA 94304-1392

ABSTRACT

Many chip-based microcolumn separation systems require serpentine channels to obtain longer separation lengths within a compact area. However, analyte bands traveling through curved channels experience increased dispersion and thus, serpentine channels reduce the intended benefit of the increased channel length. We have used simulation tools to predict this increased band dispersion and then to design a corner geometry that causes less dispersion. Experiments using bleached- and caged fluorescence visualization were conducted to verify the proposed turn designs. We demonstrate experimentally that the new corner design greatly reduces dispersion of analyte bands. Quantitative and qualitative comparisons can be made between the simulations and experiments, and these results suggest that the simulations can be used effectively as design tools.

INTRODUCTION

Pioneered in the early 1990's [1], on-chip capillary electrophoresis (CE) remains an important separation technique for chemical/biochemical microfluidic devices. This technique separates a mixture of unknown analytes based on their mobilities in an applied electric field. The basic scheme involves introducing a small volume of mixed, unknown analytes into a separation channel. An electric field is applied along the length of the separation channel and the various analytes move at different speeds based on their mobilites. After some time, the analytes separate into individual bands, as depicted schematically in Figure 1. The arrival of individual analyte bands can be detected downstream, often using a single point fluorescence or absorption sensor. The presence of a particular analyte in the unknown mixture can be determined by using markers injected with the sample or by first calibrating the system with a mixture of known analytes.

Several studies have demonstrated the potential benefits of miniaturizing capillary electrophoresis on microfabricated chips. Among these benefits are portability, reduced reagent use, and increased opportunities for parallelism [2]. However, since separation efficiency increases with the length of the separation channel, it is desirable to use longer microchannels within a small area, thus requiring turns in the microchannel. Recently, Culbertson *et al.* [3] have shown that such turns increase the dispersion of samples flowing in the microchannels and thus cancel the benefit of the additional length.

THEORY

Consider the electrophoretic separation of a fluorescently-tagged sample as depicted in Figure 1, where the initially mixed sample has separated into its two components some time after the application of the axial electric field. The two analyte bands are separated by a distance ΔL and the Gaussian bands have a standard deviation of σ. We can define a resolution as:

$$resolution = \frac{\Delta L}{\sigma} \sim \frac{E \cdot \Delta\mu \cdot t}{\sqrt{D \cdot t}} \qquad (1)$$

where $\Delta\mu$ is the difference between the electrophoretic mobilities of the two analytes, D is the diffusion coefficient and t is the time of separation. In the common case where different analytes have different diffusion coefficients, the largest σ can be used to define the resolution.
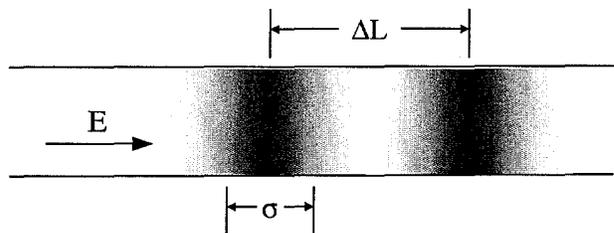
Figure 1. Schematic of an electrophoretic separation. Initially, the two bands shown above were mixed together. After the application of an electric field, the bands move from left to right and separate based on the difference in their electrophoretic mobilities.

The bands will travel a distance L proportional to the separation time and therefore, we can also state that

$$\frac{\Delta L}{\sigma} \sim E \cdot \Delta\mu \cdot \sqrt{L} \qquad (2)$$

Eq. 2 shows that the resolution will vary linearly with the field strength and mobility difference. The mobility difference is a property of the analytes and is not controlled. It is often advantageous to increase the field strength, but in some cases, further increases in the field strength will cause excessive Joule heating [4]. In such thermally limited cases, increasing the separation length will enhance the separation resolution. However, serpentine channels are required to increase separation length within a compact area.

Unfortunately, not all serpentine channels will increase separation resolution. Culbertson *et al.* [3] have presented a study of dispersion caused by constant radius turns where the radius of

curvature is larger than the width of the channel. The general effect is illustrated in Figure 2.
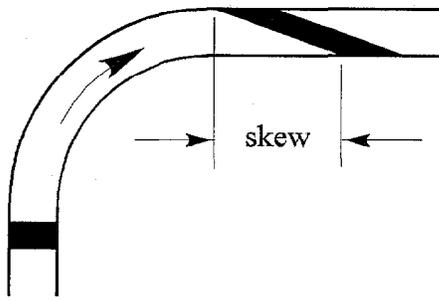
Figure 2. Effect of 90 degree, constant radius turn on an analyte band. The band is "skewed" as it travels through the turn with molecules near the inside of the turn racing ahead of the molecules near the outside of the turn. This skewing of the analyte band has been called the race-track effect.

The skew seen in Figure 2 is caused by the fact that the pathline along the inner portion of the turn is shorter than the pathline along the outer portion of the turn, and that the electric field is stronger along the inside of the turn. Using the approximate expression given by Culbertson et al. [3] the skew, as defined in Figure 2 and in the limit of no diffusion, is equal to

$$skew = 2\theta w \qquad (3)$$

where w is the width of the channel and θ is the angle completed by the turn (in radians). To address how diffusion will affect the amount of skew caused by a turn, we refer to the two-dimensional, non-dimensionalized advection-diffusion equation (Eq. 4) where u', c', x', and y' are the normalized velocity, concentration, direction along the channel and direction across the channel, respectively.

$$\underbrace{\frac{\partial c'}{\partial t'} + u'\frac{\partial c'}{\partial x'}}_{advection} = \frac{1}{Pe_w}\left[\underbrace{\frac{w}{L}\left(\frac{\partial^2 c'}{\partial x'^2}\right)}_{\substack{axial\\diffusion}} + \underbrace{\frac{L}{w}\left(\frac{\partial^2 c'}{\partial y'^2}\right)}_{\substack{transverse\\diffusion}}\right] \qquad (4)$$

The Peclet number, Pe_w is a non-dimensional parameter that characterizes the relative importance of diffusion of analytes compared to advection of analytes; $Pe = U w / D$ with U being the average electrokinetic velocity around the turn, w the width of the channel and D the diffusivity of the analyte traveling through the channel. L is a length scale in the axial direction along the channel. In this study, L is the axial length of the turn, which is on the order of 10 times w.

When Pe_w is very small, the diffusion terms on the right side of Eq. 4 will dominate the evolution of the shape of the analyte band. In this paper, L / w is on the order of 10, so we see that when Pe_w is greater than 100, the advection term dominates the shape of the analyte band. Designing the geometry of a compensating turn can be interpreted as an attempt to control the final shape of the analyte band through u' which appears in the advection term. Therefore, a compensating corner is most relevant and most effective when considering the case of large Pe_w (large molecules and/or large electrokinetic velocity). Note that when Pe_w is order 10, the advection and the transverse diffusion terms are on the same order, and we cannot state, in general, that a corner designed for high Pe_w will be equally effective in this regime. In our approach we will design the corners to perform best at high Pe_w, because as stated above, the modified u' will have the most effect in this regime. Furthermore, high Pe_w is representative of most on-chip separation systems. For example, a chip based separation system with 100 μm wide channels, separation velocities on the order of 100 μm/sec, and analytes with diffusivities on the order of 10^{-10} m^2/sec (e.g. serum albumin) would have a Pe_w equal to 1000. In the work presented in this paper, we have attempted to perform all the experiments and simulations in the limit high Pe_w on the order of 100 or higher.

The model for skew presented above does not directly account for the effect of electroosmosis in the turn. Most microchannel surfaces form a charged electric double layer when exposed to ionic solutions. Electroosmosis describes the bulk motion of fluid that results due to electrostatic body forces in the double layer when an axial electric field is applied [5]. In general, solving for the electroosmotic velocity requires solution of the electric field and charge density in the channel, followed by a solution to the Navier-Stokes equation.

Cummings et al. [6] suggest that in the case of uniform zeta potential, the electroosmotic velocity is everywhere parallel to the electric field. This result is valid when: the electric double-layers are thin compared to the channel width, no current passes through the walls of the channel, and the flow is parallel to the channel at any inlet or outlet to the region of interest. Utilizing this simplification, we propose to model the flow field by considering the electrophoretic component which, away from the electrodes, is similar to the electric field and the electroosmotic flow field. We demonstrate the validity of this approximation by using our flow measurement techniques.

SIMULATION APPROACH

Using the approximations described above, we have employed simulation tools (MEMCAD 4.6, Microcosm Technologies, Inc.) to design a low dispersion turn.
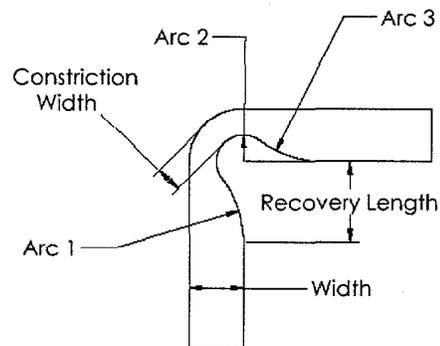

Figure 3. Parameterization of turn geometry. The corner is fully defined by the ratios of constriction to width and recovery length to width. Arcs 1, 2 and 3 are all tangent to each other and Arcs 1 and 3 are tangent to the sides of the straight channels away from the corner. These final tangent conditions fully define the geometry once the geometric ratios have been defined.

To compensate for the race-track effect, we propose to lengthen the pathline along the inner portion of the curve and to reduce transverse field gradients, using a modified turn as depicted in Figure 3. The corner design, as shown in Figure 3, is fully defined by specifying two ratios: the ratio of the width to recovery length and the ratio of the constriction width to the full channel width. This simple parameterization enables optimization of the corner design.

The simulated electric field strength in a compensating corner is compared to that of a constant radius corner in Figure 4. The compensating corner has a lengthened pathline along the inside of the turn and furthermore, a charged molecule that approaches the turn along the inside edge remains in a region of lower electric field strength (note the tilt of the contour lines) before reaching the higher field strength at the constriction. These two effects combine to help equalize the travel times of particles on the inside and outside edges of the corner, resulting in less skew.

Once the geometry is defined, the final shape of the analyte band is fully specified by the Peclet number. Recall that high Pe_w number corresponds to large molecules and/or high field strengths, and low Pe_w number corresponds to small molecules and/or low field strengths.
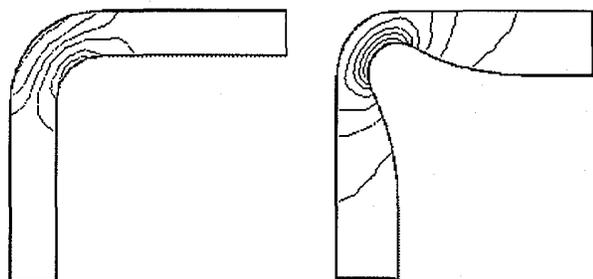

Figure 4. Comparison of electric field contours in a constant radius (left) and a compensating corner (right). Contours of constant electric field strength are shown for the two geometries.

To determine the final shape of an analyte band after it has passed through a corner design, the numerical solver performs the following steps. First, the electric field inside the channel is solved assuming electrically insulating sidewalls, and the inlet and outlet of the channel are held at constant potential. The straight sections of the channel before and after the turn must be long enough that imposing a uniform potential at the inlet and outlet will not perturb the solution (typically 3 to 4 channel widths). Next, the electrokinetic velocity in the channel can be determined by simulating only electrophoresis after appropriately modifying the mobility of the tracer dye to include the effects of electroosmosis. Finally, an initially Gaussian distribution of dye is numerically injected near the inlet of the corner and then tracked as it travels through the corner.

EXPERIMENTAL SETUP

To validate our design process and investigate the performance of our designs, channels were fabricated in polymethyl methacrylate (PMMA) using a 125-micron diameter end-mill. A PMMA coverslip was thermally bonded at 120 °C onto the machined piece to complete the fluidic channels. Figure 5 shows the custom fixture that allows for pressure filling of liquids into the channels and that provides convenient wells for inserting electrodes.

The dispersion caused by the compensating corner was observed experimentally using bleached fluorescence [7] and caged fluorescence [8, 9] techniques as depicted in Figure 6. Both techniques provide an "optical injection" of a tracer. The caged fluorescence technique relies on caged dyes that are not fluorescent until uncaged with a UV pulse. The channels are filled with
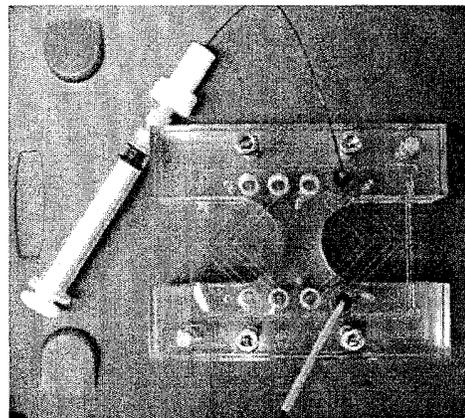
Figure 5. Custom fixture used during experimentation. This photograph shows the fixture that was used to position the microchannel chip and to provide seals for getting liquids into the chip. The eight channels can be seen as the faint traces on the 2 X 3 inch PMMA chip held in the fixture. The cutouts in the fixture provided clearance for the microscope objective used to image the corner designs.

a 400 μM solution of caged fluorescein (Molecular Probes, Inc.) dissolved in de-ionized water. The uncaging is performed using a single pulse from a frequency tripled Nd:YAG laser (355 nm, 500 μJ) focused to an approximately 100 μm by 2 mm area. For the bleaching technique, the channels are filled with a 10 μM solution of a 2 MDalton dextran-fluorescein conjugate (Molecular Probes, Inc.) dissolved in de-ionized water. A 0.5 Watt argon ion continuous wave laser beam is focused through a 10X microscope objective (NA = 0.3) and translated across the channel to create a bleached timeline.
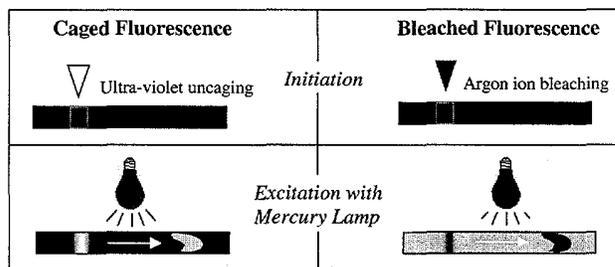
Figure 6. Schematic of caged (left) and bleached (right) fluorescence visualization techniques. The top row describes the "optical injection" stage of both techniques. The bottom row depicts the interrogation phase of the techniques when the marked region of the flow is tracked using epi-fluorescent microscopy.

After initiation, the marked region is excited with a Mercury lamp and images are collected at 10X magnification with an epi-fluorescent microscope and a 652 x 488, video rate, back-illuminated CCD camera (PixelVision Inc.).

The bleaching method has the advantage of using any fluorescent dye, but it requires a relatively long time to bleach a region (order of tens to hundreds of milliseconds). The caged

method requires custom dyes, but the UV pulse is short and therefore the uncaging can be performed while the electric field is applied.

RESULTS AND DISCUSSION

To evaluate the effectiveness of each design we use two quantities to describe the dispersion caused by the corner: the skew parameter and the full width at half maximum (FWHM) of the band. The FWHM of the band gives some indication of the dispersion caused by the corner by convolving a theoretical line detector with the band after it has traveled through the turn.
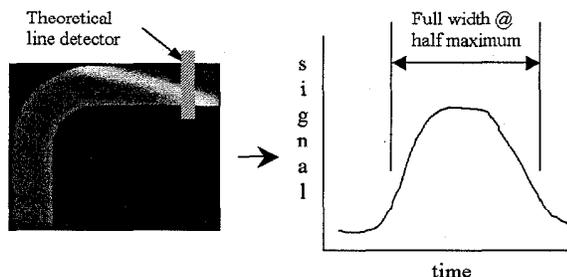
Figure 7. Determination of FWHM. The FWHM calculation as shown here simulates convolving the perturbed analyte band shape with an ideal line detector.

The theoretical line detector can be approximated by binning the rows of the analyte band at a given time and then calculating the FWHM value from the resulting graph. This process is depicted in Figure 7.

The skew gives direct indication of the stretching caused by the corner. For turn designs with an arbitrary geometry, the stretching of the analyte band may not be linear as it is in Figure 2. A large skew suggests that the band will broaden more than the FWHM indicates, since diffusion will eventually force the analyte band back to a Gaussian plug.
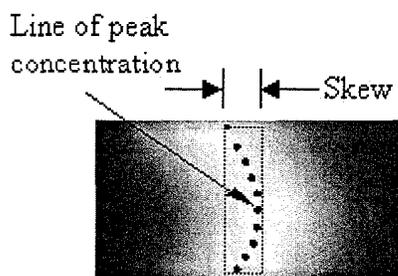
Figure 8. Determination of skew from band shape with non-linear stretching. The axial width of the line of peak concentration is used to give an estimate of the skew. This modified definition gives the same result for analyte bands that are linearly skewed.

Simulations using MEMCAD verify the large skew predicted by Eq. 3. A sample simulation is shown in Figure 9 and compared with bleached fluorescence imaging in the same geometry. The extracted skews for the simulation and experiment shown in Figure 9 are 750 μm and 800 μm, respectively.

Results from visualization of a compensating corner design are compared to MEMCAD simulations in Figure 10. A closer
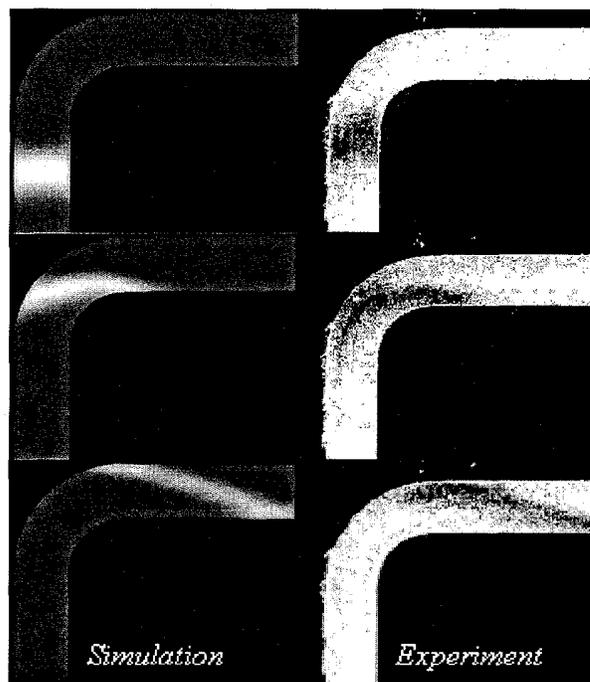
Figure 9. Simulation (left) and bleached fluorescence visualization (right) of an analyte band traveling around a constant radius corner. In both cases the channel is 250 μm wide. In the experiments, the channels were approximately 40 μm deep.
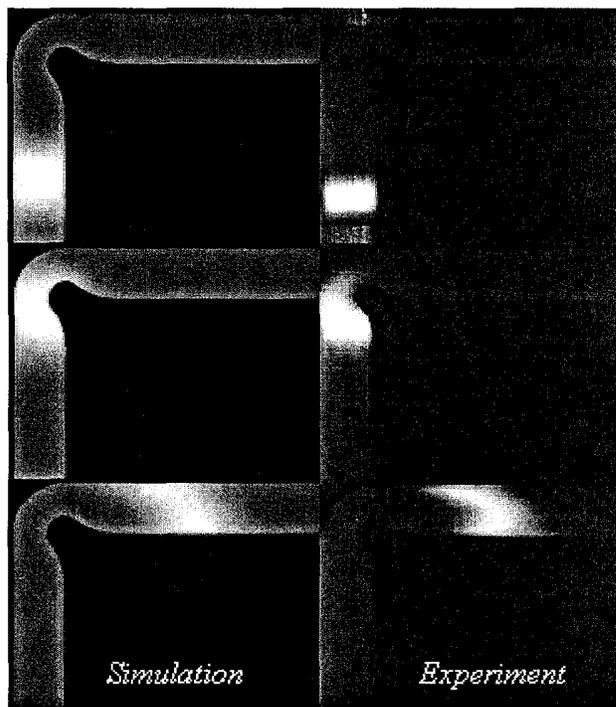
Figure 10. Simulation (left) and caged fluorescence visualization (right) of an analyte band traveling around a compensating corner design. In both cases the channel is 250 μm wide. In the experiments, the channels were approximately 40 μm deep.

look at the final band shape, as seen in Figure 11, reveals that the simulation accurately reproduces details of the band shape seen in the experiment.
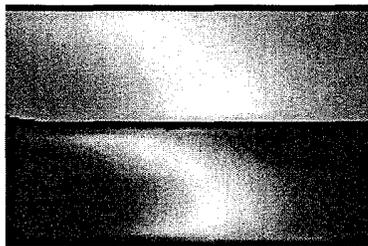
Figure 11. Close-up comparison of the band shape seen in the final frames of Figure 10. *Simulation image is shown above the experimental image. Note that the parabolic shape, although more evident in the experimental images, is also apparent in the simulation. Close inspection of the trailing end of the analyte bands shows good agreement in the amount that the bottom edge of the band leads the top edge.*

Table 1 summarizes the performance of four compensating corner designs and compares the skew extracted from matching simulations and experiments. Both experiments and simulations show that the compensating corners reduce the skew by as much as 75% when compared to the constant radius corners. However, Table 1 reveals some disagreement between simulation and experiment. Much of this error can be attributed to small amounts

Table 1. Comparison of skews obtained from simulations and experiments for a sample of compensating corners. *The starred entries indicate measurements that were slightly perturbed by pressure driven flow. The last row gives skew extracted from the experiments and simulations shown in Figure 10.*

Constriction Ratio	Recovery ratio	Skew (exp) μm	Skew (sim) μm
0.5	1	N/A	100
0.5	2	200[*]	85
0.5	3	170[*]	110
0.65	1	140	135

of pressure-driven flow, as we have found that slight imbalances in the wells (height differences of a 1-2 millimeters) can significantly disturb the skew measurements. We are currently working on new techniques to help more easily eliminate this pressure-driven flow.
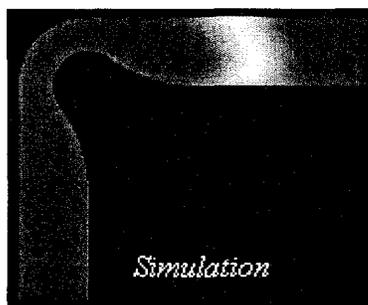

Figure 12. Simulation of a more optimized design. *In this figure, only the final shape of the analyte band is shown. By adjusting the parameters for the optimized geometry, decreases in the skew and FWHM have been obtained.*

Figure 12 demonstrates a design that is superior to the one shown in Figure 10. This design was the intended shape for the turn shown in Figure 10 but inaccuracies in the prototyping process resulted in a less successful design. We are currently working to improve these manufacturing methods.

We have also begun working to develop compensating corner designs for 180 degree turns, because we believe this design will provide more efficient use of space on a compact microfluidic device. Eq. 3 shows that the skew for a 180 degree turn is approximately 6 times the width of the channel. This large skew is evident in Figure 13.

Since the skew will decrease linearly with the reduction in the width of the channel, one method of minimizing skew is to taper the channel to an arbitrarily small width before the turn, and then return to the full width after the turn. Even if the tapering can be performed without increasing dispersion, there are many reasons why this technique may be problematic. The voltage drop across the turn will increase linearly with the reduction of the
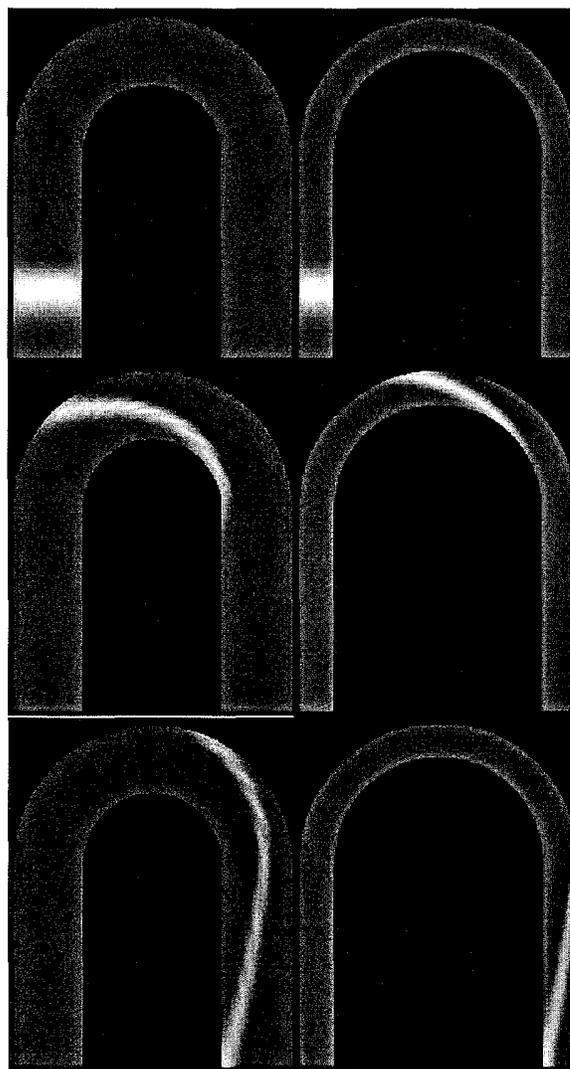
Figure 13. Simulations of an analyte band traveling through two constant radius, 180 degree turn designs. *For the purpose of comparing to compensating corner designs, both 250 μm and 125 μm wide designs are shown.*

channel width and therefore very narrow turns will reduce the field strength in the straight portions of the separation channel. Furthermore, the increased field strength in the turn will cause unwanted Joule heating and temperature rise, possibly leading to bubble formation in the turns.

Figure 14. presents simulations of an analyte band traveling through two different compensating corner designs. Although Figure 14 suggests that there is room for improvement, this design already reduces the skew by at least a factor of 5 as compared to the constant radius turns shown in Figure 13.
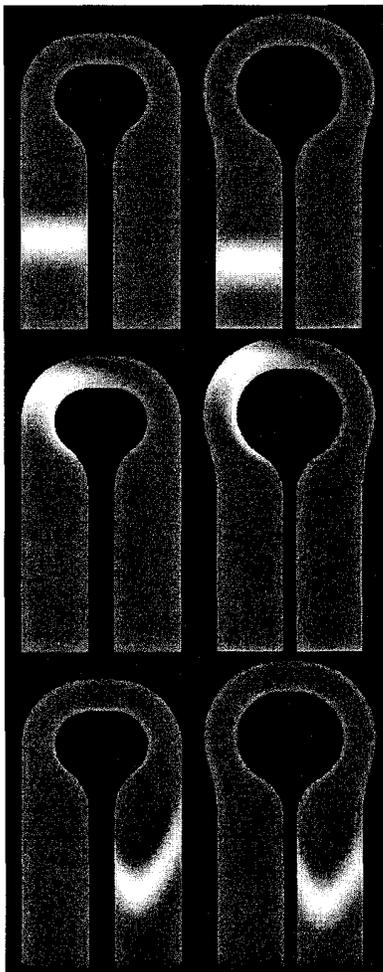
Figure 14. Simulations of two compensating, 180 degree turn designs. The first turn is essentially two of the 90 degree turns connected by a straight section. The second design incorporates an additional curvature along the outside of the turn.

Note that both of the designs shown in Figure 14 perform better than a constant radius turn having a width equal to the constriction width in the compensating designs. This enhanced performance is achieved by combining width reductions with geometry that manipulates the electric field as discussed in Figure 4. Therefore it is possible to reduce skew without the detrimental effects caused by extremely narrow constant radius turns.

CONCLUSIONS

This paper has described compensating corner designs that greatly reduce the sample dispersion caused by turns in a serpentine separation channel. Simulations were used to design the geometry of the compensating corners, and caged- and bleached fluorescence visualization techniques were used to validate the designs. The new compensating corners should be useful in realizing miniaturized electrophoresis systems that require long channels within a small area. Because of the good agreement between the simulations and experiments, we believe that simulation-based design of complete multi-turn CE separation channels can be carried out, resulting in microanalytical systems with greatly improved separation efficiencies.

ACKNOWLEDGEMENTS

REFERENCES

1. A. Manz, D. J. Harrison, E. Verpoorte, *et al.*, "Planar Chips Technology For Miniaturization of Separation Systems : a Developing Perspective in Chemical Monitoring", *Advances in Chromatography*, *33* 1-66 (1993).

2. C. S. Effenhauser, "Integrated chip-based microcolumn separation systems", *Topics in Current Chemistry*, *194* 51-82 (1998).

3. C. T. Culbertson, S. C. Jacobson, and J. M. Ramsey, "Dispersion sources for compact geometries on microchips", *Analytical Chemistry*, *70*(18) 3781-3789 (1998).

4. E. Grushka, R. M. McCormick, and J. J. Kirkland, "Effect of temperature gradients on the efficiency of capillary zone electrophoresis separations", *Analytical Chemistry*, *61*(3) 241-246 (1989).

5. R. F. Probstein, *Physicochemical hydrodynamics : an introduction.* 2nd ed. 1994, New York: John Wiley & Sons. 400.

6. E. B. Cummings, S. K. Griffiths, R. H. Nilson, *et al.*, "Conditions for similitude between the fluid velocity and electric field in electroosmotic flow", *Analytical Chemistry*, (in review).

7. B. P. Mosier and J. G. Santiago. "Bleached Fluorescence Imaging", *to be presented at the 2000 ASME International Mechanical Engineering Congress and Exposition*, Orlando, FL.

8. P. M. St. John, M. Deshpande, J. Molho, *et al.* "Metrology and Simulation of Chemical Transport in Microchannels", *Solid State Sensor and Actuator Workshop*, Hilton Head Island, South Carolina, Transducers Research Foundation (1998).

9. P. H. Paul, M. G. Garguilo, and D. J. Rakestraw, "Imaging of pressure- and electrokinetically driven flows through open capillaries", *Analytical Chemistry*, *70*(13) 2459-2467 (1998).

A SINGLE-MASK PROCESS FOR MICROMACHINED MAGNETIC DEVICES

Florent Cros, Kieun Kim, and Mark G. Allen
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0250

ABSTRACT

A method of fabrication is presented in which interlinked, electrically isolated magnetic coils and cores can be realized using a self-aligned, single-mask process. The process relies on two electroplating techniques: the differential filling and subsequent removal of metals in different aspect-ratio trenches of the same mold, in combination with more conventional plate-through-mold techniques. Since both the electromagnetic coil and the magnetic core are fabricated using the same mold, only one photolithographic process is required. Different materials can be deposited into the different aspect ratio portions of the mold. For example, in the fabrication of magnetic actuators, copper is used to fill in the high aspect-ratio winding trenches while NiFe permalloy is used to fill in the low aspect-ratio core trenches. Since the process results in electrically-self-isolated structures, the molds do not need to be removed and subsequent planarization issues (e.g., for surface micromachined devices) are greatly ameliorated. Actual geometrical limits (differences in aspect ratio) for which the process still yields functional devices were also investigated. Utilizing this process, functional electromagnetic coils of coil line widths of 15 and 50 μm have been fabricated and tested.

INTRODUCTION

Micromachined magnetic components, sensors, and actuators have application in many areas, including integrated passive components [1], magnetic field sensing [2], large-stroke actuation [3], and electrical power generation on the microscale [4]. One major drawback of magnetic microdevices, however, is the relatively complex fabrication procedure required to interlink and isolate conductors and core materials. In addition, the relatively large thickness (10-100 μm) of coils and cores required for efficient operation of these devices could present substantial planarization challenges.

As an example, consider the process developed by Taylor et al. [3] for the fabrication of magnetically-actuated, surface-micromachined relays. This process is based on conventional plate-through-mold techniques [5]. In the fabrication of the electromagnetic driving element for this relay, three masks are required: one to define lower permalloy magnetic cores, one to define copper conductors, and one to define magnetic side cores. In addition, alignment between the copper conductors and the magnetic cores is critical to prevent electrical shorting. Since thick metal lines for coil and core are required, these alignment tolerances must be maintained through thick polymer molds. The extra space required for these tolerances limits the winding packing efficiency. Moreover, in order to benefit from 'taller' (i.e., larger cross-sectional area and high aspect ratio) conductors, while

simultaneously leaving the upper surface reasonably planar for subsequent magnetic via etching and surface micromachining steps, difficult replanarization issues must be addressed. A poor degree of planarity may, for example, jeopardize subsequent processing steps such as the critical dry etch through the thick polymer molds in order to define wells for the magnetic side cores.

In this work, a method that avoids many of the above issues is presented. This process is based on conformal mold-filling techniques. Although conformal mold-filling techniques for other materials, e.g. metals [6] and silicon/polysilicon [7,8] have been previously developed, this approach differs from the previous approaches in that the deposited material is not removed by an anisotropic removal step such as chemical-mechanical polishing or reactive-ion etching. Instead, an isotropic wet etch and the differences in aspect ratios are utilized to selectively and completely (e.g., from sidewalls as well as the bottoms of trenches) remove the deposited metal from the lowest aspect ratio trenches. In addition, the mold can remain as an integral part of the final structure. This technique can also be extended (e.g., for magnetic devices) by combining it with a conventional plate though mold process. At the conclusion of this process, both the conductor and the ferromagnetic material are micromachined out of a single mold and remain embedded in the mold. Since it is not necessary to remove the molds to isolate the cores and windings, replanarization of thick electroplated structures is not required. In addition, the lack of necessity of mold removal makes the use of hard-to-remove, but otherwise desirable, materials (such as SU-8 epoxy) very attractive. These two advantages greatly simplify the fabrication process, potentially removing many of the obstacles to commercially viable magnetic component fabrication.

PROCESS

A schematic of the single-mask fabrication process is described in Figure 1. This process requires the use of a robust material able to yield molds with both high and low aspect ratio trenches. A variety of polymer micromolding techniques, such as conventional photoresists, dry-etched polyimide, and thick, negative tone epoxy resists are suitable candidates. For example, conventional thick photoresists such as AZ 9260 allow structuring of thick mold layers and offer simultaneous high and low aspect ratio molds [9]. However, the major advantage of these materials in conventional plating approaches, i.e. that they are relatively easy to remove, is not required by this process (and in fact their relatively poor stability compared with epoxy-based resists can be a drawback in this process). Thus, in this work, SU-8 epoxy photoresist is used to fabricate the mold. As demonstrated in the literature, this polymer can yield very high aspect ratio trenches with nearly vertical side-walls [10-12].

A detailed process example is given below. A 1000 μm thick NiFe (permalloy) substrate wafer is coated with a thin (5 μm) layer of SU-8 epoxy photoresist to insulate the magnetic substrate. The layer is pre-baked, cross-linked and post-baked [10], in order to create a uniform isolation layer. On top of this layer, a thicker SU-8 layer is deposited and lithographically patterned to create a micro-mold (Fig 1.1). A key characteristic of this micromold is that the sections which will ultimately define windings are of a higher aspect-ratio (thickness:width) than the sections which will ultimately define the magnetic cores. The mold is conformally coated with a seed layer of Ti/Cu/Ti (100 nm/1000 nm/100 nm) using DC sputtering (Fig 1.2). Copper is then electroplated over the entire wafer, and the high aspect-ratio molds are filled first (Fig 1.3, 1.4). Once the high aspect ratio molds are filled, the electrodeposited Cu layer is etched back (Fig 1.5). A variety of recipes were used in order to etch back the Cu (Fig 1.5); the SU-8 mold proved to withstand both acidic and basic etchants.

Note that the copper that coats any surface other than in the high aspect-ratio trenches, including that copper in the core areas, is much thinner than that in the high aspect-ratio molds. As a result, the copper (and seed layer) is selectively removed from lower aspect ratio molds as well as any flat opened surface in the field of the micromold, without the need for a lithography step (Fig 1.5). An additional benefit is that the copper lines are thereby electrically self-isolated. The insulation layer in the low aspect ratio areas is then removed in a self-aligned blanket dry etch down to the metallic surface of the substrate (Fig 1.6). Note that because the insulation layer is much thinner than the polymer mold, the blanket dry etch step does not significantly affect the height of the mold. Using the substrate as an electrical contact, NiFe is selectively electroplated, from the bottom of the NiFe substrate, filling the core mold (Fig 1.7). When the NiFe electroplating is complete, the surface has become reasonably self-planarized.
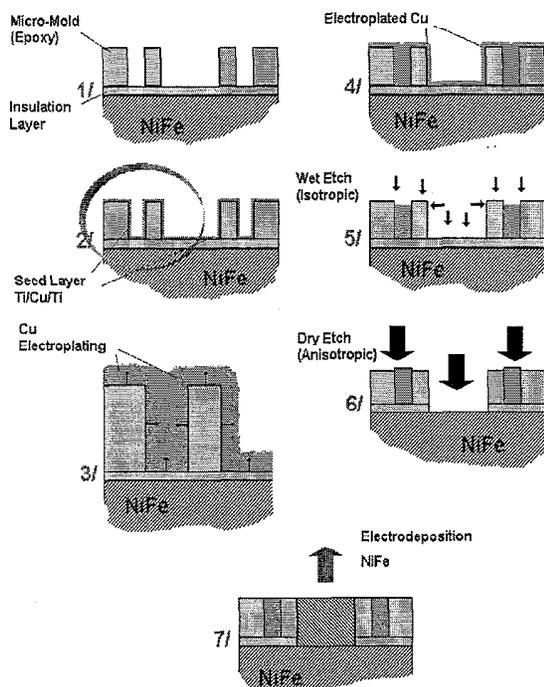
Figure 1: Schematic of the single mask process

DISCUSSION

In order to generate a high magnetic force yet maintain a relatively low driving current in a micromachined magnetic device, it is necessary to increase the number of turns in the coil (thereby increasing the number of ampere-turns in the magnetic circuit). At the same time, in order to make efficient use of the real-estate available, the highest possible packing density for the coil structure is desirable. Therefore, it is of interest to determine the actual geometrical limits (difference between low and high aspect ratio) for which the process still yields functional devices. The limit depends on the successful removal of the deposited metal from the lower aspect ratio trench while keeping the metal in the higher aspect ratio trench. Several test structures were created to probe these limitations. A 50 μm thick epoxy mold with height-to-width aspect ratio trenches ranging from 5 to 0.25 were fabricated, and subjected to steps 1-5 of Figure 1. The etch back was stopped when it was observed that the copper within the highest aspect ratio trenches began to be removed. Cu was successfully completely removed from trenches as narrow as 60 microns (aspect ratio close to unity) while higher aspect ratio trenches are still filled with copper. Figure 2.a and 2.b illustrates the differential filling and removal of copper from different aspect-ratio trenches.
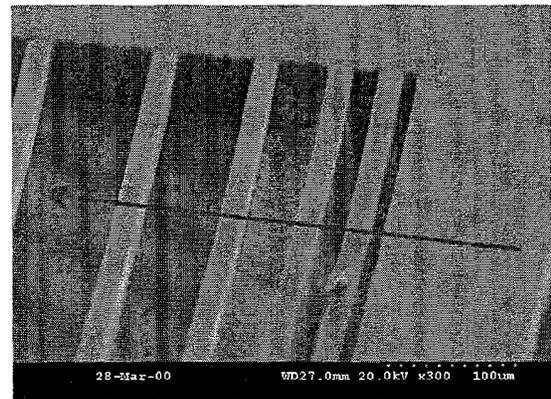
Figure 2.a: *SEM photomicrograph of different aspect-ratio trenches. Notice that the wider trenches are devoid of metal whereas the narrower ones are still partially filled with them.*
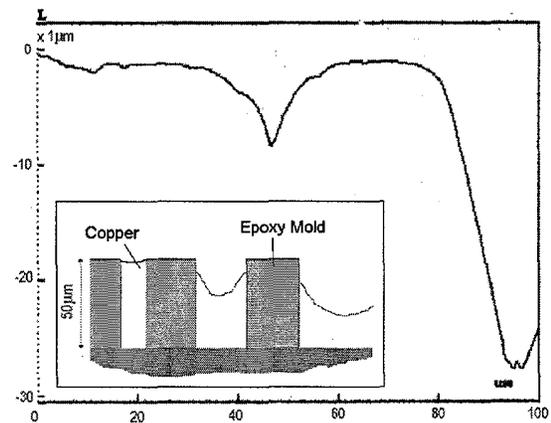
Figure 2.b: *Profile of the filled trenches as measured using a surface profilometer. Although the trench depths measured in this manner are only semiquantitative, due to the finite curvature of the profilometer stylus, they corroborate the visual SEM image of filled trench trends shown in Figure 2.a.*

It was observed that when the copper was etched back and stopped prematurely, some residual copper was found accumulated at the bottom of the low aspect-ratio trenches. This was especially true at the corners where the sidewall meets the trench floor. This phenomenon can be attributed to the fact that the copper thickness is greater at the corners than at the planar surfaces.

For a 50 μm thick mold, it was observed that in order to keep the copper intact in a trench with an aspect-ratio of 5, the lowest aspect-ratio that can be etched is 0.8. In other words, simply as a function of dimensional differences, the ratio between the narrowest width to the widest width should be 1-to-5 for this mold thickness. This will allow for optimal utilization of the single-mask process.

Figures 3 and 4 show photomicrographs of functional electromagnets fabricated using this single-mask technique and the full process of Figure 1. The copper line width is 50μm and the initial height of the epoxy mold is also 50 μm. Pads for electrical contact were created by electroplating Cu around and in between an array of epoxy posts.
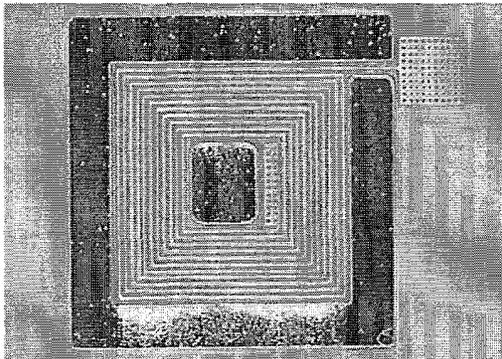
Figure 2: Photomicrograph of fabricated device with 50μm coil line width. Silver-colored areas represent NiFe magnetic cores.
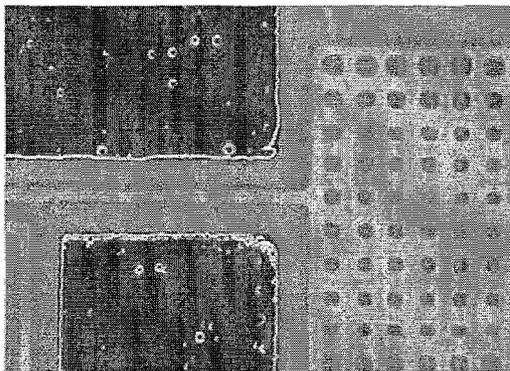
Figure 3: Close view of Fig 2.

Micro-coils with smaller line widths have also been successfully fabricated using the single-mask process. Figure 4 shows a detail of 15 μm wide, electrically isolated copper coils embedded in a 45 μm thick epoxy mold. The upper surface of the copper lines has been etched back close to the surface of the epoxy. NiFe has been subsequently partially plated in the large opening. For maximum probability of successful fabrication, openings for permalloy material were at least an order of magnitude larger than the actual with of the copper wires.
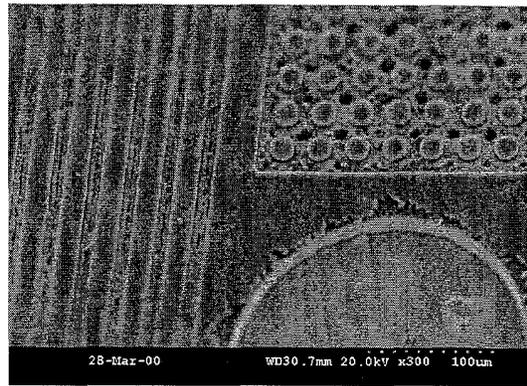
*Figure 4: SEM micrograph of a micro-electromagnetic coil with 15μm Cu coil linewidths. The semi-circle in the lower right of the micrograph is the tip of the partially filled NiFe magnetic core.*

To demonstrate the functionality of these magnetic elements, measurement of the magnetic flux generated by these elements as a function of applied coil current were performed. The probe of a Bell 9550 Gauss/Teslameter was positioned over the central pole of the electromagnet of Figure 2 and the generated flux measured as a function of applied current. Figure 5 shows the dependence of generated magnetic flux as a function of applied coil current. Due to the large size of the gaussmeter probe relative to the coil, only a relative measurement of flux is possible. However, as expected, the current-flux relationship is linear, demonstrating proper operation of the electromagnet
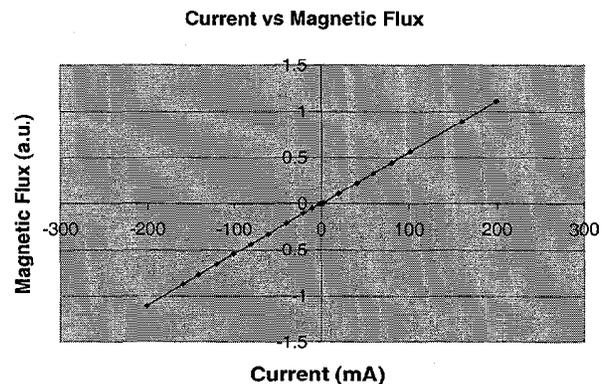
Figure 5: Flux generation as a function of current for the coil of Figure 2.

CONCLUSIONS

A simplified methodology for the microfabrication of integrated electromagnetic structures using a self-aligned, single-mask process has been developed. It exploits differences in aspect-ratios for the differential filling of mold trenches using a combination of conformal and standard electrodeposition-through-mold techniques. The electroplating mold is retained after processing as a planarized surface for subsequent surface micromachining. Utilizing this technology, integrated micro-coils of 15 and 50 μm have been fabricated and their magnetic performance measured. Although this process has been presented for magnetic micromachined devices incorporating windings and magnetic cores, it is also applicable to a wide variety of multimaterial structures and devices, as well as highly three-dimensional structures such as inductors.

REFERENCES

1. Park, J.Y., Bhattacharya, S.K., Allen, M.G., "Fully integrated passives modules for filter applications using low temperature processes", Proceedings. 1997 International Symposium on Microelectronics, Philadelphia, PA, USA, 14-16 Oct. 1997, SPIE vol.3235 , p.592-7

2. Liakopoulos, T.M., Ming Xu; Chong Ahn, Nelson, T.L., "Ultrahigh resolution DC magnetic field measurements using microfabricated fluxgate sensor chips", Proceedings of Conference on Precision Electromagnetic: Measurements, 1998 Conference on Precision Electromagnetic Measurements Digest, Washington, DC, USA, 6-10 July 1998, p.630-1

3. Taylor, W.P., Brand, O., Allen, M.G., "Fully integrated magnetically actuated micromachined relays", Journal of Microelectromechanical Systems, vol.7, no.2, June 1998, p.181-91

4. Epstein, A.H., Senturia, S.D., Anathasuresh, G., Ayon, A., Breuer, K., Chen, K.S., Ehrich, F., Gauba, G., Ghodssi, R., Groshenry, C., Jacobson, S., Lang, J., Mehra, C.C., Mur Miranda, J., Nagle, S., Orr, D., Piekos, E., Schmidt, M., Shirley, G., Spearing, S., Tan, C., Tzeng, Y.S., Waitz, I., "Power MEMS and microengines", Transducers 97. 1997 International Conference on Solid-State Sensors and Actuators: Digest of Technical Papers, Chicago, IL, USA, 16-19, vol.2, June 1997, p.753-6

5. Romankiw, L.T., "Evolution of the plating through lithographic mask technology", Proceedings of the Fourth International Symposium on Magnetic Materials, Processes, and Devices. Applications to Storage and Microelectromechanical Systems (MEMS), Chicago, IL, USA, 9-12 Oct. 1995, p.253-72

6. Carey, D.H., Pietila, D.A., Sigmond, D.M., "Trenching Techniques for Forming Channels, Vias, and Components in Substrates", U.S. Patent US05219787.

7. Richardson, D.L.K., Knowles, J.C., "Integral Honeycomb-like Support of Very Thin Single Crystal Slices", U.S. Patent US03936329.

8. Keller, C., Ferrari, M. "Milli-scale polysilicon structures", Technical Digest Solid-State Sensor and Actuator Workshop, Hilton Head Island, SC, USA, 13-16 June 1994, p.132-7

9. Conedera, V., Le Goff, B., Fabre, N., "Potentialities of a new positive photoresist for the realization of thick moulds", Journal of Micromechanics and Microengineering vol.9, no.2, June 1999, p.173-5

10. Lorenz, H., Despont, M., Fahrni, N., Brugger, J., Vettiger, P., Renaud, P., "High-aspect-ratio, ultrathick, negative-tone near-UV photoresist and its applications for MEMS", EEE Tenth Annual International Workshop on Micro Electro Mechanical Systems. An Investigation of Micro Structures, Sensors, Actuators, Machines and Robots, Nagoya, Japan, 26-30 Jan. 1997, vol.A64, no.1, p.33-9

11. Chang H.K., Kim Y.K., "UV-LIGA process for high aspect ratio structures using stress barrier and C-shaped etch holes", Transducers 99, June 7-10, Sendai, Japan, vol.2, p1428-1431

12. Lorenz, H., Despont, M., Fahrnl, N., LaBianca, N., Renaud, P., Vettiger, P., "SU-8: a low-cost negative resist for MEMS", Journal of Micromechanics and Microengineering, vol.7, no.3, p.121-4

ON-CHIP EDDY CURRENT SENSOR FOR CRACK DETECTION IN METALS

Daniel J. Sadler and Chong H. Ahn
Center for Microelectronic Sensors and MEMS
Department of Electrical and Computer Engineering and Computer Science
University of Cincinnati
P.O. Box 210030, Cincinnati, OH 45221-0030 U.S.A

ABSTRACT

In this paper, we present a new integrated eddy current sensor for the detection of microcracks on the surface of metals. The device consists of two stacked planar coils fabricated onto a glass substrate and encapsulated on one side by a Ni/Fe permalloy magnetic core. Fabrication of the device is achieved by a UV-LIGA thick photoresist lithography process which involves the lithographic patterning of 15-25 μm thick molds using AZ-4000 series photoresist. The introduction of the permalloy core coupled with the thick conductor lines produces a high inductance, low resistance device capable of generating large magnetic fields at low driving currents. Aluminum and titanium testing standards containing cracks 7 mils wide and 8 to 40 mils deep have been used to confirm the excellent operation of this new sensor. The device has been tested in the frequency range of 10 kHz - 500 kHz and has been shown to be capable of clearly detecting cracks with depths of as little as 8 mils (200 μm) at a low input power of 30 mW. Results show an extremely linear relation between crack depth and output signal voltage with a reasonably high level of unamplified signal strength.

INTRODUCTION

Eddy current testing (ECT) is a very effective and convenient way to nondestructively evaluate a metallic surface for defects. A variety of industrial applications currently exist for nondestructive eddy current testing, especially in the aerospace industry. Although many excellent and sophisticated devices are commonly used in industry already, MEMS technology is being investigated as a way to potentially lower the cost, size, and improve the performance of these existing devices [1-4]. Recently, a variety of MEMS-based eddy current sensors have been reported for use as proximity sensors and position indicators [5-6], as well as for use in nondestructive testing for cracks and flaws in metals [1-3].

In this paper, we present a new integrated eddy current sensor for the detection of microcracks on the surface of metals. The device consists of two stacked planar coils fabricated onto a glass substrate and encapsulated on one side by a Ni/Fe permalloy magnetic core. A UV-LIGA thick photoresist lithography process, which involves the lithographic patterning of 15-25 μm thick molds using AZ-4000 series photoresist, was used to fabricate the new eddy current sensor [7-9]. Although this new device is similar in structure and operating principle to some other previously reported devices [1-3], there are several important distinctions and improvements. First, the entire device is fabricated on one substrate with no hand-wound or attached coils. Also, the introduction of the permalloy core coupled with the thick conductor lines produces a high inductance, low resistance device capable of generating large magnetic fields at low driving currents. For these reasons, the devices presented in this work show promise for use as low cost, highly sensitive on-chip eddy current sensors.

THEORY

Two common methods exist for eddy current testing, a single coil and a dual coil method. Both rely on the principle that as a coil is excited with an AC voltage, flux lines emanate from one face of the coil, arch around in the space surrounding the coil, and enter the opposite face of the coil as shown in Figure 1. The total amount of flux generated is directly proportional to the amount of current flowing through the coil. The flux lines will close on themselves without crossing any other lines, but if any flux line impinges upon a conducting object, a small whirlpool like eddy current will be created at the surface of the object under test. These eddy currents, in turn produce a magnetic field which opposes that field which was generated by the driving coil. The ANSYS simulation of Figure 2 shows the change in magnetic flux linking the device that results when a conducting object is brought into the vicinity of a coil. In the single coil approach, this change in magnetic field is sensed as a change in coil inductance. The dual coil method, on the other hand, employs a second coil, called the sensing coil, which is magnetically linked to the driving coil due to its close proximity. Because the drive and sensing coils are linked, and because the generated flux is time changing (AC), an electromotive force (V_{emf}) is induced in the sensing coil as per Faraday's law which is given as follows:

$$V_{emf} = -N\frac{d\Psi}{dt},$$

where Ψ is magnetic flux and N is the number of coil turns. Therefore, by simply monitoring the voltage coupled onto the sensing coil, changes in the magnetic field can easily be detected.
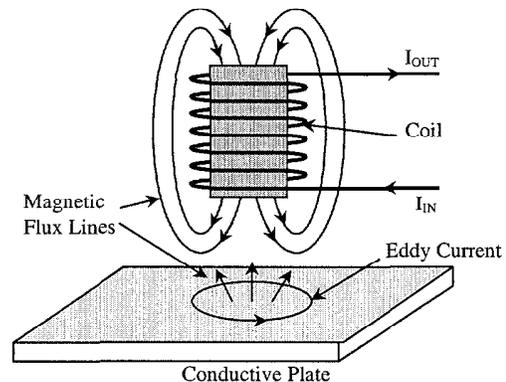
Figure 1. Magnetic fields produced by an inductor interact with the surface of a conductor to produce eddy currents. The eddy currents then produce a magnetic field which opposes the field generated by the inductor.

Although the above theoretical description is more applicable to a proximity sensor, the same analysis can be readily extended to the understanding of eddy current crack detection. Sensing the presence of the metal object can be thought of as the first order effect, while detecting defects in that object are a second order effect. Crack detection is possible by the aforementioned method due to the fact that defects in the metal object under test impede the path of eddy current flow and cause an additional change in the magnetic flux in the vicinity of the defect. Therefore, by fixing the distance between the sensor and the target, deviations in the

measured signal can be interpreted as surface flaws in the object under test, making it possible to detect the presence of a surface flaw or crack and even characterize its dimensions.
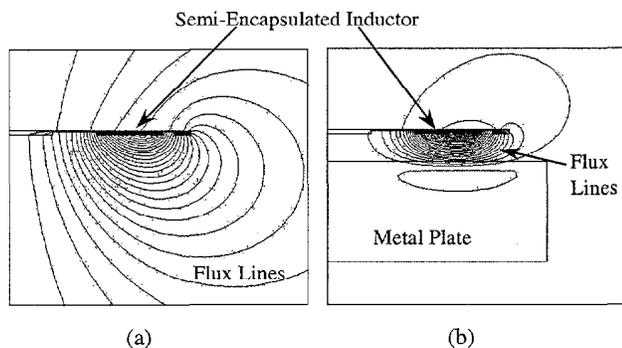
Figure 2. ANSYS simulation of flux generated by a semi-encapsulated spiral inductor (a) isolated from a metal surface and (b) in the vicinity of a metal plate. Detection of this difference in flux profile (magnetic field) provides the physical basis for proximity sensors and crack detection sensors.

In this work, we have chosen the dual coil method for two reasons. First, output signals are typically larger for this method than for the single coil method [2]. Second, the actual measurement technique of simply measuring an induced voltage is more straightforward. Although this method is simple, typical problems include difficulty in device integration, low output signal strength (hundreds of microvolts or less), and high power dissipation in the driving coil. In this paper, we intend to show that through innovative design and fabrication methods, our device is able to overcome these problems and has the potential for even further improvement.

DESIGN

A schematic representation of the sensor is shown in Figure 3. As shown in the diagram, the driving coil is located directly above the sensing coil, and the two are separated by a thin dielectric layer. In this way, the two coils are isolated electrically, but coupled magnetically due to their close proximity. For optimum performance as an eddy current sensor, coils should have both a high inductance for generating a large magnetic field and a low electrical resistance to minimize ohmic losses. In order to maximize the inductance value, a large number of turns and a tight coil packing (small conductor spacing) are required. Conversely, resistance is minimized by using conductors with a large cross-sectional area and a low number of turns. In addition, device size requirements and fabrication limitations must also be taken into consideration. Because some of these requirements are contradictory, engineering tradeoffs must be made.

In the current design, both the sensing and driving coils contain 13 full turns of copper windings. Conductor line width is 50 μm, thickness is approximately 20 μm, and spacing is 30 μm for both coils. This results in an overall device size of about 7 mm x 7 mm. The above dimensions were chosen in an effort to create a relatively high inductance and relatively low resistance device without unnecessarily complicating fabrication. All dimensions are easily achievable using our UV-LIGA process [7-9], and we believe that by further optimizing dimensions, a device with similar or better electrical characteristics and much smaller size can be produced with little additional fabrication effort. In addition, this technology could be extended to the development of an array-type sensor as pictured in Figure 4 which would likely be required in an actual industrial application.
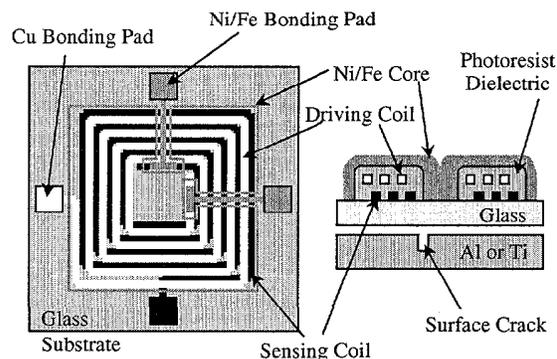

Figure 3. Top view and cross-sectional view schematic diagrams of the eddy current sensor. The entire sensor is integrated on a glass substrate including the lower sensing coil, the upper driving coil, and the semi-encapsulating magnetic core. Layers are separated by a photoresist dielectric.
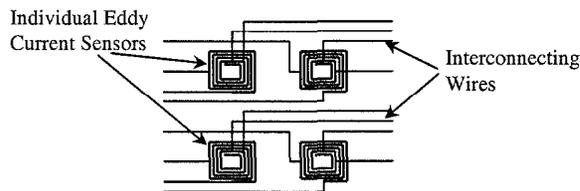

Figure 4. 2 x 2 array of the eddy current sensors. Individual sensors are combined to form an array for high speed scanning of a metal surface as would be required in an industrial nondestructive testing application.

FABRICATION

Brief fabrication steps for the device are shown in Figure 5. Pyrex 7740 glass wafers two inches in diameter and 250 μm-thick were used as substrates. Each two-inch glass substrate yielded sixteen eddy current sensors. The process began with e-beam evaporation of a Ti (300 Å)/Cu (3000 Å) seed layer. AZ-4000 series photoresist was then used to produce a 20 μm thick electroplating mold. Next, copper lines were electroplated to the top of the mold using standard electroplating techniques to form the sensing coil. After removing the photoresist and seed layer, a new layer of photoresist was spun, via openings were patterned, and the photoresist was hard cured at 220 °C to form a permanent and planarizing dielectric layer. This dielectric layer electrically insulates the lower sensing coil from the driving coil to be patterned directly above it. Another seed layer was then deposited from which the driving coil was electroplated using the same exact method as described previously. Then photoresist and seed layers were removed and another hard curing was performed to insulate the driving coil from the semi-encapsulating magnetic core. Another seed layer was deposited and AZ-4000 series photoresist was again used, this time to pattern a mold for the semi-encapsulating magnetic core. The core was formed by electroplating Ni/Fe (81%/19%) permalloy to a thickness of about 20 μm. It should be noted that in addition to serving as a magnetic material, this Ni/Fe permalloy layer is also an electrical conductor and is used to form a bonding pad for making contact to the inside of the coil. This eliminates one mask from the fabrication process. Finally, photoresist and seed layers were removed to complete device fabrication. Upon completion of the microfabrication steps, wafers were diced and individual sensors packaged for testing. A microphotograph of one completed device is shown in Figure 6.
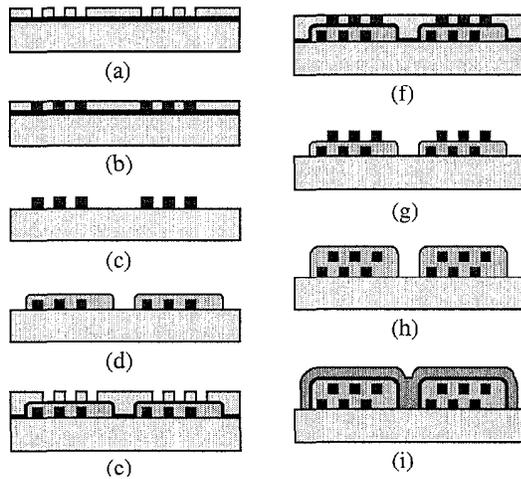
Figure 5. Brief fabrication steps: (a) Photolithography for sensing coil; (b) Copper electroplating; (c) Resist and seed layer removal; (d) High temperature curing; (e) Photolithography for driving coil; (f) Copper electroplating; (g) Resist and seed layer removal; (h) High temperature curing; (i) Seed layer evaporation and top magnetic core electroplating.
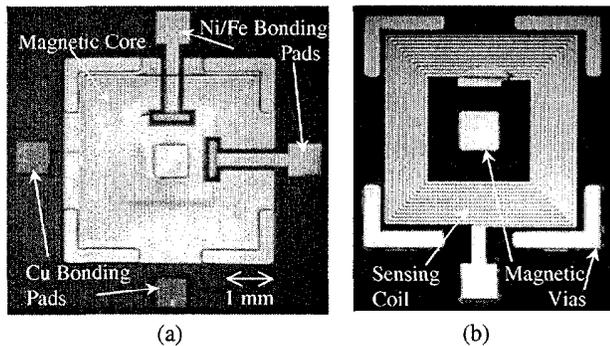

Figure 6. (a) Top view, and (b) Rear view looking through the substrate of the eddy current sensor fabricated on a Pyrex glass wafer using dual-stacked spiral inductors. Copper and Ni/Fe pads allow for connection to each of the inductors. In (a), the driving and sensing coils are hidden beneath the semi-encapsulating Ni/Fe permalloy core. Overall sensor size is about 7mm x 7mm.

RESULTS AND DISCUSSION

Inductance and resistance measurements were performed on both the driving and sensing coils of the eddy current sensors by using a Hewlett Packard precision LCR meter. Because thickness and geometry of sensing and driving coils are the same, they have nominally the same resistance and inductance values. Figure 7 shows a graph of average inductance and resistance as a function of frequency for several sensors. The relatively low resistance (3 ohms) and high inductance (1.6 µH) can be attributed respectively to the thick conductor lines and the semi-encapsulating Ni/Fe permalloy core. This combination results in the generation of large magnetic fields with low power dissipation from ohmic losses.

After verifying the excellent electrical properties of the eddy current sensors, these devices were then tested for their ability to detect surface cracks in metals. Aluminum and titanium testing standards were chosen for the device characterization because these materials are used extensively in aerospace applications. The testing standards, pictured along with a sensor in Figure 8, contain cracks 7 mils wide and 8 to 40 mils deep. Devices were tested in the frequency range of 10 kHz - 500 kHz and were shown

to be capable of clearly detecting cracks with depths of as little as 8 mils (200 µm) at a low input power of 30 mW. Figures 9 and 10 show the device response as a function of input signal frequency when used to detect cracks of various sizes in both aluminum and titanium. In these figures, output signal refers to the RMS-voltage induced on the sensing coil when the sensor is placed over top of a crack minus the RMS-voltage induced in the sensing coil when the sensor is placed over a flawless metal surface. Overall, these results show a reasonably high level of unamplified signal strength at low input power. Additionally, testing results at a frequency of 200 kHz are shown in Figure 11. These results show an extremely linear relationship between crack depth and output signal voltage. The excellent results achieved from these tests show that these devices have much promise for use as low cost, highly sensitive eddy current sensor arrays.
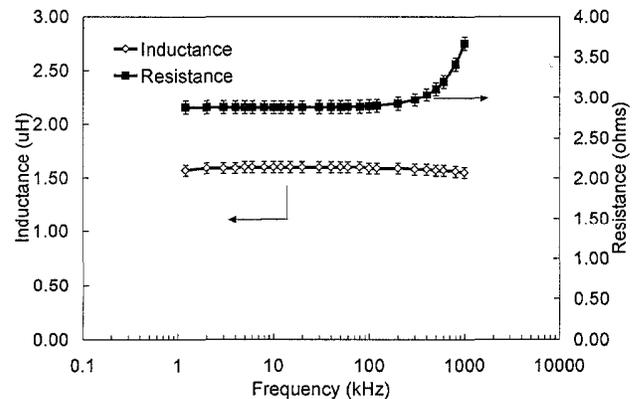
Figure 7. Inductance and resistance values as a function of frequency for spiral coils.
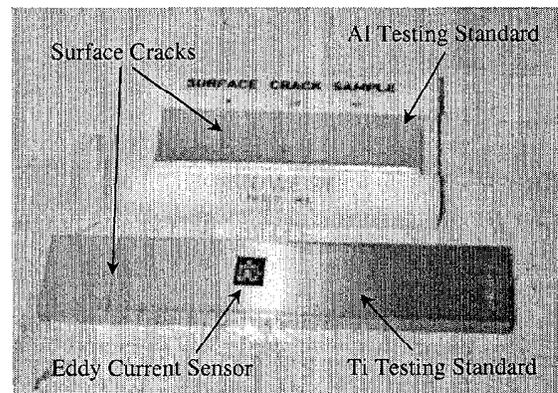
Figure 8. Aluminum and titanium testing standards pictured along with and eddy current sensor. The aluminum testing standard has three cracks with depths of 8, 20, and 40 mils, while the titanium standard has four cracks with depths of 10, 20, 30, and 40 mils.

CONCLUSION

In this work, a new integrated eddy current sensor for the detection of microcracks on the surface of metals has been designed, fabricated, and tested. This new sensor consists of two stacked planar coils fabricated onto a glass substrate and encapsulated on one side by a Ni/Fe permalloy magnetic core. The sensor is fabricated by a UV-LIGA thick photoresist lithography process and is fully integrated in that it contains no hand-wound or attached coils. This UV-LIGA method allows for the formation of

thick conductor lines, thus resulting in a low resistance device. These thick conductors, coupled with the addition of the semi-encapsulating Ni/Fe Permalloy core, produce a high inductance, low resistance device capable of generating large magnetic fields at low driving currents. Current overall device size is 7 mm x 7 mm, however, by further optimizing dimensions, a sensor with similar or better electrical characteristics and much smaller size should be possible with little additional fabrication effort.
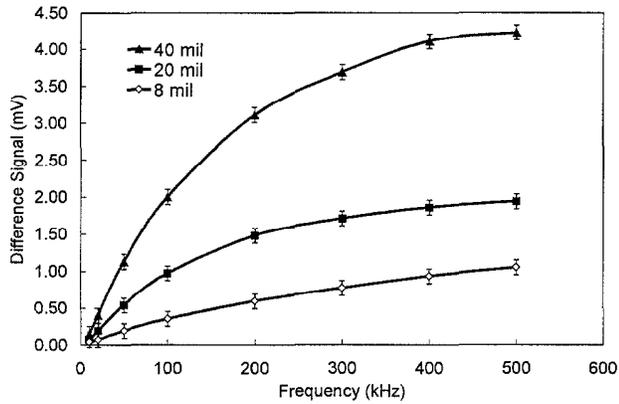
Figure 9. Difference signal vs. frequency for cracks of various depths in aluminum. Input signal power is 30 mW.
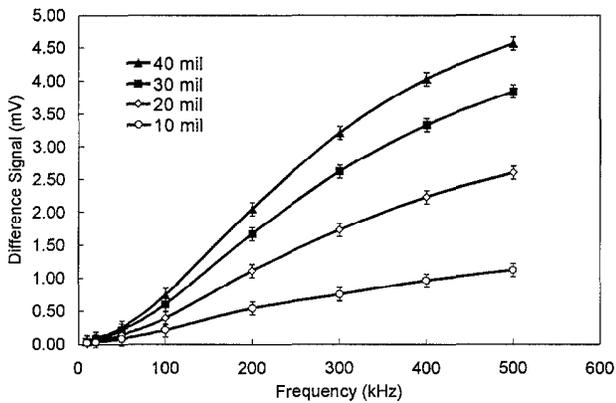
Figure 10. Difference signal vs. frequency for cracks of various depths in titanium. Input signal power is 30 mW.
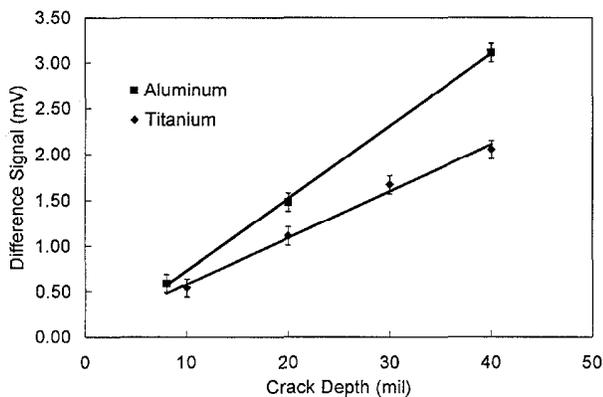
Figure 11. Output signal as a function of crack depth in both aluminum and titanium at an input frequency of 200 kHz and input power of 30 mW.

The sensors have been tested for their ability to detect surface cracks using both aluminum and titanium testing standards which contain cracks 7 mils wide and 8 to 40 mils deep. Tests on these samples has confirmed that these new sensors are capable of clearly detecting cracks with depths of as little as 8 mils (200 μm) at a low input power of 30 mW. Results show an extremely linear relation between crack depth and output signal voltage with a reasonably high level of unamplified signal strength. These initial results are very promising and show that these devices have great potential for use as low-cost, highly sensitive eddy current sensor arrays.

ACKNOWLEDGEMENT

REFERENCES

1.   M. Uesaka, T. Nakanishi, K. Miya, H. Komatsu, K. Aoki, and K. Kasai, "Micro Eddy Current Testing by Micro Magnetic Sensor Array", *IEEE Transactions on Magnetics,* **Vol. 34 No. 4**, 1998, pp. 2287-2297.

2.   S. Yamada, M. Katou, M. Iwahara, and F. P. Dawson, "Eddy Current Testing Probe Composed of Planar Coils", *IEEE Transactions on Magnetics,* **Vol. 31 No. 6**, 1995, pp. 3185-3187.

3.   Y. Hamasaki, and T. Ide, "Fabrication of Multi-Layer Eddy Current Micro Sensors for Non-Destructive Inspection of Small Diameter Pipes", *IEEE Microelectromechanical Systems Workshop*, Amsterdam, The Netherlands, 1995, pp. 232-236.

4.   W. J. Ballantyne, "Mechatronic Applications of Eddy Current Sensing", *SENSORS,* January 1995, pp. 48-52.

5.   P. A. Passeraub, P. A. Besse, A. Bayadroun, S. Hediger, E. Bernasconi, and R. S. Popovich, "First Integrated Inductive Proximity Sensor with On-Chip CMOS Readout Circuit and Electrodeposited 1 mm Flat Coil", *Sensors and Actuators,* **A 76**, 1999, pp. 273-278.

6.   G. Y. Tian, Z. X. Zhao, and R. W. Baines, "The Research of Inhomogeneity in Eddy Current Sensors", *Sensors and Actuators,* **A 69**, 1998, pp. 148-151.

7.   D. J. Sadler, W. Zhang, C. H. Ahn, H. J. Kim, and S. H. Han, "Micromachined Semi-Encapsulated Spiral Inductors for MEMS Applications", *IEEE Trans. on Magnetics,* **Vol. 33 No. 5**, 1997, pp. 3319-3321.

8.   D. J. Sadler, T. M. Liakopoulos, and C. H. Ahn, "A New Micromachined Electromagnet with Electroplated Ni/Fe Permalloy Through-Holes," *Proc. Electrochemical Society: The 5th International Symposium on Magnetic Materials, Processes, and Devices*, Boston, MA, 1998, pp. 377-388.

9.   T. M. Liakopoulos, M. Xu, and C. H. Ahn, "A Micro Fluxgate Magnetic Sensor Using Micromachined 3-Dimensional Planar Coils", *Proc. Solid-State Sensor and Actuator Workshop*, Hilton Head, SC, 1998, pp. 19-22.

LATCHING MICRO ELECTROMAGNETIC RELAYS

M. Ruan and J. Shen[#]

Dept. of Electrical Engineering and Center for Solid State Electronics Research, Arizona State University, Tempe, AZ 85286. [#]Correspondence: Email: jshen@asu.edu; Phone: (480)-965-9517

C. B. Wheeler

Microlab, Inc., 6401 E. Hummingbird Ln., Paradise Valley, AZ 85253-3608

ABSTRACT

This paper describes a new type of latching micro magnetic relay that has recently been demonstrated. The device is based on preferential magnetization of a permalloy cantilever in a permanent external magnetic field. Switching between two stable states is accomplished by a short current pulse through an integrated coil underneath the cantilever. Some key features are summarized as follows. (1) Latching (bistable); (2) low energy consumption during switching ($\lesssim 93$ µJ, switching current ~ 60 mA, minimum switching pulse width ~ 0.2 ms); (3) low voltage operation ($\lesssim 5$ V); (4) maximum DC current >500 mA; (5) capable of various switch configurations [single-pole-single-throw (SPST), multi-pole-single-throw (MPST), or multi-pole-double-throw (MPDT)]; (6) low contact resistance ($\lesssim 50$ mΩ); (7) operation in ambient environment; (8) lifetime expected to be comparable to other micro relays, (9) batch fabrication using planar processing methods.

INTRODUCTION

Latching relays are widely used in applications such as aerospace, RF communications, and portable electronics, where power consumption must be minimized. Conventional electromechanical relays typically operate by energizing an elecromagnet that actuates a soft magnetic armature to make or break a contact. When the magnet is deenergized, a spring restores the armature to its quiecent position. Recently, similar principles have been applied to microelectromechanical (MEMS) relays using the microlectronics fabrication methods [1-4]. Application of MEMS technology to relays should enable drastic reductions in fabrication costs, physical size and design complexity. To our knowledge, a latching magnetic MEMS relay has not been demonstrated.

In this work, the preferential magnetization of a permalloy cantilever is utilized to achieve two stable states. In a constant, nearly perpendicular magnetic field, a cantilever can have either a clockwise or a counter-clockwise torque depending on the angle between the cantilever and the field, which leads to the bistability. To switch the relay, a second magnetic field (generated by a current pulse through a coil in this case) realigns the magnetization of the cantilever and changes the direction of the magnetic torque, causing the cantilever to flip.

PRINCIPLE OF OPERATION

The basic structure of the microswitch is illustrated in Fig. 1. The device consists of a cantilever, an embedded planar coil, a permanent magnet, and the necessary electrical contacts. The cantilever is a two-layer composite consisting of a soft magnetic material (e.g., NiFe permalloy) on its top side and a highly conductive material, such as Au, forms the bottom surface. The cantilever is fixed at one end and is free to deflect at the opposite end (Figs. 1 and 2a). A second design variation that has also been explored has a cantilever supported by torsion flexures from the

two sides (Fig. 2b). The contact end to the right of the cantilever can be deflected up or down by applying a current through the coil. When it is in the "down" position, the cantilever makes electrical contact with the bottom conductor, and the switch is "on" ("closed"); when the contact end is "up", the switch is "off" ("opened"). The permanent magnet holds the cantilever in either the "up" or the "down" position after switching, making the device a latching relay.
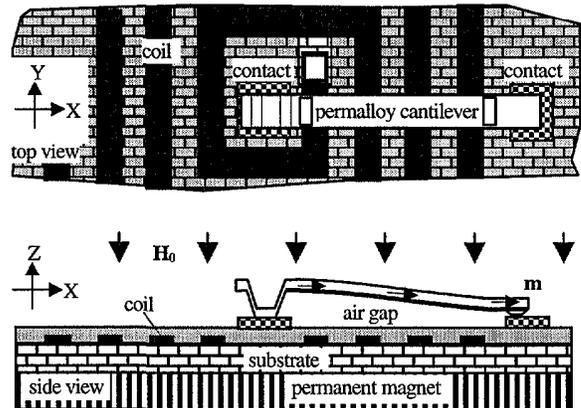
Figure 1. Top and side views of the latching relay. Planar coils are integrated on the same substrate underneath the cantilever. A permanent magnet provides the constant magnetic field H_0 which holds the cantilever to either of its stable states.
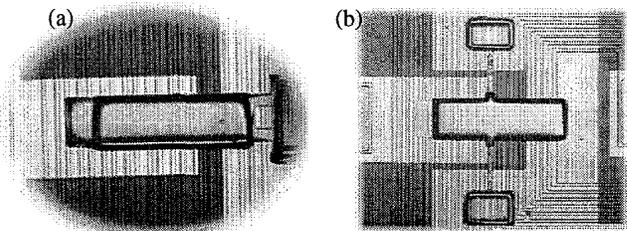
Figure 2. Photographs of the fabricated latching relays. Two types of design [(a) one-end fixed; (b) center-hinged] have been investigated and both types are successfully demonstrated. Typical beam size: 800 µm × 200 µm × 25 µm.

(i) Method to produce bistability

When the length L of a permalloy cantilever is much larger than its thickness t and width w, the direction along its long axis (L) becomes the preferred direction for magnetization (easy axis). When such a cantilever is placed in a uniform permanent magnetic field, a torque is exerted on the beam. The torque can be either clockwise or counterclockwise, depending on the initial orientation

of the beam with respect to the magnetic field. When the angle (α) between the beam axis (ξ) and the external field (H_0) is smaller than 90°, the torque is counterclockwise; and when α is larger than 90°, the torque is clockwise (Fig. 3). The bi-directional torque arises because of the bi-directional magnetization (by H_0) of the beam (from left to right when α<90°, and from right to left when α>90°). Due to the torque, the beam tends to align with the external magnetic field (H_0). However, when a mechanical force (such as the elastic torque of the beam, a physical stopper, etc.) preempts to the total realignment with H_0, two stable positions ("up" and "down") are available, which forms the basis of latching in the relay.
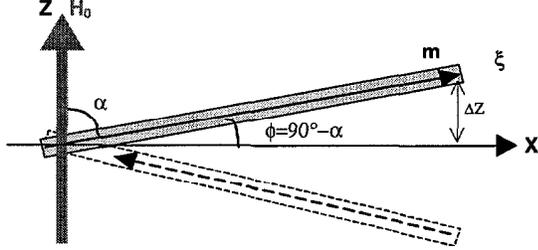
Figure 3. Schematic drawings of the relay for discussions. H_0 is the magnetic field provided by the permanent magnet, m is the magnetic moment of the cantilever, α is the angle between H_0 and m, $\phi=90°-\alpha$, and Δz is the deflection height at the end.

(ii) Electrical switching

If the bi-directional magnetization along the easy axis of the beam arising from H_0 can be momentarily reversed by applying a second magnetic field to overcome the influence of (H_0), then it is possible to achieve a switchable latching relay. This scenario is realized by situating a planar coil under the cantilever to produce the required temporary switching field. The planar coil geometry was chosen because it is relatively simple to fabricate, though other structures (such as a wrap-around type) are also possible. The magnetic field (H_{coil}) lines generated by a short current pulse loop around the coil. It is mainly the ξ-component (along the beam, see Fig. 3) of this field that is used to reorient the magnetization in the beam. The direction of the coil current determines whether a positive or a negative ξ-field component is generated. After switching, the permanent magnetic field holds the beam in this state until the next switching event is encountered. Since the ξ-component of the coil-generated field ($H_{coil-\xi}$) only needs to be momentarily larger than the ξ-component [$H_{0\xi}=H_0\cos(\alpha)=H_0\sin(\phi)$, $\alpha=90°-\phi$] of the permanent magnetic field and ϕ is typically very small ($\phi\leq 5°$), switching current and power can be very low, which is an important consideration in micro relay design.

DESIGN AND FABRICATION

(1) Torque and bistability

Assuming a uniform magnetization of the beam, the magnetic moment can be expressed as $m=MV$, where M is the magnetization in the beam and is determined by the external magnetic field (H_0), the beam geometry and magnetic properties, and V is the volume of the beam. The magnetic bending moment τ_m is obtained using

$$\tau_m = m \times B_0 = \mu_0 m \times H_0 , \qquad (1)$$

For the cantilever design with one fixed end, a key question is whether the magnetic force is large enough to overcome the mechanical force so that the beam will deflect and make contact with the contact pad on the substrate.[1] According to the elasticity theory, the elastic torque needed to cause a deflection Δz at the free end of the cantilever is [5]

$$\tau_e = wt^3 \, \Delta z \, E/4L^2 , \qquad (2)$$

where E is the Young's modulus of the beam. For a beam of L=600 µm, w=10 µm and t=1 µm, and assuming E=2×10^{11} N-m^{-2} for the permalloy [6], we have for a deflection of Δz=12 µm (ϕ=1.6°),

$$\tau_e = 1.8 \times 10^{-11} \text{ N-m} . \qquad (3)$$

According to (1), we numerically calculated the magnetic torque in a magnetic field of 0.037 Tesla and obtained a value of

$$\tau_m = 8.4 \times 10^{-10} \text{ N-m} , \qquad (4)$$

nearly 4.7 times as large as τ_e. In the calculation, we have used the permalloy B-H curves and included the demagnetization factors. In principle, the weight of the beam causes the beam to bend downward, but this bending can be shown to be negligibly small.

(2) Switching field

The current-induced coil magnetic field has both an X and a Z component. It is mainly the X component that is used for switching.

The magnetic field distribution around the planar coil can be calculated using the Biot-Savart Law:

$$\vec{B}_{coil} = \int d\vec{B} = \frac{\mu_0}{4\pi} \int \frac{I \cdot d\vec{l} \times \vec{r}}{r^3} \qquad (5)$$

A numerical routine was developed to calculate the field distribution. The calculated magnetic field at positions Z=4 µm above the X-axis is calculated and shown in Fig. 4. B_z and B_x are the field components along Z- and X-axis respectively. In the calculation, we used 25 turns for the coil. The spacing between the adjacent turns is Δx=20 µm. The current in the coil is 50 mA. Because the beam lies in the positive side of the X-axis, we show in more detail the X-component of the magnetic field in that region in Fig. 4 for several different heights (z=4-80 µm). The oscillatory behavior in B for small Z (=4 µm) is due to the proximity effect and the discreteness of the coil line separation. The average field strength is in the approximate range between 0.001 Tesla and 0.002 Tesla.

To evaluate under what conditions the coil current can switch the relay, the magnetization process needs to be examined in more detail. As mentioned above, when the aspect ratio of a beam is large (i.e., L>w>>t), the magnetization primarily aligns in the direction of L. Much smaller components of the magnetization also exist along the directions of w and t, but can be neglected due to the large demagnetization in these directions. Thus the beam's magnetization is primarily determined by the external field component along its long axis. To switch the relay, the coil

---

[1] This is not an issue for the hinged-type design because the mechanical torsion can be very small in that case.

current-induced magnetic field $(B_{coil-\xi})$ must overcome the external field $(B_{0\xi})$ along the cantilever's length, at least momentarily.
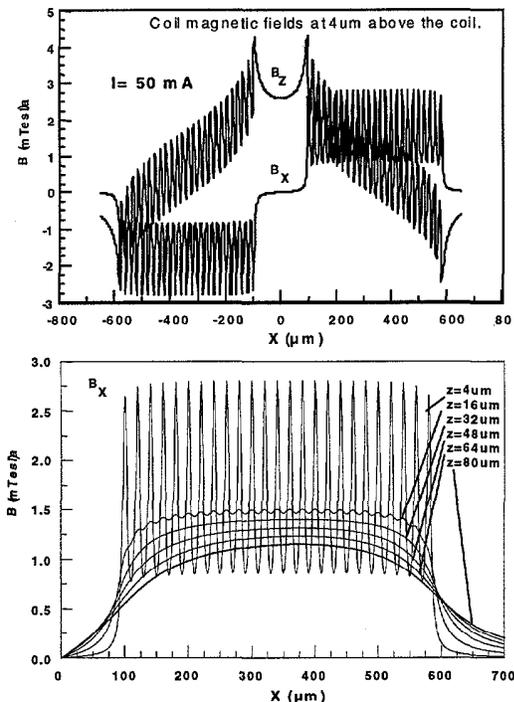

Figure 4. Calculated magnetic flux generated by a current of 50 mA though the planar coil. The x-component of the magnetic field is used to switch the magnetization orientation, and thus the relay. The oscillatory behaviors in B are due to the proximity gaps between the coil lines. The results are used to determine the necessary magnitude of switching current and the number of turns in the coil in order to switch the relay.

Consider the center-hinged type cantilever as an example to estimate the magnetic field strengths. We assume L=800 μm, w=200 μm, t=25 μm, the stage height (air gap spacing) h=12 μm, and the external magnetic field B_0=0.037 Tesla (equivalent to H_0= 370 Oersted). The key purpose is to verify that $B_{coil\_\xi} > B_{0\xi}$ when the cantilever is in the "close" position (the result also applies to the "open" case because of the symmetry). $B_{0\xi}$ depends on both B_0 and the angle α $[B_{0\xi}=B_0cos(\alpha)]$. When the cantilever is in the "close" position, we obtain α=180°-$arccos$(12/0.5L)=180°-$arccos$(12/400)= 91.72°. So, we have $B_{0\xi}$= -0.0011 Tesla. On the other hand, $B_{coil-\xi} \approx B_{coil-X} sin(\alpha) \approx B_{coil-X}$, and its average value is around 0.0014 Tesla at the height of 12 μm (Fig. 4). We can see that under these conditions, a 50 mA current in the planar coil is able to generate a magnetic field large enough to reverse the magnetization in the cantilever and switch the relay. The analysis applies to other cases where the supporting torsion flexure is not located exactly at the center of the cantilever, and the switching current can be different between the two switching events. Another point worth noting is that because the current-induced coil magnetic field only needs to overcome the ξ component of the external field (B_o) and not the Z component, the switching current can be relatively small.

(3) Fabrication

Conventional surface micromachining techniques were used to fabricate the latching relays. The device was fabricated on a Si-

substrate covered with an insulating dielectric. A Ag coil was formed by e-beam evaporation and wet etching. A polyimide layer was spin-cast on to cover the coils. Then bottom Au contact layers were deposited and patterned. Photoresist was used as a sacrificial layer. The cantilever was formed by electroplating NiFe permalloy on a Au seed layer. The cantilever was then released, and the relay diced and finally mounted on a permanent magnet.

RESULTS AND DISCUSSIONS

Latching micro relays of both the center-hinge and fixed-end type have successfully been demonstrated. A number of different beam dimensions - ranging from 10 μm ≤ w ≤ 600 μm, 80 μm ≤ L ≤ 1000 μm, 10 μm ≤ t ≤ 30 μm, and 12 μm ≤ h ≤ 18 μm - have been evaluated for both cantilever types. A 20-turn coil was used for all the devices.

Photographs of two types of micro magnetic relays (hinged and fixed-end) are shown in Fig. 2. Both types are fully functional; demonstrating both latching operation and electrical switching.

(1)   Resistance

Resistance was measured using the four-probe method. For a typical hinged-type relay (L=1000 μm, w=500 μm, t=30 μm, h=18 μm), the measured lead-to-lead resistance was 52.8 mΩ under a 410 Oersted permanent magnet field with a switching current of 118 mA.

(2)   Switching characteristics

Electrical switching characteristics are shown in Fig. 5. A waveform generator (HP33120A) was used to produce the coil current pulses for switching. A DC voltage of 3.2 V was applied to the relay with a load resistor (R_L=5 kΩ) in series. Both the input signal (V_{in}) and the relay signal (V_R on the load) were recorded on
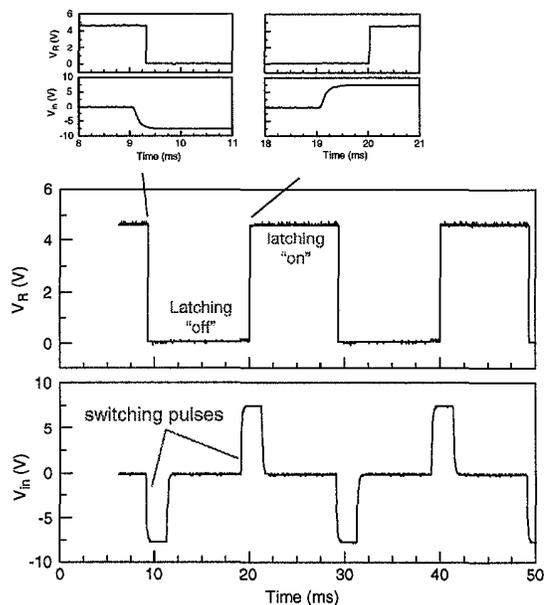

Figure 5. Typical switching characteristics of the latching relay. A square-wave-type current pulse (peak amplitude of 100 mA) was used to switch the relay. As can be seen, the relay latches after switching. The top enlarged plots show the on and off transient characteristics. Beam dimensions: 800 μm × 200 μm × 25 μm.

an HP Oscilloscope (HP54615B). The input pulse width was 3 ms (the shortest measured input pulse width is 0.2 ms) and the duty cycle was 20 ms. The magnitude of the driving pulse was 7.5 V, which corresponds to a current of 100 mA with a coil resistance of 75 Ω. As can be seen, a negative current pulse through the coil turns off the relay and a positive pulse turns on the relay. Between the pulses, the relay latches in a stable state without any power consumption. The turn-on and -off transient characteristics are shown in the enlarged plots in Fig. 5. Oscillations in the circuit are also observed (Fig. 6) and can be due to the parasitic capacitive and inductive elements in the device and/or measurement setup.
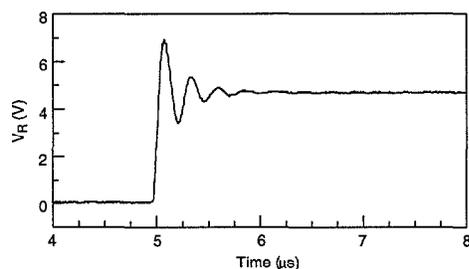

Figure 6. Oscillation behaviors are also observed at the turn-on. Parasitic LC elements in the measurement setup and/or the device may contribute to the oscillation.

(3) Reliability

The maximum DC current is >500 mA. Lifetime tests were performed on a center-hinged-type relay with the following dimensions: L=1000 μm, w=500 μm, t=30 μm, air gap h=12 μm. The permanent magnetic field strength was 370 Oersted and the switching current was 100 mA. No mechanical or electrical degradation (permanent contact resistance change) were observed after the relay was switched for 4.8×10^6 cycles in ambient conditions with an "on" current of 240 μA. Momentary resistance changes were occasionally observed during test and might have been due to particles in the air. Further evaluations are under way to better understand the long-term contact properties.

V. SUMMARY

This paper describes a new type of latching micro magnetic relay that has recently been demonstrated. The device is based on preferential magnetization of a permalloy cantilever in a permanent external magnetic field. Switching between two stable states is accomplished by a short current pulse through an integrated coil underneath the cantilever. Some key features are summarized as follows. (1) Latching (bistable); (2) low energy consumption during switching ($\leq$93 μJ, switching current ~ 60 mA, minimum switching pulse width ~ 0.2 ms); (3) low voltage operation $\leq$ 5 V); (4) maximum DC current > 500 mA; (5) capable of various switch configurations [single-pole-single-throw (SPST), multi-pole-single-throw (MPST), or multi-pole-double-throw (MPDT)]; (6) low contact resistance ($\leq$ 50 mΩ); (7) operation in ambient environment; (8) lifetime expected to be comparable to other micro relays, (9) batch fabrication using planar processing methods.

The most significant advantage of this relay is its bistability and consequent elimination of power consumption in the quiescent states. Other desirable properties (such as low contact resistance, low operation voltage, high off-state isolation, simplicity of design and fabrication, fast switching and large actuation force) also make the device a good candidate for applications in automatic testing equipment, smart interconnects, and other fields that require latching relays.

REFERENCES

[1] J. W. Judy and R. S. Muller, "Magnetically actuated, addressable microstructures," J. Microelectromechanical Systems **6**, 249 (1997).

[2] W. P. Taylor, O. Brand, and M. G. Allen, "Fully integrated magnetically actuated micromachined relays," IEEE J. Microelectromech. Syst. **7**, 181 (1998).

[3] J. A. Wright, Y.-C. Tai, and G. Lilienthal, "A magnetostatic MEMS switch for DC brushless motor commutation," Proceedings: Solid-State Sensor and Actuator Workshop 1998, Hilton Head Island, South Carolina, pp. 304-307 (1998).

[4] H. A. C. Tilmans, E. Fullin, H. Ziad, M. D. J. Van de Peer, J. Kesters, E. Van Geffen, J. Bergqvist, M. Pantus, K. Baert and F. Naso, "A fully-packaged electromagnetic microrelay", Proceed. 12th Ann. Int. Conf. MEMS, pp 25-30 (1999).

[5] P. P. Benham, R. J. Crawford, and C. G. Armstrong, *Mechanics of Engineering Materials*, 2nd edition (Longman, Essex, 1996), pp. 189.

[6] The Young's modulus for the permalloy was not available to the authors at the time of writing this paper. The value for nickel was therefore used. See *CRC Handbook of Chemistry and Physics*, 78th edition, pp. 12-204.

POLYSILICON XYLOPHONE BAR MAGNETOMETERS

Dennis K. Wickenden, John L. Champion, R. Ben Givens, Thomas J. Kistenmacher, Jay L. Lamb, and Robert Osiander

The Johns Hopkins University
Applied Physics Laboratory
Laurel, MD 20723

ABSTRACT

The availability of miniature orientation or spin sensors capable of being mounted directly on a projectile would greatly simplify testing and evaluation procedures. It would also allow for the development of smart munitions by providing essential fuzing and guidance functions. One of the simplest orientation sensors is a magnetometer. The recently developed Johns Hopkins University Applied Physics Laboratory xylophone magnetometer is shown to be suitable for such applications. The xylophone magnetometer, based on a free-free resonating bar and utilizing the Lorentz force to measure vector magnetic fields, is intrinsically linear and has a wide dynamic range, measuring field strengths from nanoteslas to teslas. Furthermore, since its sensitivity, for a constant Lorentz force, is independent of size for resonating bars of the same material and aspect ratio, it is ideally suited for miniaturization. Various polysilicon xylophone bars have been designed, fabricated at the Cronos MUMPs foundry, and characterized. The collective output response has verified the size-independent scaling law and sensitivities of order 100 nT have been achieved with drive currents as low as 20 µA. This drive current is limited by the sheet resistance of polysilicon and directly affects the sensitivity and dynamic range. Aspects of the various designs for polysilicon xylophone bar magnetometers are described, together with experimental results and projections for improving sensitivity and dynamic range.

INTRODUCTION

Typical techniques for measuring the orientation or spin rate of various types of munitions during the free flight phase of their test and evaluation include spin cards, high-speed video, spark range photographs, x-rays, and radar. All of these techniques are complex, expensive to implement, require special test ranges, and do not readily provide continuous recording [1]. The availability of miniature orientation sensors capable of being mounted directly on the munition would greatly simplify test and evaluation procedures and would also expand their application in munition systems. For example, they could serve two important functions. The first, as a compass, would measure the direction of travel of the munition with respect to the local Earth's magnetic field, and aid in its guidance. The second, as an angular rate sensor, would have fuzing applications. For example, the spin-up that occurs as the munition is fired could trigger an arming switch to activate a fuze. Alternatively, it could act as a timing device to air burst the munition after traveling a predetermined number of revolutions. A processor, in combination with the magnetometer, could be programmed to keep track of the number of revolutions the munition makes as it travels down range. If the munition's velocity is known, then the number of revolutions could be set to initiate the burst at the desired distance.

One of the simplest orientation or spin sensors is a magnetometer. In either application, the sensor measures the vector components of the Earth's magnetic field as a function of time. A multiple sensor device is needed to avoid nulls that will occur when the sensitive axis aligns with magnetic field lines. To date, the use of magnetometers to assist with the determination of a projectile's motion or heading has been severely limited. Two obvious reasons are the low sensitivity of standard compasses and the size and weight constraints of standard magnetometers. The multi-sensor requirement leads naturally to the necessity of further miniaturization and microelectromechanical system (MEMS)-based implementation. The goal of developing MEMS-based magnetometers has lead to the development of other new devices such as piezoresistive [2] and magnetostrictive [3] cantilevers, magnetoresistive films [4], and magnetometers based on electron tunneling effects [5]. While these methods can be used to create miniature magnetometers, their sensitivity has not reached the nanotesla range required for precision instruments. This article describes the development of MEMS-based devices based on the Johns Hopkins University Applied Physics Laboratory's xylophone bar magnetometer [6], with particular emphasis on their use in projectile applications.

XYLOPHONE PRINCIPLES

The resonating xylophone bar magnetometer uses the Lorentz force as a measure of the magnetic field vector. It offers the potential of being lightweight and operating in a low power mode. True to its name, the device is based on a classical resonating xylophone bar. The principle of operation is shown in Figure 1.
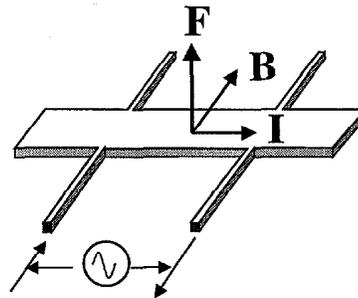

Figure 1 Principle of operation of xylophone bar magnetometer

The active element is a conductive bar supported at the nodes of its fundamental mode, of frequency f_o, of mechanical vibration. An alternating sinusoidal current $[I\cos(2\pi f t)]$ is fed along the bar through the supports. In the absence of a magnetic field the bar will remain stationary.

In the presence of a magnetic field, B, the Lorentz force F, normal to the bar and the drive current, is given by:

$$F = I \times B \cos(2\pi f t)$$ (1)

This force causes the xylophone bar to vibrate, with a displacement, d, at the mid-point of the bar being given by:

$$d = \frac{d_{dc}}{\sqrt{\left[1 - \left(\frac{f}{f_0}\right)^2\right]^2 + \left(\frac{f}{Qf_0}\right)^2}}$$ (2)

The static deflection of a beam, of length l, width a, and thickness b, simply supported at both ends due to the Lorentz force uniformly distributed along its length is given by [7]:

$$d_{dc} = \frac{5Fl_2^4}{384EI_B}$$ (3)

where I_B = area moment of inertia = $ab^3/12$, E = Young's modulus, and l_2 is the distance between the supports.

The phase angle between the driving current and the deflection is given by:

$$\phi = \arctan\left(\frac{f \cdot f_0}{Q\left(f_0^2 - f^2\right)}\right)$$ (4)

In the case of an ideal free-free bar, where the effect of the supports can be neglected, the resonance frequency of the fundamental mode of the bar is given by:

$$f_0 = \frac{22.4}{2\pi}\sqrt{\frac{EI_B}{\rho abl^4}}$$ (5)

It can be seen from these equations that at resonance ($f = f_0$), the deflection of the xylophone bar is linearly proportional to the amplitude of the drive current, I, the mechanical Q of the structure, and the magnetic flux density, B. Furthermore, since the xylophone bar includes no magnetic material, the sensor does not saturate. This results in a magnetometer with an extremely wide linear dynamic range. Macro-devices, chemically milled form Cu-Be foils, with dimensions of 5 mm x 0.5 mm x 0.025 mm, Q factors in the range 1000 - 5000, and drive currents of the order of one amp, have demonstrated dynamic ranges in excess of 80 dB and a noise floor below 0.1 nT/$\sqrt{\text{Hz}}$ [8].

RESULTS

The xylophone bar magnetometer design readily lends itself to miniaturization, and importantly, as equation (3) shows, that a scaling of all dimensions (length, thickness, and width) does not

change the sensitivity (deflection) of the xylophone bar for a constant Lorentz force. Several generations of MEMS- processed polysilicon magnetometers have been designed and fabricated using Multi-User MEMS Processes (MUMPs) foundry service at MCNC. A schematic diagram of a typical device is shown in Figure 2. The lower electrode consists of a 0.5-μm thick
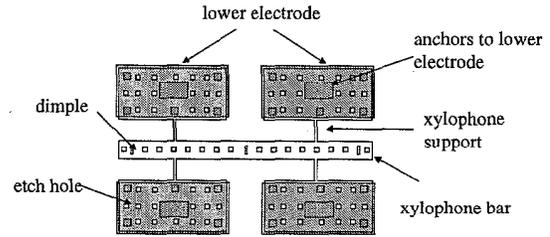

Figure 2 *A schematic diagram of a typical device*

polysilicon layer patterned on the silicon nitride-coated silicon substrate. The xylophone bar, support electrodes, and mounting pads are fabricated from 2 μm thick polysilicon suspended (after release of the sacrificial silicon dioxide) 2 μm above the nitride layer or lower electrode. The mounting pads are attached to the lower electrode by indicated patterns to provide a rigid anchor that minimizes vibrational coupling with the xylophone bar [9]. Contact between the xylophone bar and the nitride layer is minimized by having 0.75 μm deep "dimples" in the center and toward the ends of the xylophone bar. Lastly, etch holes are incorporated to facilitate a clean release of the completed structure. An SEM micrograph of a packaged 500 x 50-μm xylophone bar is presented in Figure 3. The micrograph clearly displays the anchor pattern, the support arms, the mounting pads, and the topographical evidence of the dimples. It also demonstrates the clean release of the xylophone structure.
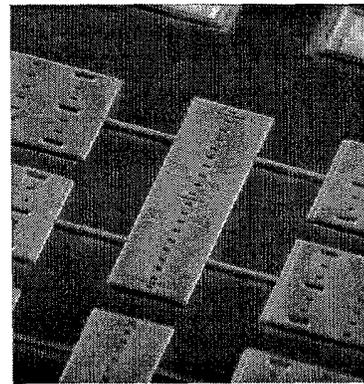
Figure 3 *SEM micrograph of a packaged 500 x 50-μm xylophone bar*

Evaluation of these polysilicon devices has been undertaken in a vacuum test chamber using an optical beam deflection setup. All devices tested to date have worked very well in static magnetic fields and exhibit mechanical Q-factors in excess of 20,000 at reduced pressures of 15 – 20 mTorr.

Figure 4 illustrates results obtained by scanning the frequency of the drive current through a 500 x 50 μm xylophone bar with 45 μm long x 4 μm wide supports operating in a static magnetic field. The rms current through the device was 22 μA, the magnetic flux

density was on the order of mT (from a nearby permanent magnet) and the test chamber pressure was 35 mTorr. From this data, the resonance frequency was found to be 78.15 kHz. The predicted frequency from equation (5) is 68.4 kHz.
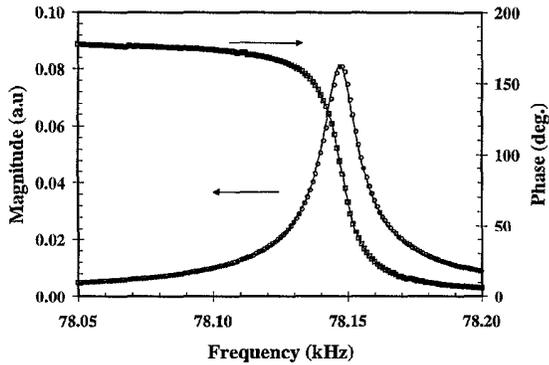
Figure 4 Frequency response of 500x50-μm xylophone with 4-μm wide support arms

The difference between the observed and predicted resonance frequency was attributed to the increased torsional stiffness of the xylophone support arms. This was confirmed by a detailed analysis of the mechanical properties of the structures. In this analysis, each of the support arms was modeled as a combination of vertical and torsional springs (K_v and K_r, respectively) and a concentrated mass (M); and, the resonator was modeled as a beam using Bernoulli-Euler theory. Use of this theory, rather than the more complex Timoshenko theory, was justified because the wavelength of vibration is much larger than the thickness of the beam. The differential equation governing the motion of a Bernoulli-Euler beam undergoing harmonic motion $w(x,t)$ at an angular frequency ω is given by:

$$EI_B w^{iv} + \omega^2 \rho A w = 0 \qquad (6)$$

The effects of the two springs and concentrated mass enter into the expressions for the boundary conditions. Here, the boundaries include both ends of the beam as well as the support points. The zero-moment and zero-shear boundary conditions are used at both of the free ends. The motion of the beam is restrained to some degree in both displacement and rotation by the two elastic springs. The concentrated mass also adds inertia to the beam. Because the springs and mass are attached to the nodal points, neither K_v nor the mass M has any influence on the first bending mode of vibration. Only the torsional spring K_r has an influence on this mode and results in the observed functional dependence of resonance frequency on support arm width, as illustrated in Figure 5. In this figure the solid line denotes the results of the analysis and the open circles represent the experimental results from 500x50x2 μm xylophone bars with different support arm widths. The agreement is excellent and well within the manufacturing tolerances of the MEMS processing.

A plot of the output response of the polysilicon xylophone magnetometer with 4 μm support arm widths as a function of impressed magnetic flux density up to 150 μT is shown in Figure 6. The superimposed line is a least-squares fit to the data, and it is apparent that the magnetometer response is linear over the range of magnetic flux density shown. An estimate of the ultimate noise
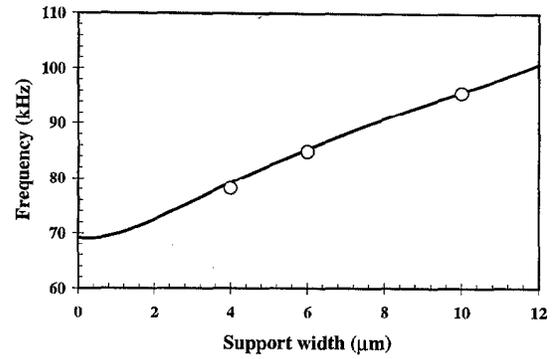

Figure 5 Resonance frequency as a function of support arm width for 500x50-μm polysilicon xylophone 2 μm thick

floor of the xylophone magnetometer, based on thermomechanical noise (Brownian motion due to the Langevin force) and, to a lesser extent, Johnson noise (current noise that couples with the magnetic field), yields a value of the order of 100 pT•A/√Hz. This estimate calls out the very important role that the amplitude of the drive current plays in determining the ultimate field sensitivity of a xylophone bar magnetometer.
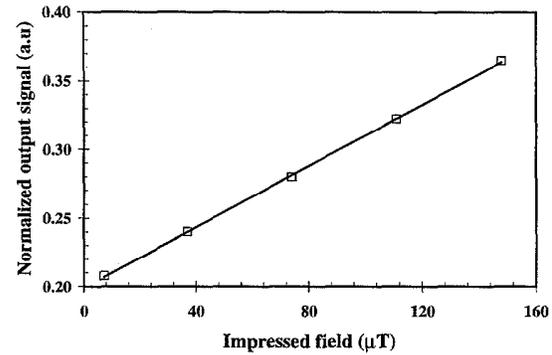
Figure 6. Output response of 500 x 50-μm xylophone magnetometer as a function of impressed magnetic flux density.

The length and width dimensions of the polysilicon xylophone bar magnetometer presented here are proportionally scaled from the Cu-Be devices described above, while the thickness is an order of magnitude smaller than for direct scaling. The decreased sensitivity is identified as being a direct result of the high sheet resistivity of the polysilicon material (10 Ω/square) and its lower current carrying capability.

Current work is concentrated on reducing the resistance of the polysilicon devices. Approaches being adopted include feeding the drive current to the magnetometer from both sides of the xylophone bar, giving an immediate doubling of the drive current capability and improving the thermal symmetry of the device. In addition, magnetometers have been manufactured from 3.5-μm thick poly1/poly2 layers, and the use of gold stripes deposited on various parts of the xylophone has been examined. An SEM micrograph of a released and packaged poly2 devices with gold deposited on the electrodes, the support arms, and as stripes along the xylophone bar between the support arms, is reproduced in Figure 7. Poly1/poly2 500 x 50 x 3.5-μm bars with gold deposited on only the electrodes have resistances of approximately 150 Ω, resonance frequencies of 165 kHz and Q-factors of between 30,000 and 40,000 at low
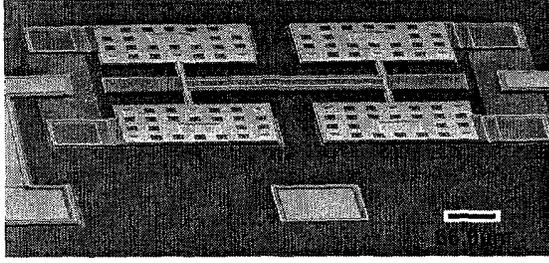
Figure 7. SEM micrograph of xylophone bar with the drive current fed from both sides of the xylophone and with gold deposited on the electrodes, the support arms, and along the bar.

pressures. The same bars with full gold strips on the support arms have resistances of approximately 80 Ω, increased resonance frequencies of 168 kHz, due to the increased stiffness of the support arms, and Q-factors of 10,000. The bars with the gold strips along the xylophone bar have resistances of approximately 40 Ω, decreased resonance frequencies of approximately 164 kHz, due either to mass loading of the bar or to the decreased effective Young's modulus, and Q-factors in excess of 10,000. Investigations are continuing in order to determine whether the drop in Q-factor observed when gold is deposited on the support arms or along the bar is a real effect or not.

CONCLUSIONS

Polysilicon devices based on the Johns Hopkins University Applied Physics Laboratory's xylophone bar magnetometer have been shown to be suitable for use as sensitive orientation or spin sensors for projectile applications. Such sensors will not only greatly simplify the test and evaluation of projectiles undergoing current development but could also lead to the appearance of smarter future munitions. 500x50x2-μm devices have confirmed the scaling laws expected for xylophone bar magnetometers and have a predicted ultimate noise floor of the order of 100 pT•A/$\sqrt{Hz}$. The drive current capable of being passed by the polysilicon limits the sensitivity. The drive current can be increased by the use of gold overlays but early indications are that these techniques may degrade the mechanical Q-factors.

ACKNOWLEDGEMENTS

REFERENCES

1. B. S. Davis, T. Harkins, and L. Burke, "Flight test results of miniature, low-cost, spin, accelerometer, and yaw sensors," *AIAA 35th Aerospace Sciences Meeting and Exhibit*, AIAA 97-0422, Reno, NV, 1997.

2. C. Rossel, P. Bauer, D. Zech, J. Hoder, M. Willemin, and H. Keller, "Active microlevers as miniature torque magnetometers," *J. Appl. Phys.*, **79**, 8166-8173 (1996).

3. R. Osiander, S. A. Ecelberger, R. B. Givens, D. K. Wickenden, J. C. Murphy, and T.J. Kistenmacher, "A microelectromechanical-based magnetostrictive magnetometer," *Appl. Phys. Lett.*, **69**, 2930-2931 (1996).

4. F. Y. Yang, K. Liu, K. Hong, D. H. Reich, P. C. Searson, and C. L. Chien, "Large magnetoresistance and field sensing characteristics of electro-deposited single-crystal bismuth thin films," *Science* **284**, 1335-1337 (1999).

5. L. M. Miller, J. A. Podosek, E. Kruglick, T. W. Kenny, J. A. Kovacic, and W. J. Kaiser, "A μ-magnetometer based on electron tunneling," *Proc. IEEE Workshop on Micro Electro Mechanical Systems*, (IEEE, New York) pp. 467-472 (1996).

6. R. B. Givens, J. C. Murphy, R. Osiander, T. J. Kistenmacher, and D. K. Wickenden, "A high sensitivity, wide dynamic range magnetometer designed on a xylophone resonator," *Appl. Phys. Lett.*, **69**, 2755-2757 (1996).

7. W. C. Young, *Roark's Formulas for Stress and Strain*, 6th Edition, McGraw-Hill, New York, 1989.

8. R. B. Givens, D. K. Wickenden, D. A. Oursler, R. Osiander, J. L. Champion, and T. J. Kistenmacher, "Heterodyne detection of alternating magnetic fields with a resonating xylophone bar magnetometer," *Appl. Phys. Lett.*, **74**, 1472-1474 (1999).

9. D. Sherman, "*An Investigation of MEMS Anchor Design for Optimal Stiffness and Damping*," Master's Project, University of California, Berkeley, May 1996.

MONOLITHIC INTEGRATION OF GaAs SAW CHEMICAL MICROSENSOR ARRAYS AND DETECTION ELECTRONICS

Stephen A. Casalnuovo, Vincent M. Hietala, Edwin J. Heller, Gregory C. Frye-Mason, Albert G. Baca, and Joel R. Wendt

Sandia National Laboratories
Albuquerque, New Mexico 87185-1425

ABSTRACT

We describe the integration of an array of surface acoustic wave delay line chemical sensors with the associated RF microelectronics such that the resulting device operates in a DC in/DC out mode. The microelectronics design for on-chip RF generation and detection is presented. Both hybrid and monolithic approaches are discussed. This approach improves system performance, simplifies packaging and assembly, and significantly reduces overall system size. The array design can be readily scaled to include a large number of sensors.

INTRODUCTION

In many of chemical sensing applications, sensor size is a critical system constraint, where smaller size generally translates into lighter weight, smaller sample volumes, lower power consumption, and greater portability. Small sensors are also more readily arrayed into a system with a high degree of chemical discrimination, capable of operation in the complex chemical backgrounds found in real-world environments. However, a reduction in the size of the chemical transducer is not sufficient to reduce the overall sensor system size significantly, since system size is often determined by the control electronics, electrical interconnections, and packaging. To bring about a significant reduction in system size, these other components must be considered as well. For chemical transducers that can be produced on semiconductor substrates, the monolithic integration of the transducer with its control electronics, through the use of integrated circuit (IC) microfabrication technology, can bring about the desired size reduction. Other benefits as well often accrue with monolithic integration, including enhanced manufacturability and reliability, minimized temperature dependence, simplified packaging and assembly, and reduced cost.

Gallium arsenide (GaAs) is an intriguing semiconductor material for the particular case of acoustic wave chemical sensors, when viewed in terms of integration. GaAs is intrinsically piezoelectric, which permits electrical generation of acoustic waves. A mature GaAs IC technology also exists, with commercial foundries able to provide application-specific integrated circuits (ASICs) from a customer's design. Coincidentally, this IC industry has driven two developments critical to integrated acoustic wave sensors: high purity, low defect density, crystalline GaAs substrates suitable for acoustic wave generation and high frequency analog ICs required to generate and detect the acoustic signals. Thus, the foundational technologies required for integrated GaAs acoustic wave chemical sensors are in place.

Acoustic wave chemical sensors operate at high frequencies, typically between 100 MHz and 500 MHz, because chemical sensitivity generally increases with acoustic frequency [1]. This frequency range underscores the importance of monolithic integration for this particular application. In addition to the advantages listed above, a device that contains all high frequency electronic and acoustic components on a single substrate is expected to have improved performance because of a reduction in losses associated with transferring high frequency signals between discrete components. GaAs is particularly well suited for high frequency microelectronics integration, making complete direct current (DC) -in, DC-out operation on a single chip a realizable goal.

SURFACE ACOUSTIC WAVE SENSOR DESIGN

As shown in Figure 1, surface acoustic wave (SAW) delay line sensors consist of a piezoelectric substrate and two interdigitated transducers (IDTs) formed by photolithographic patterning of a thin metal layer. Application of an alternating voltage to the input transducer generates an alternating strain field that launches a surface acoustic wave. The acoustic wave travels along the substrate surface before being converted back into an electrical signal by the output transducer. The time delay resulting from the transit of the acoustic wave between the IDTs gives rise to the name of these devices.
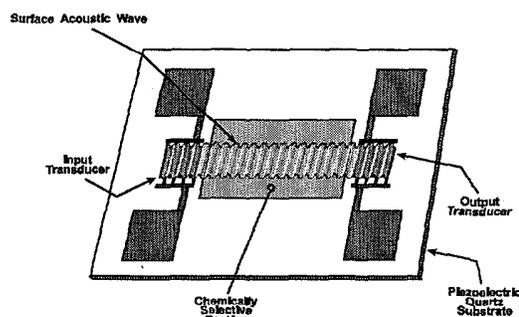
Figure 1. Schematic of a SAW delay line sensor.

The velocity and attenuation of the propagating wave are very sensitive to properties such as mass, temperature, and viscoelasticity of thin films formed on the device surface. For example, increases in surface mass loading decrease the SAW velocity. This property can be used to detect picogram mass changes [1]. By coating the acoustic path with a material that sorbs a chemical species of interest, this sensitivity can be used to develop chemical sensors [1-3]. When configured in an oscillator circuit, as shown in Figure 2, changes in the delay line acoustic velocity, ΔV_R, which is equivalent to a phase change, $\Delta\theta$, can be observed as a frequency shift, Δf. This oscillator approach converts a very small velocity change into a measurable frequency increment, which can be determined with parts per billion resolution. In the limit of small perturbation, the response to these surface changes can be expressed as:

$$\frac{\Delta f}{f_o} = \frac{\Delta\theta}{\theta_o} = \frac{\Delta V_R}{V_{Ro}} = -k_m\frac{\Delta m}{m_o} - k_T\frac{\Delta T}{T_o}, \qquad (1)$$

in which m is the surface mass, T is the temperature, k_m is the mass sensitivity and k_T is the temperature sensitivity.

approach has converted the velocity change into a voltage change. This allows for easy measurement with low-cost electronics, but also undoubtedly loses resolution compared to a time measurement.

Figure 3. Block diagram of a dual-channel (3-port) SAW sensor system.

This design compensates for changes in the sensor temperature. Since the two acoustic paths are on the same substrate, they can be assumed to be at exactly the same temperature. Therefore, referring back to Equation (1) and assuming that only the test channel only experiences the mass uptake, the temperature effects will exactly cancel and only the mass uptake of the test channel will be present at the output. This compensation is almost perfect except for the slight change in mass sensitivity as the reference channel center frequency changes over temperature. Additionally, this approach is tolerant to "mode-hopping" (changes in the longitudinal mode number) due to the subtractive nature of the technique. In practice, the necessary application of the chemically selective thin film to the test channel will change the temperature coefficient of the test channel and compromise the complete elimination of the temperature terms in the subtraction. This unfortunate effect will be an issue for any measurement approach and substrate material.

SAW ARRAY ELECTRONICS DESIGN

The dual-channel temperature compensated SAW sensor approach is readily extended to large arrays of acoustic sensors. A schematic of a four-element SAW delay line array configuration is shown in Figure 4. A large bi-directional central transmitting IDT spans each of four smaller receiving transducers. Appropriate chemically sorbent coatings are applied to three of the regions between the transmitting IDT and the receiving IDTs, with the fourth left uncoated to provide a reference phase for the array. As in the dual sensor system, an oscillator circuit drives the transmitter, launching an acoustic wave to the receivers. Phase comparator circuitry measures the relative phase of each of the receivers with respect to the reference and provides a DC output voltage proportional to the relative phase. The block diagram of a four-channel system, including electronic circuitry, is shown in Figure 5. Extension to n-channels simply requires splitting the reference signal 2n-1 ways (n inputs for each acoustic channel and mixer drive for n-1 channels).

The electronics diagramed in Figure 5 incorporate several novel design features. All signals represented by single lines in this figure are implemented differentially in the actual design. The amplifier blocks shown as triangles are specially designed limiting differential amplifiers providing small phase errors at large over-drive conditions. The amplifiers also provide a log output signal, which allows for an accurate power measurement of the signals from each SAW output port. This permits measurement of the insertion loss of each SAW channel. The mixers are implemented as Gilbert cell mixers, which homodyne the test signal to DC. The input signals to the mixers are hard limited (square waves) by the


Figure 2. Schematic of a 470 MHz SAW oscillator circuit.

GaAs SAW delay lines are fabricated on (100) semi-insulating substrates with the delay lines oriented in the [011] direction. These substrates are standard for microelectronics applications and the [011] direction maximizes the coupling between electromagnetic and acoustic waves. These delay lines operate between 100 MHz and 500 MHz, where the pitch of the IDT finger pairs determines the frequency of the device. Typically, each IDT contains 50 finger pairs. The input and output IDTs are separated by 190 acoustic wavelengths and are 30 wavelengths wide. IDTs consist of a 500 Å to 1000 Å thick gold layer (depending on device frequency) on top of a 200 Å thick titanium adhesion layer, deposited by thermal evaporation.

SAW SENSOR ARRAY

Because of the limited selectivity of the sorbing coatings applied to SAW chemical sensors, a single SAW sensor is insufficient to identify analytes in a realistic background of interferants. In practice, an array of sensors, each with a different selective coating, is used to make reliable chemical identification. An array of sensors also allows redundancy and error-checking, important advantages in the overall system. Multiple sensors are also commonly used to compensate for undesired sensor responses. In a classic dual sensor system, two sensors are maintained in identical environments except one is introduced to the desired stimulus. The resulting responses (frequencies) can then be subtracted to allow for removal of the undesired background responses due to temperature or other surface changes.

Figure 3 shows a block diagram of a novel dual-sensor approach that allows for integration and can be scaled to larger array size. In this system, phase is compared (subtracted) rather than frequency. One SAW device (Ch R) is used as a reference channel and is assumed isolated from the surface stimulus of interest. The reference oscillator drives a second acoustic channel (termed test channel, Ch T) on the same substrate. Note the unique use of the acoustic wave propagating in either direction from the center IDT to provided the necessary power split of the reference and test channels. Multiplying the test signal with the reference signal and observing the DC component of the result performs the phase difference. A parasitic second harmonic is present at the phase output and is filtered off. It should be noted that this

amplifiers so in effect the mixer operates digitally, performing the exclusive-or function. This provides linear phase detection and provides two-quadrant phase detection as shown in Figure 6. This detection response provides a unique output if the two signals' phase difference remains within the appropriate two quadrants of phase space (1 and 2 or 3 and 4).
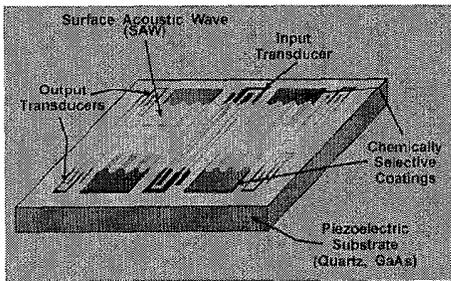

*Figure 4. Schematic of an array of four surface acoustic wave devices on a single substrate.*
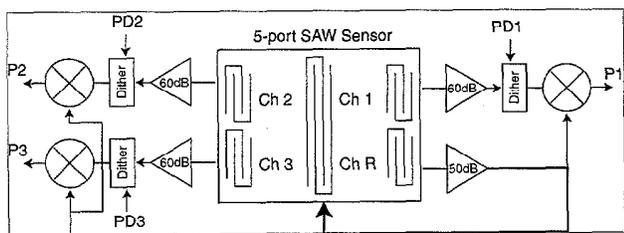
Figure 5. Block diagram of a four-channel (5-port) SAW sensor system.
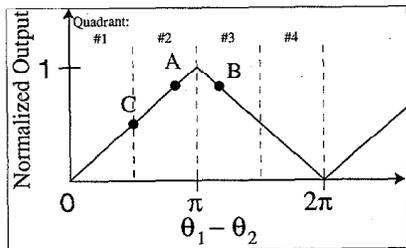

Figure 6. Diagram of the normalized DC component from the Gilbert cell mixer.

To achieve the desired full four-quadrant phase detector, the phase "dither" circuits as shown in the block diagram were added. The phase dither circuits allow for a given test channel to be selectively advanced or retarded in phase by a fixed predetermined amount. Inputs are provided to cycle through three phase states to determine the phase uniquely over the complete 2π phase space. Operation of the dither sequence can be illustrated by an example shown by the three labeled points in Figure 6. Assume the actual phase is at the point A. From a single DC output measurement, one cannot determine if the actual phase is in quadrant #2 or quadrant #3 (normalized output is the same at points A and B). Using the phase dither, the response curve can be "shifted" to determine the "slope" of the output and hence the correct quadrant. Assuming that the dither shifts the output from point A to point C, the negative slope indicates the actual point is in quadrant #2. Three dither states are provided to handle the event where the measured and dithered points straddle a quadrant boundary (as with points A and B in Figure 6). In general all three dither states must be considered to guarantee unique phase detection over all phase space.

HYBRID FOUR-ELEMENT SAW ARRAY

An array of SAW sensors and the associated microelectronics was first produced using a hybrid packaging approach rather than through monolithic integration [4]. Although it requires more assembly than the monolithic approach, the hybrid device prototypes can be produced more rapidly and can be used to test the sensor design and circuit functionality. It also represents a significant size reduction compared to conventional SAW sensor systems.

The SAW array is fabricated on a ST-quartz substrate. In this case, quartz was selected because the requirement for monolithic integration was removed and quartz has slightly better acoustic performance than GaAs. The oscillator amplifier and phase comparator circuitry are custom GaAs ASICs attached directly to the quartz substrate. Metal paths patterned directly onto the quartz die provide circuit interconnection. Wire bonding is used to connect the ICs to the metal paths on the quartz. Although not monolithically integrated, this device incorporates all high frequency components on the quartz substrate so that the packaged part operates in the desired DC in/DC out mode. A photograph of the multi-chip SAW array configuration is shown in Figure 7. The hybrid device requires 90 mA at 2.5 V to operate the GaAs ASICs, which can readily be provided by batteries. The SAW oscillator operates at 510 MHz and the phase comparators have a sensitivity of 1 V per 180 degrees of phase.

These devices have been tested as chemical sensors. The response of the SAW array to dimethyl methyl phosphonate (DMMP) is shown in Figure 15. The varied response of the 3 different coatings indicates their relative sensitivity to DMMP and demonstrates how the array can be used to discriminate among a number of analytes and interferants.
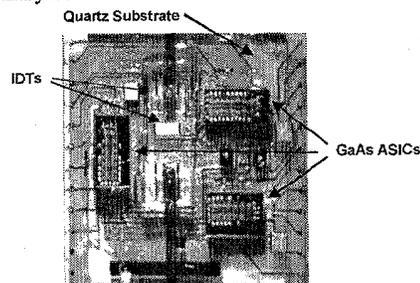

Figure 7. Hybrid version of SAW microsensor array. The ST-quartz die size is 6.9 mm by 8.6 mm. Each of the GaAs ICs is approximately 1 mm x 2 mm. The device operates in a DC in/DC out mode.
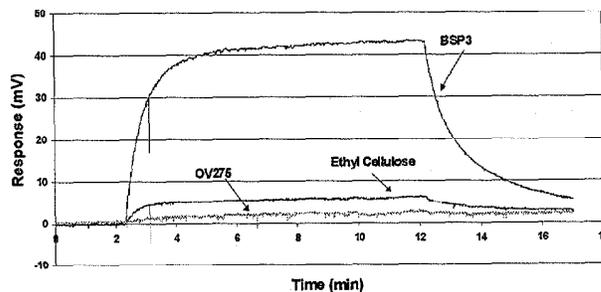

Figure 8. Hybrid SAW array response to 15 parts per million of DMMP. The three microsensor coatings were BSP3 hydrogen-bond acid, ethyl cellulose, and OV-275 (a cyano-modified polysiloxane).

MONOLITHIC SAW SENSOR ARRAY

We have completed fabrication of the first fully monolithic version of the SAW sensor array described above. This device is based on a GaAs SAW array of the same configuration shown in Figures 4 and 5 and uses the same GaAs microelectronic circuitry as the die shown in Figure 7, but puts all these components onto a single GaAs substrate. Figure 9 (a) shows one of the four-element arrays with the four delay lines arranged diagonally across the center of the die and the amplifier and phase comparator circuitry placed in the corners. The die measures 4.6 mm by 4.6 mm. The microelectronics on the monolithic devices are fabricated at a commercial GaAs IC foundry. The partially processed wafers are transferred to our laboratory for IDT fabrication. Prior to IDT fabrication, a plasma etch is used to remove several microns of dielectric material to expose the GaAs surface. The IDTs are then patterned on the GaAs surface using a metal lift-off process. Figure 9 (b) shows the frequency spectrum of the delay line oscillator portion of the device, which operates at approximately 692.5 MHz and draws 28.5 mA at 3 V_{dc}. Oscillator function demonstrates that the post-processing of the IDTs is compatible with the microelectronics fabrication process. The temperature dependence of the phase comparator output has been tested and this is insensitive to temperature, as expected. These devices are now being packaged for further electronic and chemical testing.


(a)   *(b)*
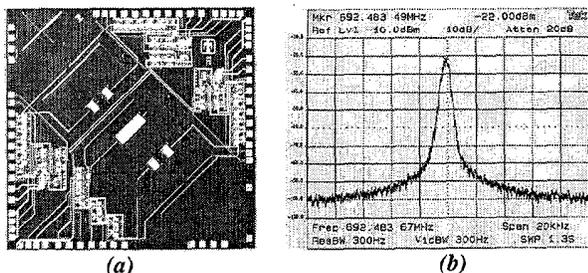
Figure 9. Monolithic GaAs SAW sensor array. (a) Photograph of die showing IDTs in the center and microelectronics in the corners. Die size is 4.6 mm x 4.6 mm. (b) Frequency spectrum of integrated oscillator, demonstrating successful delay line integration.

SUMMARY

We have described the development of microfabricated SAW chemical sensor arrays, culminating in a fully monolithic device that integrates all high frequency components onto a single GaAs substrate. These arrays operate in a DC in/DC out mode that simplifies assembly and improves performance. Hybrid arrays show good sensitivity to chemical analytes. Monolithic arrays have shown electronic functionality and are undergoing chemical tests.

ACKNOWLEDGEMENTS

REFERENCES

1. D. S. Ballantine, Jr., R. M. White, S. J. Martin, A. J. Ricco, E. T. Zellers, G. C. Frye, and H. Wohltjen, *Acoustic Wave Sensors Theory, Design, and Physico-Chemical Applications,* Academic Press, New York, (1997).

2. G. C. Frye, D. S. Blair, T. W. Schneider, C. D. Mowry, C. C. Colburn, and R. P. Donovan, "Development and Evaluation of On-Line Detection Techniques for Polar Organics in Ultrapure Water", *Journal of the Institute Of Environmental Sciences,* 39, 30 (1996).

3. J. W. Grate, S. L. Rosepehrsson, D. L. Venezky, M. Klusty, H. Wohltjen, "Smart Sensor System for Trace Organophosphorous and Organosulfur Vapor Detection Employing a Temperature-Controlled Array of Surface-Acoustic-Wave Sensors, Automated Sample Preconcentration, and Pattern-Recognition", *Analytical Chemistry,* 65, 1868 (1993).

4. E. J. Heller, V. M. Hietala, R. J. Kottenstette, R. P. Manginell, C. M. Matzke, P. R. Lewis, S. A. Casalnuovo, and G. C. Frye-Mason, "An Integrated Surface Acoustic Wave-Based Chemical Microsensor Array for Gas-Phase Chemical Analysis Microsystems," *Chemical Sensors IV,* Electrochemical Society, 99-23, (1999), pp. 138 - 142.

SIGNAL PROCESSING HIERARCHIES FOR
PORTABLE, LOW-POWER SAW-BASED CHEMICAL SENSING SYSTEMS

Denise M. Wilson
Department of Electrical Engineering, University of Washington
Box 352500, Seattle, Washington 98195-2500

Thaddeus A. Roppel
Department of Electrical and Computer Engineering, Auburn University
Auburn, Alabama 36849

ABSTRACT

This paper presents a comparison of multi-stage linear and single-stage nonlinear processing techniques for accomplishing chemical discrimination. Data variance, computational overhead, and memory storage requirements are compared between (linear) multistage principal components analysis, non-linear VERI (visually empirical region of influence) and non-linear artificial neural network techniques. For the linear techniques, data variance is reduced by 88%, compared to that of raw data during data preprocessing. Computational overhead is reduced up to 82.5% and 77% for non-linear clustering and artificial neural network techniques respectively. These improvements offer clear promise for significant reduction in power and space consumption for portable chemical sensing systems design.

INTRODUCTION

In the last decade, linear signal processing techniques for chemical sensor arrays have largely been reserved for preliminary analysis of array performance, for array optimization, or for simple sensing applications involving few analytes or few sensors (e.g. [1]-[3]). Non-linear techniques, such as artificial neural networks(e.g. [4]-[7]), non-linear clustering (e.g. [8]), genetic algorithms (e.g. [9]), fuzzy logic (e.g. [10][11]), and similar techniques. Principal component and cluster analysis, in particular, is a linear processing technique that has frequently been used to quickly evaluate, visually and quantitatively, the ability of an array to separate characteristics of individual chemicals. However, as the number of analytes associated with a particular application increases, principal components analysis quickly loses its effectiveness as cluster overlap increases, making discrimination difficult and reducing accuracy. In many cases, non-linear signal processing techniques have proven effective in solving discrimination problems that simply cannot be solved using linear processing techniques. However, most effort in both non-linear and linear signal processing targeted at chemical discrimination has been done without data preprocessing and in a single stage of processing. In this paper, we intend to support the argument that data preprocessing and multi-stage analysis can reduce superfluous information and optimize computation capability sufficiently to allow linear processing techniques to effectively solve discrimination problems.

The inherent limitations of linear signal processing are also what make them attractive to portable sensing systems. Linear signal processing, especially principal components analysis, is often conducive to hardware implementation and requires relatively few floating point operations compared to nonlinear signal processing techniques. Fewer floating point operations translate directly to decreased computational overhead and increased speed of operation. Data compression, minimal memory requirements, and low computational overhead are essential for the design of portable sensing systems that meet user-defined weight, battery-life, and size constraints.

As a proof of concept, this paper presents linear signal processing techniques for implementation in portable, low-power SAW-based chemical sensing systems. This paper provides a demonstration of how three-stage signal processing involving a combination of principal components analysis (PCA) and linear signal preprocessing can provide performance for chemical discrimination comparable to more complex, non-linear analysis techniques, but at significantly reduced numbers of floating point operations.

THEORY

A linearly separable problem, whether two-dimensional or thirty-dimensional, can be solved using linear signal processing techniques. The issue in determining linear separability, however, lies in the ability to visualize the problem, which is often not possible in dimensions greater than three or arrays that contain more than three elements. A number of software programs are available to assist in examining multi-dimensional data, but are limited in their ability to provide all possible perspectives of the data and, in addition, can only provide one (or few) of many perspectives at a given moment in time. For this reason, it is often more straightforward to allow a non-linear signal processing method to determine the separability of the data using an algorithm that has been proven mathematically to converge for a solvable problem. Efforts to discriminate among analytes using arrays of chemical sensors, more often than not, rely on such non-linear signal processing techniques as multi-dimensional, non-linear clustering, artificial neural networks, fuzzy logic, genetic algorithms and others. In this work, we seek to establish linear separability of the problem using multiple (two) stages of signal processing performed in sequence rather than the single stage that characterizes most other efforts. Once linear separability is established, the problem of analyte discrimination, in field operation, is completed more efficiently using linear signal processing techniques. Ultimately, in power and space limited applications for portable chemical sensing, we wish to minimize the number of mathematical operations and stored parameters needed to solve the discrimination problem.

In this work, we demonstrate that an analyte discrimination problem that has been solved using a single-stage, non-linear signal processing technique can also be solved using two stages of linear

signal processing. The problem is that of a 7-element SAW device array used to discriminate 16 analytes that consist of aliphatic, aromatic, and chlorinated hydrocarbons, alcohols, ketones, and organophosphorous compounds across a range of concentrations from 0.40 - 48.0 P/P_{sat}, where P and P_{sat} are the partial pressure and saturated partial pressure of the analyte of interest respectively This wide range of analytes and concentrations requires a large computing space. Our first step is to reduce the dynamic range of the data by reducing the influence of concentration. After examination of the data, 59 of 112 possible response characteristics (isotherms) across all sensors and analytes in the array are linear. Because the array characteristics are dominated by linear response curves (53%), we logically choose a linear normalization method to reduce the influence of concentration. During normalization, every point on every response curve in the training set is normalized to the value of coating 1, the SAW device coated with bare quartz; this reference sensor demonstrates linear response characteristics more frequently (69%) than other coatings for all 16 analytes and is used to normalize all response points for the remaining six coatings as follows:

$$\Delta f_{normalized(a, b)} = (\Delta f_{a, b})/(\Delta f_{1, b}) \qquad (1)$$

where Δf refers to the frequency shift experienced by the SAW device in the presence of an analyte, a refers to the coating (numbered 1 through 7) and b refers to the analyte type (1 - 16).

After normalization, a standard linear signal processing technique, principal components analysis, is used to transform the seven dimensional data into two dimensions which contain most of the variance in the original data. The data are then analyzed using standard linear regression techniques (fitting the points to a line) and ambiguous classes are determined (analytes that are not distinguishable in the first stage of processing). A second stage of linear signal processing (principal components analysis) uses the most distinguishable features of the ambiguous classes of the first stage. Linear regression of the first two principal components in this second stage then discriminates or separates the remaining analytes in the test set. Once the principal components are determined using the test set, they can be implemented in field discrimination tasks by storing small, two-dimensional matrices containing the principal components. PCA, in contrast to non-linear signal processing techniques, also involves a significantly reduced number of mathematical operations required to multiple each seven dimensional data point by its 7X2 principal component matrix and then, to evaluate its distance to the 16 lines that characterize the analytes in the training set.

Any point, once transformed, that lies further than two standard deviations from any of the 16 lines in the training set, is labelled an outlier. In this way, unusual ambient conditions, new analytes, or broken sensors can be flagged because each data point is not necessarily forced into a particular classification group. The conversion of this problem to linear signal processing guarantees a significant reduction in the computational demands of the system which can provide lower power, more compact operation or more computational capability to address erroneous or unexpected conditions in the environment or sensor array.

In summary, the 7-element, SAW device array is processed to discriminate 16 analytes using:

- Multi-stage principal components analysis: Stage 1 is pre-processing (normalization of data); Stage 2 is principal components analysis on all 16 analytes simultaneously; Stage 3 is principal components analysis only on members of classes

that could not be separated in Stage 2. Stage 3 is preceded by feature extraction which extracts the most important features of each group to discriminating members of that group.

- VERI: creates a non-linear, multi-dimensional cluster that encompass all training points for a particular analyte. Distance from each testing point to all 16 clusters is then evaluated to determine which (if any) cluster the training point belongs.

- Artificial Neural Network: uses a fully interconnected, feed forward, back propagation trained perceptron-style architecture (7 input nodes, 14 hidden nodes, 16 output nodes) to train on the data set.

In the next section, we discuss the results of linear analysis of a test set for all 16 analytes and compare computational overhead (clock cycles and memory) for this technique and the two other common, non-linear techniques (VERI, Artificial neural networks) used to process the outputs of the SAW array.

EXPERIMENTAL DETAILS

The array of sensors used in this demonstration is a 7-element array of SAW sensors fabricated and tested at Sandia National Laboratories, each coated with a different material and tested across a wide range of concentrations (0.40 - 48.0 P/P_{sat}) for 16 analytes of interest (see Figure 1). The 7-element array is a reduced form of a 30-element array optimized using a nonlinear signal processing program called VERI (visually empirical region of influence) which relies on human visual perceptions of clusters to separate analytes from one another; although VERI works as well with linear clusters as non-linear, most of the clusters in this analysis are determined to be non-linear. The seven coatings used in this array are a bare quartz coating, a dendrimer coating, two polymer coatings (SE-30, PECH), and three metal coatings (CuO, Au, Pt). Details of the SAW device fabrication and the array optimization and non-linear clustering routine (VERI) can be found in 12 and 8 respectively.

The first stage of the linear signal processing process is preprocessing, which is designed to reduce the influence of parameters that are not relevant to analyte discrimination. In this case, the largest variance in the raw data. is concentration, which is reflected in the principal component analysis of the raw data in Figure 1a. Normalization according to the most linear coating (the bare quartz reference) results in an 88% reduction in data variance (corresponding to variations in the first principal component) as shown in Figure 1b.

The second stage of processing is a straightforward principal components analysis of the data set after normalization (Figure 1b, 1c, 1d). After PCA, all but 5 of the 16 analytes separate into clearly defined clusters, where members of each cluster can be identified within 100% accuracy using a simple linear regression analysis. The remaining 5 analytes separate into two classes, one containing three analytes, cyclohexane(a), hexane(b) and isoctane(c) and the other containing two analytes, propanol(k) and DMMP(p), which are differentiated in the third and last stage of signal processing (Figure 1e, 1f) by first preprocessing the normalized signals used as the input to stage 2. Preprocessing involves highlighting the differences between members of a group. It is the differences between coatings 6 and 3, 7 and 6, and 4 and 1 that dominate the distinguishing features of the response curves in each of the two

classes. After preprocessing to extract only these features, another principal components analysis yields the results in Figures 1e and 1f where the five remaining analytes are clearly distinguishable.

Comparisons in computational overhead for these three types of processing these 16 analytes in this SAW array are made in Table 1. The Motorola MC68HC12 is used to translate the number of add, multiply, and compare operations in each technique to clock cycles which translate directly to processing time and power consumption. The linear processing techniques consume 82.5% and 77% less processing time than the VERI and ANN methods respectively, with no loss in accuracy. The linear techniques also consume 91% and 71% less memory than the VERI and ANN methods respectively. All three pattern recognition techniques achieve 100% discrimination accuracy in this experiment.

Table 1: Computational Overhead Comparison

| Method	Clock Cycles	Memory
PCA (linear)	12,142	344 bytes
VERI	69,390	4 kbytes
ANN	51,740	1.2 kbytes

CONCLUSIONS

We have demonstrated the use of linear signal processing techniques to perform chemical discrimination at the same accuracy as comparable non-linear techniques using a multi-stage hierarchy of processing. Multi-stage processing facilitates the removal of superfluous information to the chemical discrimination problem (e.g. concentration) and also allows the computational effort to be focused on areas of high activity in the chemical signature space. Speed improvements up to 82.5% and memory reductions up to 91% have been computed for the principal components analysis-based techniques presented on SAW-device based arrays in this paper. Comparable improvements can be achieved using other arrays for other analytes. The focus of increased speed, unlike many other sensing applications, is not in improving system response time, but rather in reducing power consumed by the overall chemical sensing system. These algorithms, as presented, can be implemented in a simple microcontroller, such as the Motorola MC68HC12, evaluated here to generate a self-contained system that is truly portable for field applications.

ACKNOWLEDGEMENTS

References

1 C. Di Natale, A. Macagnano, F. Davide, A.D'Amico, A. Legin, Y. Vlasov, A. Rudnitskaya, and B. Selezenev, "Multicomponent analysis on polluted waters by means of an electronic tongue," *Sensors and Actuators B*, vol. 44 (1997) pp. 423-428.

2 J.W. Gardner, "Detection of Vapours and Odours from a Multisensor Array Using Pattern Recognition: Part 1: Principal Component and Cluster Analysis," *Sensors and Actuators B*, vol. 4 (1991) pp. 109-115.

3 J.W. Gardner, T.C. Pearce, and S. Friel, "A Multisensor System for Beer Flavour Monitoring using an Array of Conducting Polymers and Predictive Classifiers," *Sensors and Actuators B*, vol. 18-19, (1994) pp. 240-243.

4 M. Hoummady, D. Hauden, B. Hivert, J. Henrious, P. Mielle, and P. Etievant, "SnO_2 Sensor Array for Flavour Recognition with Neural Networks," *Eurosensors VIII:* Toulouse, France, 1994.

5 P.S. Barker, J.R. Chen, N.E. Agbor, A.P. Monkman, P. Mars, and M.C. Petty, "Vapour Recognition using Organic films and Artificial Neural Networks," *Sensors and Actuators B*, vol. 17 (1994) pp. 143-147.

6 G. Huyberechts, P. Szecowka, J. Roggen, and B.W. Licznerski, "Simultaneous quantification of carbon monoxide and methane in humid air using a sensor array and an artificial neural network," *Sensors and Actuators B*, vol. 45 (1997) pp. 123-130.

7 Eduard Llobet, Jesus Brezmes, Xavier Vilanova, Jesus E. Sueiras, and Xavier Correig, "Qualitative and quantitative analysis of volatile organic compounds using transient and steady-state responses of a thick-film tin oxide gas sensor array," *Sensors and Actuators B*, vol. 41 (1997) pp. 13-21.

8 G.C. Osbourn and R.F. Martinez, "Empirically Defined Regions of Influence for Clustering Analyses," *Pattern Recognition* vol. 28 no. 11, (Nov. 1995) pp. 1793-1806.

9 John S. Wagner, Michael W. Trahan, Willie E. Nelson, Philip J. Hargis, Jr., and Gary Tisone, "Chemical Recognition Software," *SPIE Conf Optical Sensors for Environmental and Chemical Process Monitoring*: McLean, Virginia (1995) pp. 228-238.

10 Wang Ping and Xie Jun, "Novel Recognition Method for Electronic Nose using Artificial Neural Network and Fuzzy Recognition," *Sensors and Actuators B*, vol. B37, no. 3 (Dec. 1996) pp. 169-174.

11 Dimitrios Vlachos and John Avaritsiotis, "Fuzzy Neural Networks for Gas Sensing," *Eurosensors IX*: Stockholm, Sweden (June 1995) pp. 703-706.

12 Antonio J. Ricco, Richard M. Crooks, and Gordon C. Osbourn, "Surface Acoustic Wave Chemical Sensor Arrays: New Chemically Sensitive Interfaces Combined with Novel Cluster Analysis to Detect Volatile Organic Compounds and Mixtures," *Acc. Chem. Res.*, vol. 31, (1998) pp. 289-296.
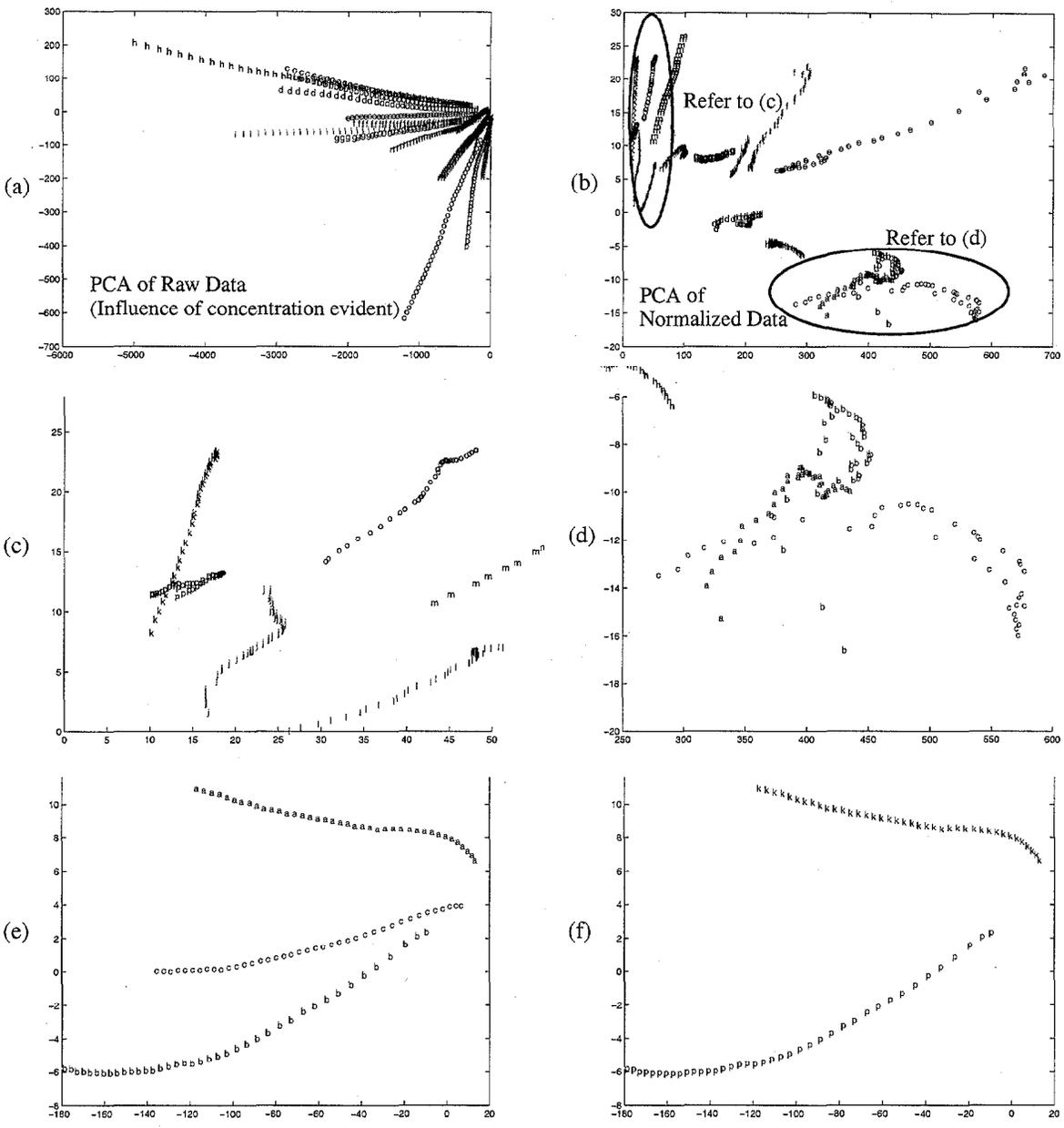
Code to Chemicals:

a: cyclo-hexane   e: benzene   i: trichloroethylene (TCE) m: acetone

b: hexane   f: toluene   j: methanol   n: methyl isobutyl ketone (MIBK)

c: isooctane   g: chlorobenzene   k: propanol   o: diisopropylmethylphosphonate (DIMP)

d: kerosene   h: carbon-tetrachloride l: pinacolyl alcohol   p: dimethylmethylphosphonate (DMMP)

Figure 1. Single and Multiple-stage Principal Component Analysis of 16 Chemicals in 7-element SAW Array
In (a) single stage principal components analysis, the analytes tend to group according to the magnitudes of their response curves, directly relating to concentration as well as discrimination. After (b), (c), (d) normalization, data clusters primarily according to fundamental differences in the response characteristics associated with the last six of the seven coatings. During the third stage of processing (e) and (f), members of closely related classes are separated and recognized. After clustering based on distance to the nearest line, all points cluster to their correctly identified analytes.

AN INTEGRATED CHEMICAL SENSOR ARRAY USING CARBON BLACK POLYMERS AND A STANDARD CMOS PROCESS

Jeffrey A. Dickson[1], Michael S. Freund[2], Nathan S. Lewis[2], Rodney M. Goodman[1]

[1] Department of Electrical Engineering [2] Department of Chemistry
California Institute of Technology
Pasadena, California 91125

ABSTRACT

We have developed a new chemical sensor array by combining polymer-based chemiresistors with a standard integrated circuit technology. We fabricated an array of addressable chemical sensor sites in a CMOS process, and then performed a post-processing step of electroless gold to create sensor contacts. We create sensors by spraying a mixture of nonconductive polymers and carbon black particles onto the sensor sites. We demonstrate that an array of diverse chemical sensors can perform discrimination of odors.

INTRODUCTION

This paper describes the development of an array of chemical sensors. The sensors are based on the polymer approach of Lewis et al.[1, 2] employing carbon black and non-conducting polymers. Exposure to particular analytes causes the sensors to swell, which increases the electrical resistance. By employing different polymers we can create a large number of different broadly tuned sensors. Since they can be fabricated at room temperature and a change in resistance is easily measured, this technology is attractive for integration with active circuitry. This array is capable of chemical discrimination that does not require external excitation or complicated signal processing like optical sensing[3]. Unlike SAW devices, we can integrate large arrays on the same chip[4]. Applications including environmental monitoring, narcotic and explosives detection have demanding chemical sensing requirements. Our goal is to create small, inexpensive, low power and even wearable chemical sensor arrays that rival the detection and discrimination capabilities of mammalian olfaction.

DESIGN AND FABRICATION

The sensor consists of an array of individually addressable electrical contacts, on which a polymer/carbon black mixture is deposited. The sensor technology is well suited to integration with on-chip circuitry. The array allows each sensor to be individually addressed.

Figure 1 shows the schematic of the unit sensor cell. The cell consists of a switch transistor and decoding logic. The availability of only two metals layers in the IC process required transistors at each sensor cell to perform decoding. This circuitry (M1-M4) decodes X and Y selection signals generated by shift registers on the periphery of the array. This selection signal controls a switch (M7) that toggles a current (I_{in}) through the resistive sensor. In this design only one sensor is energized at a time to reduce power consumption. To reduce noise and the switch resistance, transistor M7 occupies most of the sensor area. The decoding circuitry also selects a transmission gate (M5,M6,M8,M9) which passes the sensor voltage to a column output bus. This signal is amplified and transmitted off-chip for processing. The decoding circuitry is complicated because the sensor occupies one of two available metal layers in the fabrication process we used for this chip, precluding the use of a simplified bus scheme.
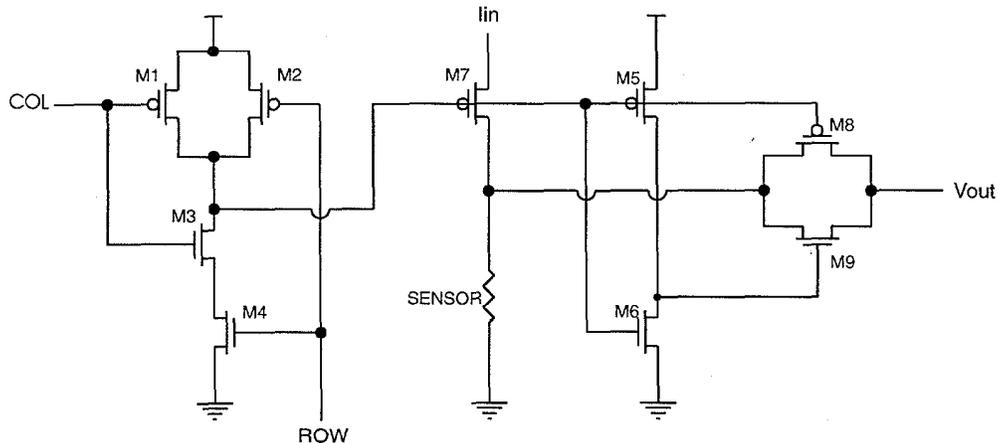

Figure 1 Schematic of three wire sensing cell. Transistors M1-M4 form a NAND gate to select the cell, M7 switches the current source on sensor resistor, and M5,M6,M8,M9 form a transmission gate to select the output on a column output line. The column output is buffered and passed off-chip.
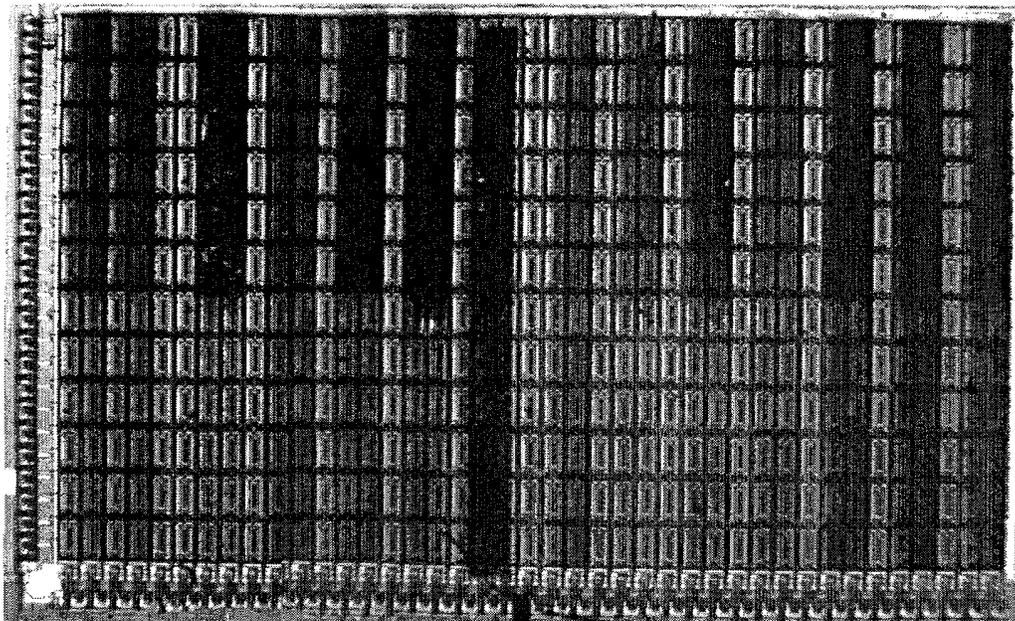
Figure 2 Photograph of the integrated sensor chip, after post fabrication electroless gold plating and polymer deposition by airbrush. The chip contains 492 sensor sites arranged in a 41X12 array. On this test chip 209 sensor sites have been covered with one of eight different polymer based chemiresistors. Shift registers located along the left and bottom of the array select an individual sensor site, whose output is amplified and passed off chip. The chip is 0.5 cm by 0.25 cm.

Figure 2 shows the 0.5 cm X 0.25 cm chip, with 492 sensors arranged as an array of 41 X 12 sites, fabricated in a 2.0 micron process through the MOSIS design service. The dark vertical bars on the chip are deposited chemical sensors, discussed later in the text.

A close-up of the individual sensor sites its shown in Figure 3. There are two contacts to the sensor: one with the drive switch M7, and another connecting the sensor to signal ground. The ground terminal is laid out as a ring around the perimeter of the cell and is common to all sensors. The interior of each cell contains the drive contact for the sensor. The ring structure was motivated by difficulties encountered in early sensor deposition trials – the carbon black particles would aggregate along the perimeter of a deposited sensor, creating a low resistance path. Moreover, the ring structure allows us to experiment with depositing mixtures of sensor materials across the chip[5]. We deposit the sensor polymer between these two contacts, directly on top of the active circuitry. We make the sensors rectangular to increase the contact area for the interior contact, as well as to reduce contact noise and 1/f noise due to the non-uniform electrical field[6].

We use a standard commercial foundry for the fabrication of the integrated circuits. The top layer of Aluminum is used for the sensor contact. Unfortunately, the native aluminum oxide that forms on the contacts prevents depositing the sensors without a post-processing step. Since dedicated wafer runs are cost prohibitive for small prototyping runs, this step must be performed on the individual chip die returned from the foundry. This precludes the use of a conventional mask based approach, since it is difficult to use a resist mask on an individual die.

To create suitable contacts we use an electroless Ni/Au process that requires no masking from Stapleton Technologies (Long Beach, California). This process can be performed easily on individual die with simple equipment and requires only seven procedures: four involving cleaning and surface preparation and

three plating steps. The surface preparation involves an acid zincate process to remove the native oxide and activate the aluminum surface. This is followed by the three plating steps. Nickel is plated first, followed by a two stages of gold plating: a monolayer process that plates the Nickel and then a build up stage that finishes the plating. Figure 3 shows three sensors, plated with 9 microns of Ni and 1 micron of Au. In addition to creating a non-reactive surface for the sensor contacts, the plating also creates wells that help constrain the sensor material during deposition.
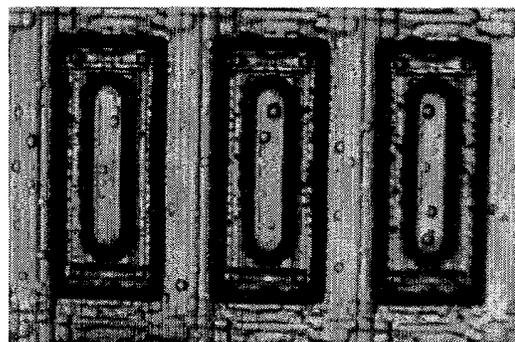

Figure 3 Picture of three sensor sites after the electroless gold plating. The central bars are the switched output node of each individual sensor. The surrounding conductor is a common ground. The sensor material is deposited on top of the chip, forming the sensor between the central contact and the surrounding ground

The sensors are a combination of a particular polymer and carbon black particles. To prepare the sensor material we combine 20 mg of Carbon Black and 80 mg of the polymer in powder form. The polymers and solvents are shown in Table 1. The Carbon Black we used is a furnace black from Cabot Co. (Billerica, MA). We place the mixture in an ultrasonic bath for a minimum of five min-

163

utes to suspend the carbon black particles before depositing the sensors.

Polymer	Solvent
PEO	Toluene
PEVA	Toluene
p-Butadiene	Toluene
p-vinyl carbazole	THF
p-vinyl acetate	Acetone
p-capraclactone	THF
p-vinyl pyrrolidone	Ethanol
p-4-vinyl phenol	THF

Table 1 Listing of the eight polymers and the corresponding solvents used in the fabrication of the chemiresistors.

To deposit the sensor material on the surface of the integrated circuit we employed an airbrush. A sheet of polyamide 50 microns thick is used as a physical mask to define the sensors. Apertures are cut in the polyamide using a computer-controlled laser. While other materials and processes are available to make this mask, the polyamide gaskets well to the surface of the chip. In addition, the ability to see through the polyamide allows us to position the mask accurately. We are able to create apertures as small as 50 microns using this technique, enabling us to spray individual sensor sites. We sprayed eight different polymers (Table 1) in columns two sensors (270 µm) wide. Figure 4 shows a close-up of the sprayed chip, demonstrating the ability to fabricate small sensors on the chip. The spraying of the polymer allows us to create thin, uniform films of sensor material. Previously we used a direct deposition technique using a fine tip that resulted in uneven deposition and thick films.
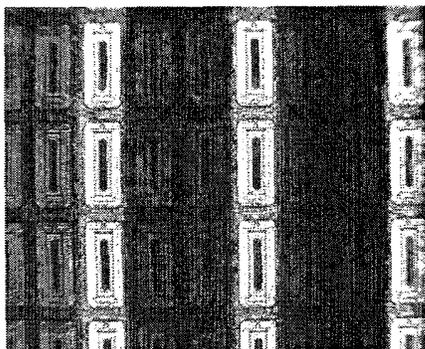

Figure 4 Deposition of sensor materials. The sensors are sprayed with an airbrush, using a laser cut polyamide mask. We sprayed the chip in columns of two sensors wide (270 µm), with a row left blank to demonstrate the ability to selectively spray the sensors at this resolution

TESTING

The current vs. voltage characteristic for an individual sensor node is shown in Figure 5. This demonstrates that we are able to successfully fabricate an individually addressable sensor pixel. The nonlinearity of the response is due to the on-chip amplifier, which is not optimized to be linear over the entire voltage range. In practice, the sensors are biased to operate at a single operating point to minimize the error due to the nonlinearity of the amplifier.
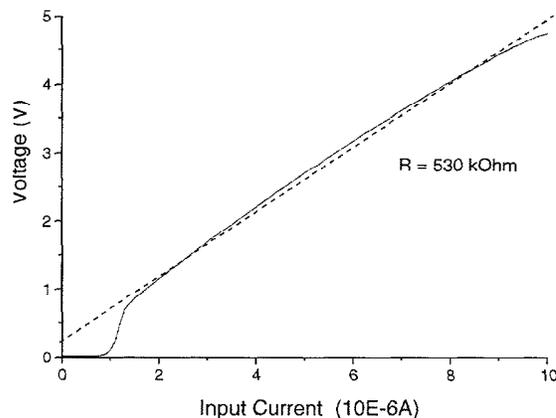

Figure 5 Voltage vs. Current sweep of an individual sensor node, demonstrating its linear resistive nature. At <1V output the column amplifier does not operate, and the deviation from linear is due to errors in the amplifier.

To test the sensors we use an automated flow system to generate solvent vapors at a specific vapor pressure. Mass flow controllers regulate a laboratory air supply through ceramic frits in glass bubblers filled with the desired solvent.

Figure 6 shows the temporal response of a p-vinyl acetate sensor to a series of random analyte exposures at 5% vapor pressure. After the analyte is removed the sensor returns to its nominal value. We use the maximum percentage change in resistance during an exposure as the output response.
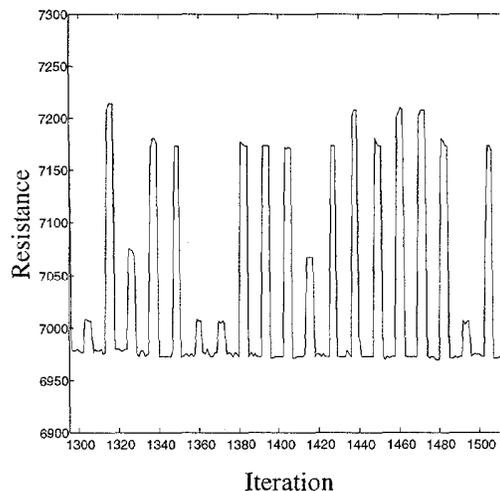

Figure 6 Temporal response of a typical polymer carbon black chemiresistor to a series of random analyte exposures. This detector was composed of p-vinyl acetate, and the response shown is the change in resistance to a random solvent exposure at 5% vapor pressure.

While we have only used eight different polymers in this paper, we can still perform classification. The array of sensors produces a characteristic fingerprint for a particular analyte, shown in Figure 7 for eight unique sensors exposed to eight analytes at 5% vapor pressure. One method of performing classification is using
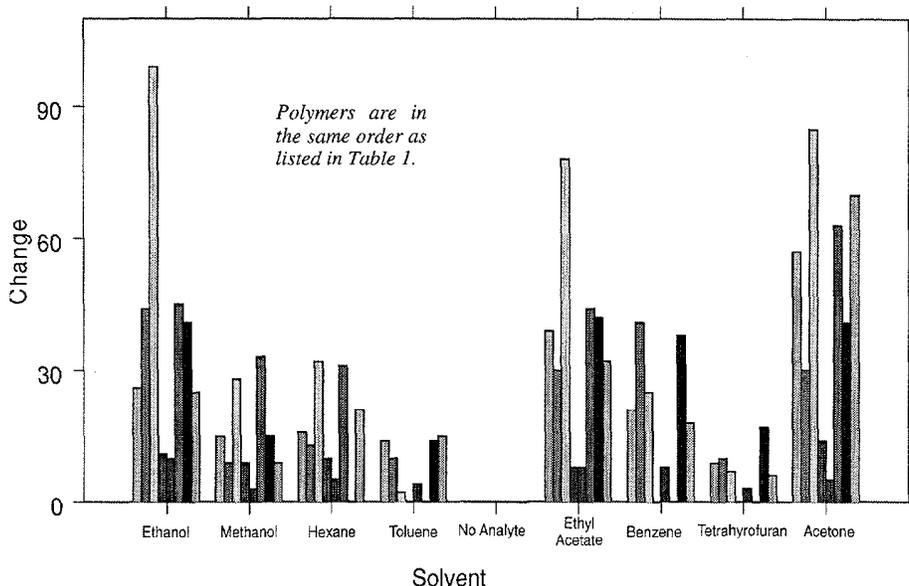
Figure 7 *Sensor responses for a variety of analyte exposures. For each exposure the maximum resistance change for one sensor of each variety (see Table 1) is shown. The data are scaled to the maximum response for all sensors and exposures. The response provides a fingerprint for a particular analyte. Statistical methods can be used to provide a classification of an unknown exposure*

principal component analysis. Figure 8 shows the second and third principal components applied to the output responses of the 209 sensors on the chip. We are able to cluster the response of the sensors to the eight different analytes, permitting classification and discrimination.
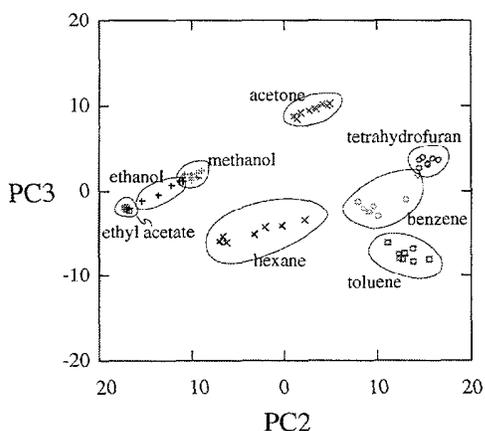

Figure 8 *Principal component analysis of the chip response. The second (PC2) and third (PC3) principal components of the maximum resistance change per exposure are shown in two dimensions. Just these two components allow discrimination and classification of the analytes.*

CONCLUSION

We have demonstrated the successful integration of a chemical sensor array with a standard CMOS process. Following a simple post processing operation we are able to deposit sensor material on the surface on an integrated circuit. By depositing different sensor materials we are able to create an array capable of discriminating analytes. We plan to create larger arrays with a large number of different polymers, and arrays with additional active circuitry such as amplification and adaptation.

REFERENCES

[1]   M. S. Freund and N. S. Lewis, "A Chemically Diverse Conducting Polymer-Based Electronic Nose," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 92, pp. 2652-2656, 1995.

[2]   E. J. Severin, R. D. Sanner, B. J. Doleman, and N. S. Lewis, "Differential detection of enantiomeric gaseous analytes using carbon black-chiral polymer composite, chemically sensitive resistors," *Analytical Chemistry*, vol. 70, pp. 1440-1443, 1998.

[3]   T. A. Dickinson, J. White, J. S. Kauer, and D. R. Walt, "A chemical-detecting system based on a cross-reactive optical sensor array," *Nature*, vol. 382, pp. 697-700, 1996.

[4]   J. W. Grate, S. J. Martin, and R. M. White, "Acoustic-Wave Microsensors .1," *Analytical Chemistry*, vol. 65, pp. A940-A948, 1993.

[5]   B. J. Doleman, R. D. Sanner, E. J. Severin, R. H. Grubbs, and N. S. Lewis, "Use of compatible polymer blends to fabricate arrays of carbon black-polymer composite vapor detectors," *Analytical Chemistry*, vol. 70, pp. 2560-2564, 1998.

[6]   H. Yoshida, "The Effect of Resistor Geometry On Current Noise," *IEEE Transactions On Components Hybrids and Manufacturing Technology*, vol. 16, pp. 344-349, 1993.

PACKAGE LEVEL INTEGRATION OF SILICON MICROFABRICATED REACTORS TO FORM A MINIATURE REACTOR TEST SYSTEM

David J. Quiram[*], Klavs F. Jensen[*], and Martin A. Schmidt[†]
[*]Dept. of Chemical Engineering, [†]Dept. of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, MA 02139

James F. Ryley, Patrick L. Mills, Mark D. Wetzel,
James W. Ashmead, Richard D. Bryson,
Daniel J. Kraus, and Alan P. Stamford
DuPont Central Research and Development
Wilmington, DE 19880-0357

Steven J. Medwin and Russell E. Mitchell
ENSER Corporation
Cinnaminson, NJ 08077

ABSTRACT

A method for packaging microfabricated reactors is described that allows for the creation of parallel arrays of operating devices for scale-up purposes. As a packaging demonstration, these microreactors are being coupled with microvalves and MEMS mass flow controllers to form a miniature microreactor test station. A novel method based on Known Good Die technology is used for the electrical and fluidic interfaces to the microreactors. The microfluidic components are mounted on standard circuit boards that have the monitoring and control circuitry along with fluidic connections. The boards are housed in a standard 6U CompactPCI chassis utilizing its backplane for the electrical connections between boards. This system will be the first demonstration of a scalable, parallel microreactor array that integrates reactors with other microfluidic devices and the electronic circuitry needed for monitoring and control.

INTRODUCTION

The use of microfabrication techniques to produce microreactors and other miniaturized devices for chemical processing is completing its first decade of active research[1, 2]. Since the establishment of an annual international conference in the field in 1997, there has been an explosive growth in progress[3-6]. As in any MEMS device, the reasons for miniaturization have to be strong to compete with conventional technologies. There is now overwhelming evidence that miniaturization offers advantages ranging from greater selectivity and conversion to the processing of new materials that could not otherwise be made[6]. A new challenge is to utilize the benefits of microreactor technology for commercial processes. In some cases, a single high throughput device may be adequate, but scale-up is generally needed to increase production rates.

The scale-up of microreactors differs from the traditional scale-up paradigm in that production capacity is increased by creating an array of microreactors running in parallel versus building a larger scale unit[1, 7]. Although obtaining greater production rates appears to be simpler, this parallelization presents new challenges that have not been addressed in previous efforts[8]. Particularly, the areas of reactor monitoring and control quickly become complex as the system size is increased to massively parallel arrays. More attention has to be placed on the microreactor packaging since both fluidic and electrical interconnections must be designed to allow for parallelization. Questions also arise as to how such systems will perform as a whole, and whether they can be designed in such a way to be financially competitive with conventional technology. Efforts in this joint MIT-DuPont program concentrate on developing a technological base for integrating microfabricated chemical reactors into large parallel arrays suitable for on-site production of small quantities of hazardous or thermally unstable chemicals, such as organic peroxides, hydrogen cyanide, organic acids, and related species. With the design and construction of prototype multiple reactor systems, a knowledge base will be created that will eventually lead to commercially viable systems.

BACKGROUND

The goal of the MIT-DuPont program is to build a prototype microreactor system that incorporates MEMS reactors along with other microfluidic devices. This system will be functionally equivalent to an industrial catalyst test station constructed using conventional technology. Figure 1 shows the layout of a typical system, which is comprised of fluidic components to form a feed gas manifold, a reactor manifold, and a product analysis system. These sections are constructed using valves, pressure regulators, reactors, and mass flow controllers. The microreactor system will be constructed entirely from MEMS devices and will include the feed gas manifold and the reactor manifold. Product analysis will be done using standard technology so the performance of this system can be directly compared to its larger counterpart.

The system being developed utilizes the thin-film, gas-phase microreactor design that originated at MIT at the beginning of this research program[9, 10]. This device has a single channel with
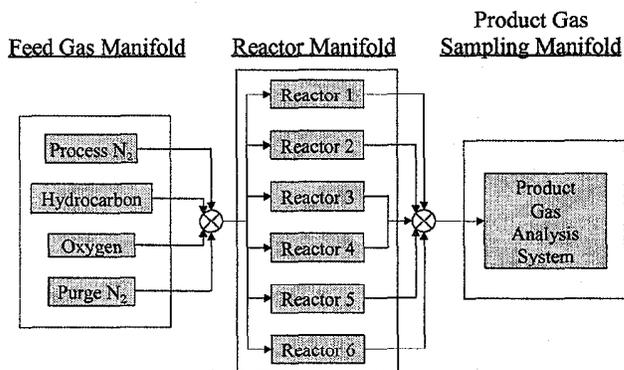

Figure 1. Catalyst test station layout.

integrated heating and temperature sensing. Figure 2 shows the top and cross-sectional views of the reactor, which has a reactant inlet and product outlet on opposite ends of the channel. The bottom of the channel is sealed by a Pyrex layer that is anodically bonded to the silicon die. The Pyrex has drilled holes for the gas inlet and outlets. This reactor differs from the original "T" microreactor in that reactant mixing takes place off-chip in the gas feed manifold. This was done to 1) improve membrane stability; 2) allow direct sampling of the reactant mixture into the product analysis system; and 3) to reduce the overall reactor size. The reaction channel is 0.5 mm wide, 0.5 mm deep, and 11 mm long. It has seven distinct zones for heating and temperature sensing. Each heater is 0.94 mm in length and the temperature sensors are 0.77 mm long. The size and number of heater and temperature sensor elements was chosen to give more resolution in the reactor temperature profile and allow greater control over this profile. An enlargement of the heater and temperature sensor configuration is shown in Fig. 2d.
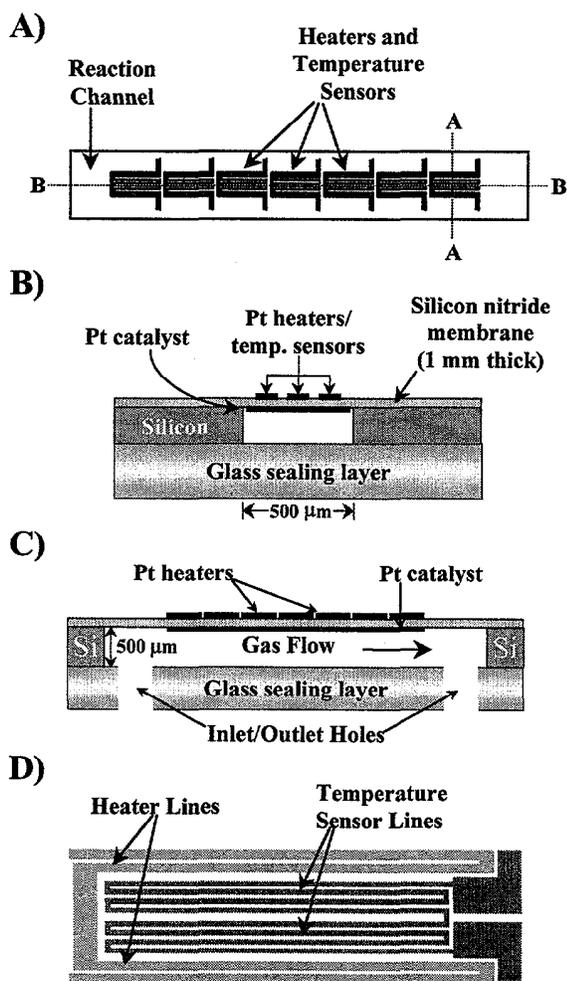
Figure 2. *MIT-DuPont scale-up slot microreactor. A) Top-view B) Cross-sectional view AA C) Cross-sectional view BB D) Enlargement of heater and temperature sensor configuration*

METHODOLOGY

The characterization of this type of microreactor has been reported in previous publications[9-12]. The emphasis here is integrating the thin-film reactor design into a hybrid MEMS for chemical production or catalyst testing. The system design can be divided into three main categories: 1) device packaging, 2) reactor design, and 3) system level integration of microfluidic components and control electronics. The approach taken by the MIT-DuPont program was to first identify methods for packaging and evaluate them. The most suitable method was selected by considering the constraints placed by the reactor design, the chemistries of interest, prototype cost, and overall size. The packaging method, reactor design, and hybridization concept were then based on these initial decisions. The following sections describe each of these components in greater detail.

Device Packaging

The critical factor in selecting the microreactor packaging was the requirement for fast reactor replacement with the constraint of consistent, robust interconnects. The thin-film reactor design has advantages due to its high level of integration, but its main disadvantage is its fragility. The primary failure mechanism of this microreactor design is membrane rupture, which is generally caused by high temperatures in the reaction zone. However, it can also fail as a result of physical shocks. The MIT-DuPont microreactor also uses a large number of electrical connections. Thus, the first level package screening was to identify a fast, easy method to make electrical contact with the die that was also compact in size. Initial testing of the "T" thin-film microreactor was done using a probe card, but this required precision aligning equipment. This resulted in a test setup with a large footprint due to the probe card size and the micropositioners used. Wire bonding was a sub-optimal choice because of the time required to replace chips (automated wire-bonding was not an option at this stage of development).
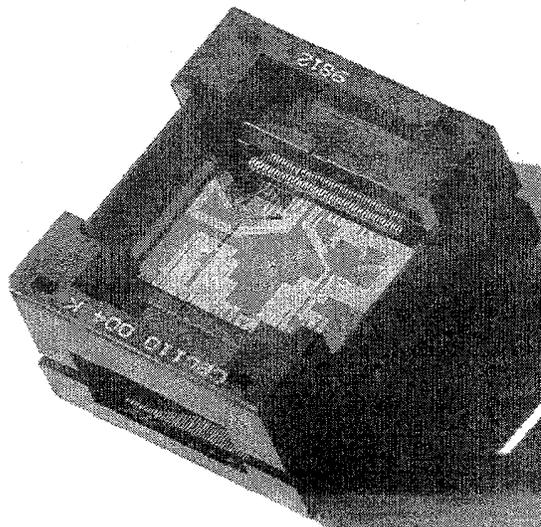

Figure 3. *TI DieMate® socket with a mounted 'Y' scale-up microreactor die.*

167
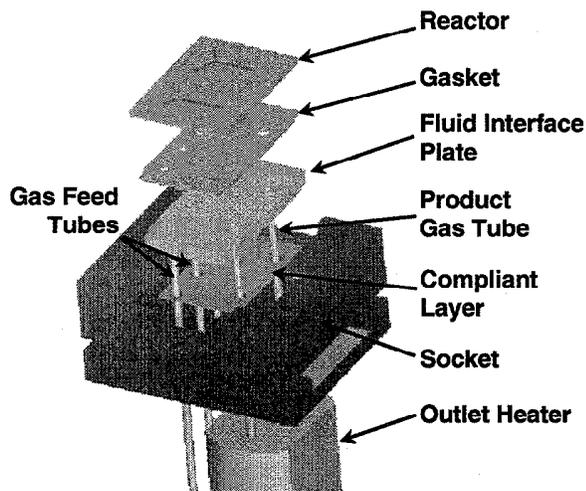
Figure 4. Microreactor fluidic connection assembly.

To solve the microreactor packaging problem, a standard Known Good Die (KGD) socket, manufactured by Texas Instruments, was chosen. This socket, which was originally intended for use in die testing on production lines, was available in a variety of sizes and pin counts. The DieMate® socket was designed to be used with a chip mounting assembly for die testing, but this assembly was not needed for our purposes. Instead, the microreactor was designed to fit directly in the socket as shown in Fig. 3. The socket functions with a series of springs that form a pressure contact between the pins and the reactor die. The springs hold the die in place, and by designing the microreactor to fit exactly inside the socket, the system provides a fast, self-aligned method for forming electrical interconnections. The socket also lends itself to easy Printed Circuit (PC) board mounting.

The entire DieMate® microreactor assembly is illustrated in Fig. 4. Since the DieMate® socket was not intended for use with fluidic systems, it was modified to provide a method for interconnecting the gas transfer lines to the microreactor. Referring to Fig. 4, a fluidic connection plate was inserted underneath the reactor die and a compliant layer was placed between these two pieces to form a gas-tight seal using the DieMate® springs to provide the sealing force. This connection plate is a piece of flat stainless steel with 1/16 in. o.d. stainless steel tubing welded into it for the reactant and product gas streams.

Reactor Design

With the DieMate® socket chosen as the packaging piece it was necessary to design a reactor to fit into the socket. Since the smallest DieMate® socket available required a die size of 20 mm by 28 mm, two single channel reactors were placed on a single die. With 110 pins, the socket had more than enough electrical interconnections for the temperature sensors and heaters. Two microreactor designs were developed for this socket. The prototype design was a 'Y' shaped reactor that had separate feed inlets for the two reactants. This microreactor can be seen in Fig. 3 mounted in a DieMate® socket. This design was later replaced with a straight channel device for the reasons mentioned previously. Figure 5 shows a picture of the straight channel reactor. Both designs were successful in that all the bond pads aligned as designed with the DieMate® pins when inserted into the socket.
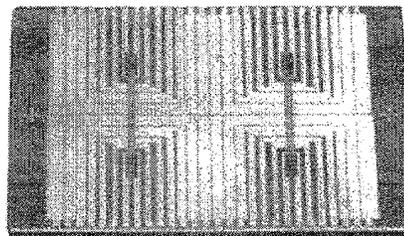

Figure 5. Scale-up slot microreactor.

Hybrid System Design

The compact reactor packaging allows the integration of the microreactors on PC boards with other MEMS devices to form catalyst test stations or small-scale production units. The hybridization concept is based on mounting devices on standard 6U CompactPCI boards that will then be inserted into a Kaparel PS6090 chassis to form the final unit as shown in Fig. 6. The individual boards will provide specific functions, such as a reactor card with a reactor and its Mass Flow Controller (MFC). The card will also have some of the electronic circuits needed to run the reactors. Additional circuits will be located on other CompactPCI boards inside the chassis. A CAD drawing of such a card containing a Redwood Microsystems Flow-ister®, a MEMS-based MFC, and a microreactor is shown in Fig. 7. Similar cards will be made for the feed gas lines. The 6U CompactPCI chassis standard was chosen because of the wide range of backplanes that are available, the larger board sizes possible, and the high number of electrical interconnections available for use. The microreactor test station will use a standard telephony backplane with minor modifications, which can be used directly with the analog signals for device monitoring and control. This CompactPCI chassis will be connected to another CompactPCI chassis containing the control and data acquisition hardware based on National Instruments' PXI/CompactPCI embedded controllers.
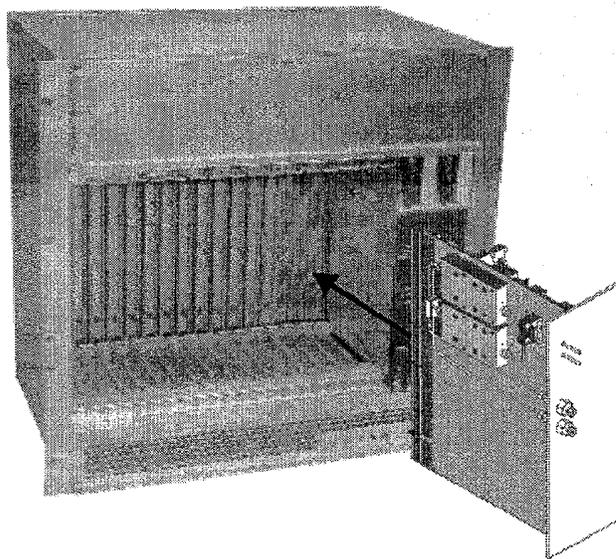

*Figure 6. Kaparel PS6090 CompactPCI chassis with a CAD drawing of a fluidic circuit board.*
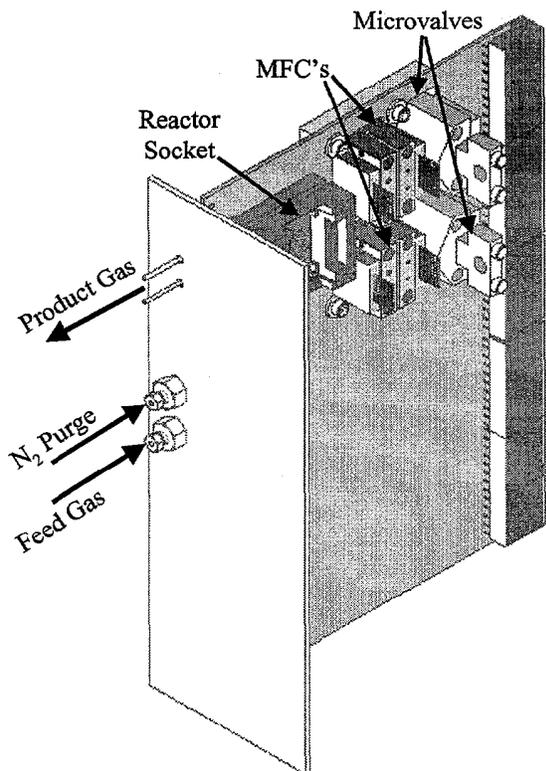
Figure 7. *6U Compact/PCI reactor board with mounted DieMate® socket, Flow-ister® MFC's, and Redwood Microsystems microvalves.*

FUTURE WORK

With the design of the system components completed, current work is focusing on assembling and testing the hybrid microreactor test station. Additional work will focus on the development of a human-machine interface to allow for turnkey style of operation. This will involve building GUI's for operation and developing robust system control algorithms. This is essential not only for ease-of-use, but also for process safety. After this system has been built and tested, a performance comparison between the miniature and conventional systems will be made.

ACKNOWLEDGEMENTS

REFERENCES

1. J. J. Lerou, M. P. Harold, J. Ryley, J. Ashmead, T. C. O'Brien, M. Johnson, J. Perrotto, C. T. Blaisdell, T. A. Rensi and J. Nyquist, "Microfabricated minichemical systems: Technical feasibility", *Microsystem Technology for Chemical and Biological Microreactors: Papers of the Workshop on Microsystem Technology, Mainz, 20-21 February, 1995*, DECHEMA, Frankfurt (1996), pp. 51-69.

2. K. P. Jäckel, "Microtechnology: Application opportunities in the chemical industry", *Microsystem Technology for Chemical and Biological Microreactors: Papers of the Workshop on Microsystem Technology, 20-21 February, 1995*, DECHEMA, Frankfurt (1996), pp. 29-50.

3. W. Ehrfeld, I. H. Rinard and R. S. Wegeng, *Process Miniaturization: 2nd International Conference on Microreaction Technology*, AIChE, New Orleans, LA (1998).

4. W. Ehrfeld, *Microreaction Technology: Proceedings of the First International Conference on Microreaction Technology*, Springer, Berlin (1997).

5. W. Ehrfeld, *IMRET 3: Proceedings of the Third International Conference on Microreaction Technology*, Springer Verlag, New York (2000).

6. I. Rinard, *IMRET 4: 4th International Conference on Microreaction Technology*, AIChE, Atlanta (2000).

7. W. Ehrfeld, V. Hessel, H. Möbius, T. Richter and K. Russow, "Potentials and realization of microreactors", *Microsystem Technology for Chemical and Biological Microreactors: Papers of the Workshop on Microsystem Technology, Mainz, 20-21 February, 1995*, DECHEMA, Frankfurt (1996), pp. 1-28.

8. D. J. Quiram, K. F. Jensen, M. A. Schmidt, P. L. Mills, J. F. Ryley and M. D. Wetzel, "Integrated microchemical systems: opportunities for process design", *Foundations of Computer-Aided Process Operations*, AIChE, New York (2000),

9. R. Srinivasan, I.-M. Hsing, J. Ryley, M. P. Harold, K. F. Jensen and M. A. Schmidt, "Micromachined chemical reactors for surface catalyzed reactions", *Technical Digest of the 1996 Solid-State Sensor and Actuator Workshop*, Hilton Head Island, SC, June 3-6, Transducers Research Foundation (1996), pp. 15-18.

10. R. Srinivasan, I.-M. Hsing, P. E. Berger, K. F. Jensen, S. L. Firebaugh, M. A. Schmidt, M. P. Harold, J. J. Lerou and J. F. Ryley, "Micromachined reactors for catalytic partial oxidation reactions", *AIChE Journal, 43*, 11 (1997), pp. 3059-3069.

11. K. F. Jensen, I.-M. Hsing, R. Srinivasan, M. A. Schmidt, M. P. Harold, J. J. Lerou and J. F. Ryley, "Reaction Engineering for Microreactor Systems", *Microreaction Technology: Proceedings of the First International Conference on Microreaction Technology*, Springer, Berlin (1997), pp. 2-9.

12. D. J. Quiram, I.-M. Hsing, A. J. Franz, R. Srinivasan, K. F. Jensen and M. A. Schmidt, "Characterization of microchemical systems using simulations", *Process Miniaturization: 2nd International Conference on Microreaction Technology*, New Orleans, LA, March 9-12, AIChE (1998), pp. 205-210.

A VERTICAL MICROMACHINED RESISTIVE HEATER
FOR A MICRO GAS SEPARATION COLUMN

Janet K. Robertson
Standard MEMS Inc.
Burlington, MA 01803

ABSTRACT

This paper presents a micromachined vertical resistive heater fabricated as part of a micro gas separation column. The complete device is formed from two silicon die electrostatically bonded to a single glass die. The glass die is conventionally machined and contains four through holes. The holes are filled with a zeolite adsorbent and thus act as separation columns. Each silicon die serves as a screen which allows air to flow through the device while holding the zeolite in place. The 5000μm diameter screens are Deep Reactive Ion Etched (DRIE) and contain 1964 square 50μm x 50μm x 350μm channels. A vertical resistor is also integrated onto one of the silicon die. The junction isolated resistor is formed from a phosphorus diffusion which occurs after DRIE. Thus, the resistor is embedded into the walls of the micromachined silicon screen. The vertical resistor is used to heat a purge gas and regenerate the zeolite adsorbent. The measured vertical resistance is between 6Ω and 7Ω. The residence time of the gas in the heater is ~60msec assuming a 1SCCM N2 gas flow rate. The theoretical heat transfer coefficient is 3000W/m²/°C. The power consumed to heat both the screen and the gas at steady state is ~33mW/die.

INTRODUCTION

The development of a micro gas separation system is a key component for the continued development of micro chemical processes. Like micro valves [1], pumps [2] and heat exchangers [3], the development of a micro gas separation system would contribute greatly to the successful development of a microreactor process. For example, a separation system could remove contaminants from feed gas entering a micro reactor and/or remove unwanted by-products from the effluent.

A micro separation system could also be used to enhance the performance of gas sensors. Current gas sensors, while having good sensitivity, suffer from poor selectivity, responding equally well to a number of different gas species. By incorporating one or more gas separation columns in the incoming flow stream, species which are not of interest could be removed, reducing the possibility of ambiguous gas sensor responses. A gas separation/gas sensing system could then be used to control a micro chemical process by the real-time separation and detection of critical product gases.

Although many separations are possible, the prototype micro gas separation column was designed to separate water from air. The column uses 5A powdered zeolite to dry 1SCCM of air. The on-stream time (i.e. the time necessary to saturate the adsorbent with water) at 25°C and 100% humidity is approximately 1hour. The zeolite is regenerated by passing a heated purge gas through the separation column, heating the zeolite and desorbing the water. By using multiple columns, some adsorbing while others are regenerating, the system can

supply a continuous stream of dry air. The zeolite is completely regenerated when the bed temperature reaches 350°C. Therefore, the purge gas needs to be heated to at least 350°C.

This paper primarily describes one part of the micro gas separation system, the vertical micromachined resistive heater.

DEVICE FABRICATION

The micro gas separation column, shown in Figure 1, is fabricated from two silicon die electrostatically bonded to a single glass die. The completed device is approximately 1.7cm x 1.7cm. The glass die is conventionally machined[1] and contains four through holes, which form the separation columns and house the zeolite adsorbent. Each silicon die is micromachined using Deep Reactive Ion Etching (DRIE) to produce a silicon screen which allows gas to flow through the device but prevents the powered zeolite from escaping. One of the two silicon die (the silicon heater die) also serves to heat the purge gas. It contains a junction isolated vertical resistive heater formed by a phosphorus diffusion which occurs after the die has been micromachined. The valves necessary for cyclic regeneration of the zeolite are not integrated with the device.
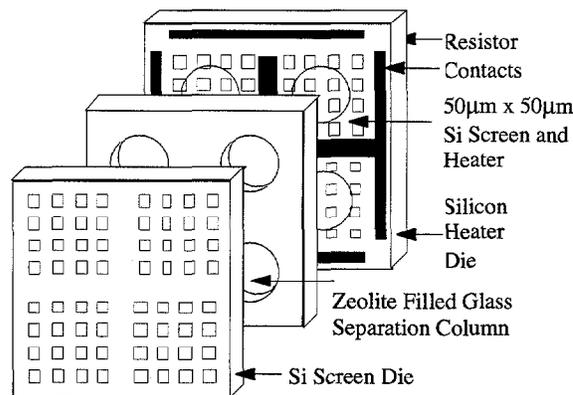

Figure 1. *Perspective drawing of the Micro Gas Separation Column.*

The resistive heater is diffused into the walls of the silicon micromachined screen. Thus, current flows vertically from the frontside Al contacts through the resistor to the backside Al contacts. Therefore, the device is inherently double-sided. The fabrication steps described below, for the resistive heater, are performed on both the back and the front of 3inch double-side polished 100 silicon wafers. The one exception is the DRIE which proceeds from one side and stops on PECVD layers deposited only on the opposite side. To facilitate connection to

---

[1] The glass wafers and the glass shadow masks were manufactured at Specialty Glass Products, Willow Grove, PA.

both the frontside and the backside contacts, the silicon heater die is larger than either the glass die or the silicon screen die.

Fabrication of the resistive heater, shown in Figure 2, begins with a photostep which patterns photoresist for a 5000Å dry plasma recess. The recess facilitates electrostatic bonding of the silicon die to the glass separation columns by recessing areas of the silicon containing steps formed from subsequent selective diffusions. The recess was omitted from the fabrication of the prototype resistive heater since these die were intended to test the function of the heater alone.

Following recess, a 1.2µm wet thermal masking oxide is grown and patterned for n+ contact diffusion. Following diffusion, the masking oxide is stripped and a new thermal oxide grown and patterned for Boron Ion Implant. After the oxide has been etched, a second thin dry oxide is grown. This oxide moves the peak of the implant doping distribution toward the surface of the silicon facilitating the formation of an ohmic contact to the lightly doped p- substrate. After Implant all oxide is removed with BHF.

The silicon screen is formed by dry etching 50µm x 50µm square channels completely through the silicon wafer. After this etch is complete, photoresist processing is far more difficult because of the presence of these through holes. Therefore, the masking oxide that will be needed after DRIE is grown and patterned before DRIE. Thus, following the removal of the implant masking oxide, 6000Å of wet thermal oxide is grown and patterned. This 'block' oxide will be used to mask the phosphorus resistor diffusion.

After the block oxide has been patterned, 5000Å of PECVD silicon nitride and 1µm of PECVD silicon dioxide are deposited on one side of the wafer. On the opposite side, a 9µm-10µm thick layer of positive photoresist is patterned for the 50µm x 50µm through holes. Following plasma descum, the wafer is etched for ~4hours in a Bosch etcher[2]. The etch proceeds through the silicon wafer, through the PECVD nitride and stops on the PECVD oxide. The stopping oxide is necessary to prevent the plasma from attacking the exposed silicon surrounding the through hole on the underneath side of the wafer. This type of attack, or etching, is sometimes called "spidering" because the etch pattern spiders out in all directions around the exit hole. A perspective of the micromachined through holes is shown in Figure 3. SEM photographs of the silicon screen are shown in Figures 4–6.

The PECVD oxide and nitride layers are then removed. This can be accomplished using either wet or dry etching processes. If dry etching is used, the frontside of the wafer resists etching since it is face down in the plasma and still covered with photoresist. Alternately, BHF can be used to remove the PECVD oxide. The photoresist is then stripped and hot phosphoric used to remove the PECVD nitride. Regardless, the block oxide on both sides of the wafer is unaffected.

It is important that the resistance of the n+ vertical resistor be larger than the resistance of the Al interconnects. The majority of the applied voltage should drop across the vertical resistor generating heat primarily in the silicon screen. Therefore, following DRIE, 600Å of dry oxide is grown on the wafer. The oxide will grow not only on the surface of the wafer but also on the surface of the square through holes. The POCOL diffusion then takes place through the oxide which lowers the

---

[2] The Bosch etching was performed at the NNUN facility at Cornell University.

dopant concentration. Note that obtaining an exact dopant concentration is not critical. It is only important that the impedance of the embedded resistor be larger than the interconnect impedance.
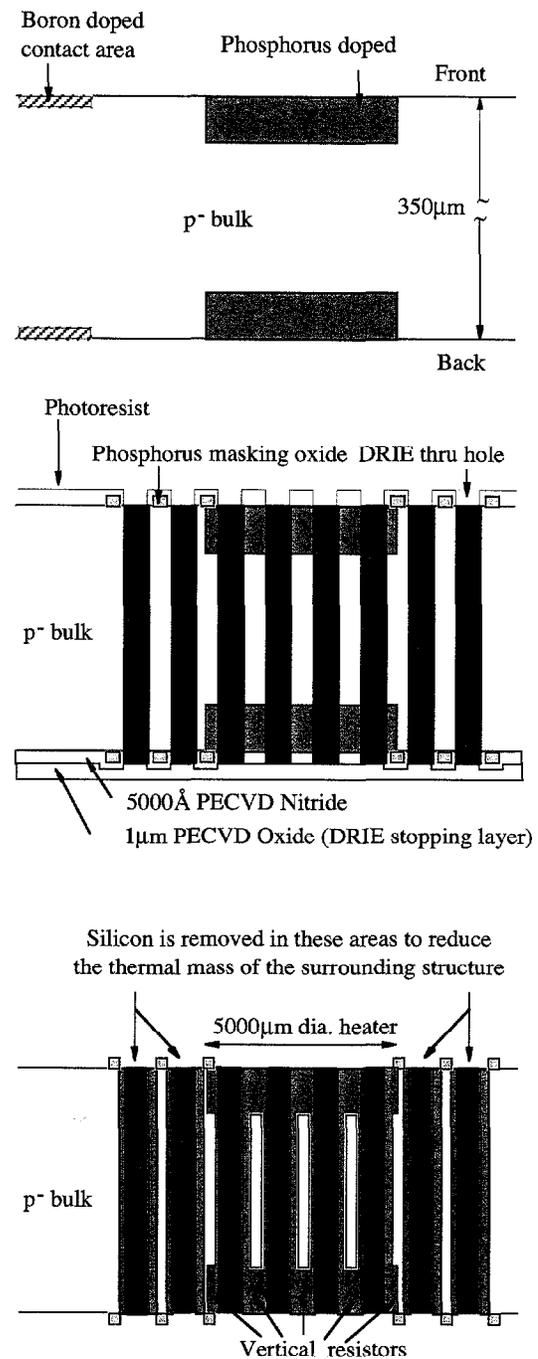

Figure 2. Fabrication Sequence for the Silicon Heater Die.

The silicon screen extends beyond the boundary of the circular glass separation column. This helps to lower the total thermal mass of the silicon die. However, it is undesirable for additional heat to be generated in this area. The block oxide is used to electrically isolate the corners of the silicon screen from the embedded resistor. Without this oxide, the entire surface of

the wafer would be doped n+ during the resistor POCOL diffusion. The through holes in the corners of the silicon screen would then become part of the embedded resistor. The block oxide is patterned in a square picture frame (15μm wide) around each DRIE through hole in the corners of the silicon screen. This prevents the resistor phosphorus diffusion from doping these areas, electrically isolating them from the embedded resistor.

After completion of the resistor phosphorus diffusion, the block oxide is removed in BHF. A glass shadow mask is then aligned and clamped to the silicon wafer. The alignment tolerances and the metal contacts are large which facilitates accurate alignment by hand through a microscope. The clamping fixture was designed to be easily inserted into a metal evaporator. 1μm of Al is then deposited. The procedure is repeated to form the backside contacts.

Fabrication of the silicon screen die begins with the growth of ~1μm of wet thermal oxide. The oxide is then selectively removed from the front side of the silicon wafer using BHF. The front of the wafer is then coated with 9μm-10μm of positive photo resist, exposed and developed opening 50μm x 50μm square holes. Following a plasma descum the wafers were etched in a Bosch etcher for ~4 hours. Removal of the photoresist and masking layers completes the process.

The glass die are machined from four inch, 800μm thick, borosilicate glass wafers. The columns are 5000μm in diameter and 800μm long. If larger zeolite volumes are desired, to accommodate higher gas flow rates, both the thickness of the glass and diameter of the columns can be increased.

CALCULATIONS

The following analysis was developed to estimate the heat transfer coefficient and the steady state power required to heat the purge gas to approximately 400°C. At steady state, the vertical resistors generate a nearly constant heat flux along the wall of each vertical channel. The bulk temperature of the gas increases linearly as it flows through the heater. An energy balance written for the control volume shown in Figure 7 yields

$$q \cdot 4s\Delta x + mc_p\left(T_{B|_x} - T_{ref}\right) - mc_p\left(T_{B|_{x+\Delta x}} - T_{ref}\right) = 0 \quad (1)$$

where q is the heat generated per unit time per unit of surface area, m is the mass flow rate, c_p the specific heat at constant pressure, and $s = 50\mu m$ is the side length of the square channel. In the limit as $\Delta x \to 0$

$$\frac{dT_B}{dx} = \frac{4sq}{mc_p} \quad (2)$$

Integrating (2) provides the average, or mixing-cup temperature of the purge gas as it travels down the length of the tube;

$$T_B = T_i + \frac{4sq}{mc_p}x \quad (3)$$

where T_B is the temperature of the gas in the tube at position x and T_i is the temperature of the gas at the inlet. Newton's Law

of Cooling states that the local heat flux is proportional to the local temperature difference, i.e.

$$q = h(T_w - T_B) \quad (4)$$

where h is the heat transfer coefficient and T_w is the wall temperature. Clearly both T_w and T_B vary linearly down the length of the channel. Combining Equations (3) and (4) yields the following expression for the silicon wall temperature as a function of the gas inlet temperature and the position x along the channel

$$T_w = T_i + q\left[\frac{1}{h} + \frac{4sx}{mc_p}\right] \quad (5)$$

For flow with a fully developed velocity and thermal profile, and a constant wall flux, h is given by [4]

$$h = 4.364\frac{k_{gas}}{D} \quad (6)$$

where D is the hydraulic diameter and k_{gas} is the thermal conductivity of the gas. Therefore, using Tables 1 and 2, the calculated heat transfer coefficient is 3000W/m²/°C.

The total power necessary to continuously heat a 1SCCM flow of nitrogen from 25°C to 400°C can be calculated by first calculating the heat flux at steady state. Using Equation (3) $q = 67.7\,W/m^2$. The total interior surface area of a single channel is $6.2 \cdot 10^{-8}\,m^2$. Thus, the heat generation rate per channel is 4.2μW/channel. The total power required to heat the purge gas is 8.24mW/bed.

The heater will not only heat the gas flowing through the channels but will also conductively heat the silicon die. The Biot number is the ratio of surface convection to internal conduction. For the silicon screen the Biot number is 0.18, which implies that the screen heats uniformly. The heat required to raise the silicon screen temperature from 25°C to 400°C is 3.11J.

Table 1. *Geometry of the Silicon Screen*

Channel side length	50μm
Hydraulic diameter	56.42μm
Channel length	350μm
Number of channels per bed	1963
Diameter of the bed	5000μm
Number of beds per die	4

Table 2. *Physical Properties of N2*

Inlet temperature	25°C
Outlet temperature	400°C
Heat capacity	1.06 kJ/kg/°K
Thermal conductivity	38.7 mW/m/°C
Mass flow rate	$2.08 \cdot 10^{-8}$ kg/sec

PRELIMINARY TEST RESULTS

The measured vertical resistance of the heater die is between 6Ω and 7Ω with the p- substrate floating. The
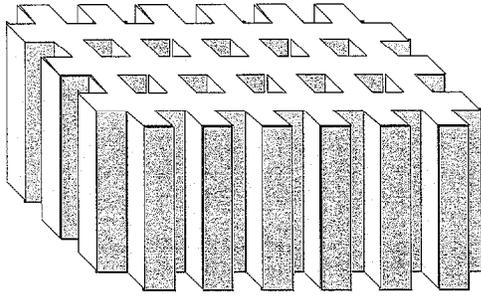
**Figure 3.** Perspective drawing of the micromachined silicon screen.
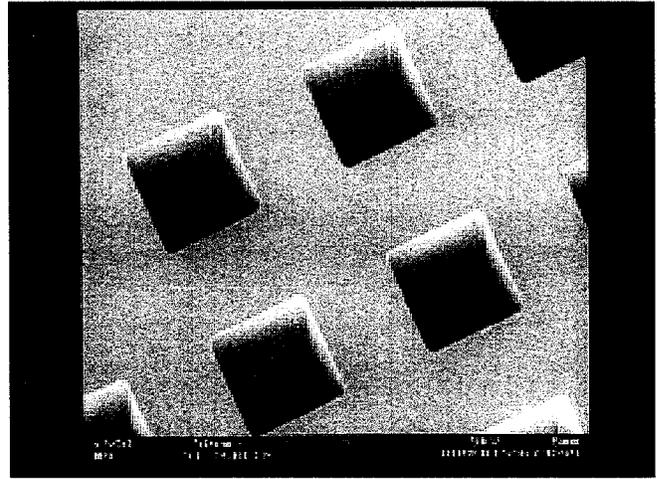
Figure 4. SEM photograph of the DRIE through holes as viewed from the frontside of the wafer. (Plasma entrance holes)
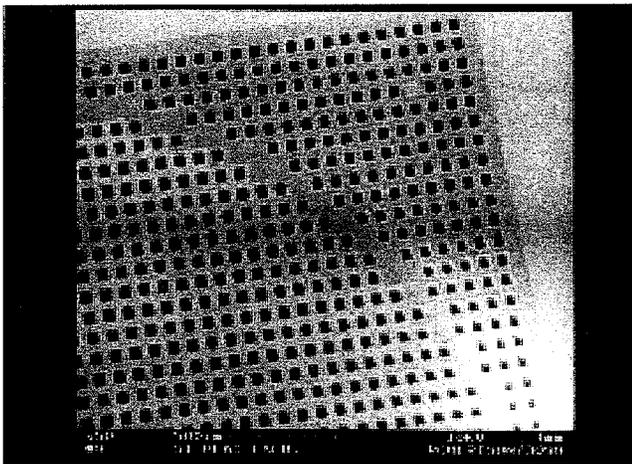

Figure 5. SEM photograph showing a portion of one silicon screen. The circular section is positioned over the zeolite filled separation column and contains the resistive heater. The corners of the structure have also been micromachined to reduce the thermal mass of the surrounding silicon.
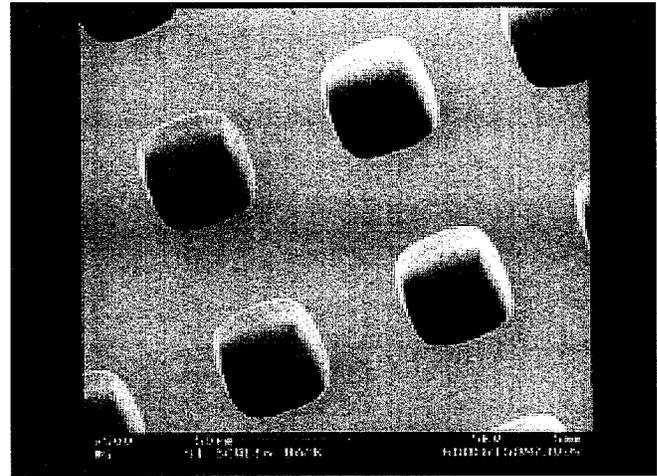

Figure 6 SEM photograph of the DRIE through holes as viewed from the backside of the wafer. (Plasma exit holes)
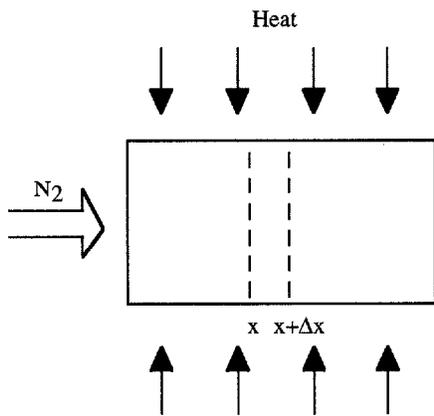
Figure 7. Control volume depicting the heat transfer through one channel.
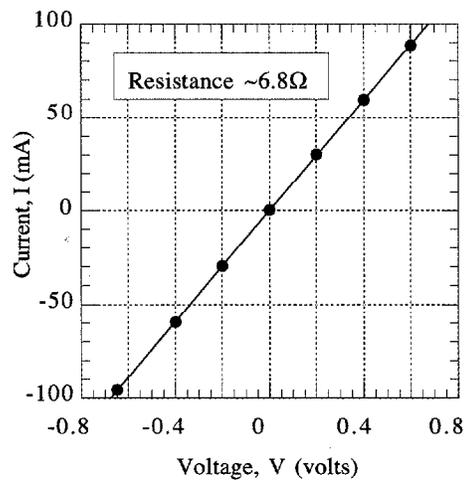
Figure 8. Vertical resistor IV characteristic measured with the substrate floating.

173

calculated resistance of the Al interconnects is < 0.025Ω. A plot of the IV characteristic is shown in Figure 8. The vertical resistor L/W ratio is approximately 1.8. Therefore, the average resistance of each channel is between 47KΩ and 55KΩ. The average sheet resistance is between 26KΩ/square and 31KΩ/square. However, sheet resistance measurements on an unetched control wafer yielded ~177Ω/square. It is possible that although significant doping occurred on the surface of the wafer (and on the surface of the control wafer) that the doping density was reduced inside the channels resulting in a larger sheet resistance for the micromachined channels.

With the substrate grounded the total resistance of the heater die was ~5Ω. The decrease is due to a large area pn junction diode formed between the n+ resistor and the p-substrate. The parasitic device should not substantially impact the operation of the heater.

At steady state, approximately 8.24mW/screen (or 32.96mW/die) is required to heat 1SCCM of N2 gas from 25°C to 400°C. Therefore, assuming an average resistance of 6.5Ω, a voltage of ~0.46Volts is required at steady state.

CONCLUSIONS

A vertical micromachined resistive heater for a micro gas separation column has been designed and fabricated. The measured vertical resistance is between 6Ω-7Ω. The calculated power required to heat the purge gas at steady state is ~33mW/die requiring a applied bias of ~0.46volts and a current of ~71mA. The calculated heat transfer coefficient is ~3000W/m²/°C.

The micromachined resistive heater presented in this paper is a demonstration of the type of vertical integration that is possible using deep reactive ion etching. The resistor diffusion occurred after the Bosch etch, thus forming the resistor along the micromachined surface. The diode is an insignificant parasitic device in this structure, but again demonstrates the use of the vertical dimension of the silicon wafer to produce a semiconductor device.

ACKNOWLEDGEMENTS

REFERENCES

1. J. K. Robertson, and K. D. Wise, "A Low Pressure Micromachined Flow Modulator", *Sensors and Actuators:A. Physical*, 71/1-2, pp. 98-106, (1998).

2. E. Meng, X. Q. Wang, H. Mak, and Y. C. Tai, "A Check-Valved Silicone Diaphragm Pump", *Proc. IEEE The Thirteenth Annual International Conference on Micro Electro Mechanical Systems,* Miyazaki, Japan, Jan. (2000), pp. 62-67.

3. N. P. Chopey, G. Ondrey, and G. Parkinson, "Microreactors Find New Niches," *Chemical Engineering*, pp. 30–33, March (1997).

4. J. D. Parker, J. H. Boogs, and E. F. Blick. *Introduction to Fluid Mechanics and Heat Transfer,* Addison-Wesley Publishing, Reading, MA (1974), p. 190.

Fabrication of Genetic Analysis Microsystems using Plastic Microcasting

P. Sethu, H. Yu[†], P. Grodzinski[†], and C. H. Mastrangelo

Center for Integrated Microsystems
Department of Electrical Engineering and Computer Science
University of Michigan, Ann Arbor, MI 48109-2122, USA
(734)-763-7162, *FAX:*(734)-763-9324
† Microfluidics Laboratory, Motorola Labs.
2100 East Elliot Road, MD EL508
Tempe, AZ 85284, USA

ABSTRACT

In this paper we present a simple, low cost, high resolution plastic microcasting method using two component, optically clear polymers for the fabrication of microfluidic systems. This method allows excellent replication of microfeatures (<1 μm), low surface roughness (< 30 nm), the ability to embed devices, and the capability of mass production. Using this method we have successfully implemented plastic microsystems capable of separating and amplifying DNA. We present the first demonstration of plastic microreaction chambers with on-chip heaters and coolers and also the fabrication and demonstration of a microcapillary electrophoresis device made by this technique.

Keywords: plastic casting, capillary electrophoresis, microPCR, genetic analysis systems.

INTRODUCTION

Miniaturized systems for biochemical analysis can significantly reduce the cycle time, reagent cost and labor intensity compared to traditional technologies [1–3]. Many of these systems are targeted for a few or even single use assay applications where it is essential that they are fabricated using inexpensive materials and technologies. Devices for these microfluidic systems have been previously made by injection molding (IM) [4] and hot embossing [5]. Casting methods for microfluidic devices were initially introduced by Whitesides [6]. Plastic microcasting is a simple, low cost fabrication technique which uses liquid polymers which are poured in a micromold and cured to form solid parts. Unlike conventional IM methods, microcasting offers excellent replication resolution but so far it has been limited to silicon rubber materials and nonplanar substrates.

In this paper we have extended the microcasting process to more rigid and less permeable epoxy materials and we have developed practical methods for the production of planar plastic casted wafers that enable further lithographic processing. Since plastic microcasting uses liquid polymers, embedding of active devices in plastics can also be done relatively easily. Our process offers excellent replication resolution (< 1 μm), a high process flexibility using a wide variety of plastic materials, the ability

to embed elements, and a low fabrication costs. High-quality sealed-channel devices are fabricated using a lamination process on top of microcasted structures. Both capillary electrophoresis and micro PCR devices with integrated heaters, coolers and temperature sensors were fabricated by microcasting.

CASTING POLYMERS

The best polymers for genetic analysis and microfluidic applications have low viscosities, high light transmission, good water barrier properties, chemical resistance to biological reagents, and stability at elevated temperatures (up to 100°C). Common castable materials include acrylics, polyurethenes, fluoropolymers, silicone rubbers and epoxies. Maximum operating temperatures for acrylics and polyurethenes are below 80°C. While the temperature and chemical resistance of fluoropolymers is excellent the light transmission is in general poor. Silicone rubbers are easily castable but have high water permeabilities [7] resulting in rapid sample loss. Epoxies offer high working temperatures, good chemical resistance, and good barrier properties (second only to fluoropolymers); therefore are the most suitable polymers for microfluidic applications.

For our applications, we selected low viscosity epoxies that were easily castable and provide an adequate working time before solidification. The epoxies used for fabricating the devices were EPOTEK 301-2 (Epoxy Technologies, Billerica, MA) with viscosity 300 cps, light transmission 98%, water absorption (room temperature, 98% RH, 30 days) 1%, and EPOTEK 301-2 FL, with viscosity 150 cps, light transmission 99%, water absorption (room temperature, 98% RH, 30 days) 1.5%. Both epoxies have a maximum operating temperatures of 125°C and a curing time of 8 hours at 45°C.

CASTING PROCESS

Figure 1 shows the simplified microcasting process. First, patterns for channels, reservoirs, chambers, and reactors are etched in a silicon wafer using a deep RIE system (Specialty Technology Systems), (SF_6:130 sscm, O_2:30 sscm, C_4F_8:80 sscm, coil power:800 Watts, and platen power:600 Watts, 30 mins.) Next, a 1 μm-thick p-xylylene (parylene-C) layer is vapor deposited on the patterned silicon substrate [8]. This thin film serves as a release layer which facilitates the separation between
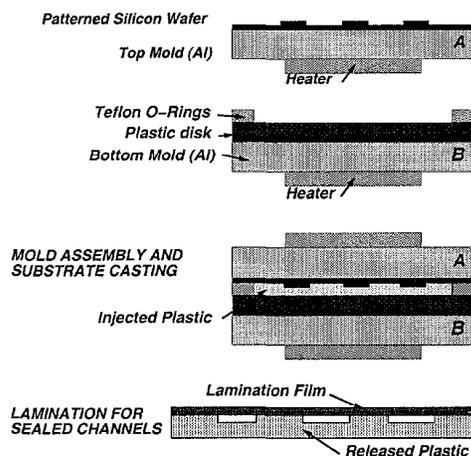
Figure 1: Microcasting and lamination process flow

the casted material and the wafer surface. The patterned silicon substrate is then used to create a mold by joining it with a supporting plastic plate and teflon ring spacer. The mold assembly is then placed between two aluminum molds as shown in Figure 2. The mold is then heated to 45 °C for 8 hrs and filled with a low viscosity catalysed epoxy (EPOTEK 301-2, EPOTEK 301-2 FL).
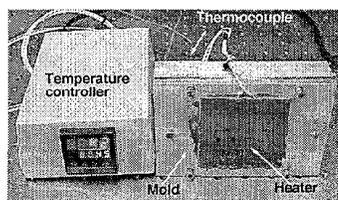

Figure 2: Complete mold setup with computer controlled heaters

Figures 3(a) and 3(b) show SEM photographs of a casted substrate 50 μm deep with channels 100 μm wide and reservoirs 5 mm wide showing the texture of the plastic surfaces which reproduce exactly the texture of the RIE etched molds (surface roughness < 30 nm). Through-substrate access holes are created either using a two mold process which uses a plastic wafer with press fitted pins on one side or by drilling holes through the substrate.

LAMINATION

In order to seal channels and reservoirs, a flexible transparent film is laminated [9] on top of the epoxy substrates. The lamination process uses a desktop roll type laminator (Kepro Circuit Systems, Inc. BLT 121-A) and a thin mylar film (2 mil, Monokote, Top-Flite, Inc.) is used to laminate the plastic substrates. The substrate is first treated with acetone, and then the film is laminated onto the substrate by passing it between two rollers. The process is carried out at room temperature and at a feed rate of 2 ft/min. Figure 4 shows the cross section of a laminated channel. These structures show excellent permeation barrier characteristics with water permeabilites P
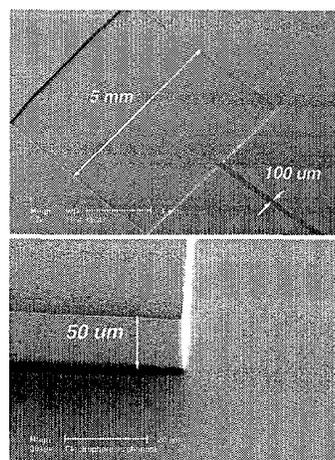

Figure 3: SEM images of (a) large-feature plastic casted microfluidic channel structures and (b) showing the side walls of the plastic casted microfluidic channel structures

< 1.7 μl/cm^2/day at room temperature. When these capillaries are filled with sample there is no seeping present at the bond interface.
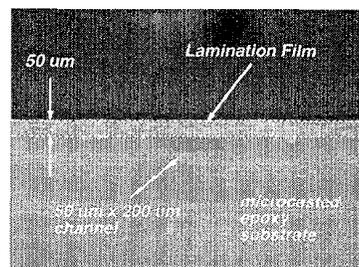

Figure 4: Cross section of a laminated channel

PCR DEVICE FABRICATION

A PCR device was fabricated using the casting process. The device is made by embedding pre-assembled components around a reaction chamber to obtain controlled heating and cooling cycles. The device consists of a reaction chamber which was made using a 1 mm diameter glass capillary (Laboratory Devices, Inc.), a resistive heating coil (Scientific Instrument services W73), a 4 × 4 × 2.2 mm^3 thermoelectric (TE) cooler (Melcor), and a 250 μm OD thermocouple (Omega Engineering, Inc.). The mold assembly for casting the PCR device is shown in Fig. 5
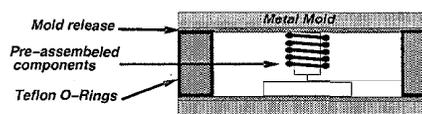

Figure 5: Mold assembly for casting the PCR device

These elements are assembled between a metal plate coated with a common injection molding release agent

and another self releasing plastic disk. All of these components are pre-assembled before setting them in the mold and embedding them in epoxy and cured. The heating coil was located around the glass capillary (reaction chamber), the TE cooler was located at the bottom of the glass capillary, and the thermocouple was placed between the glass capillary and the TE module with the tip in the bottom of the reaction chamber. Figures 6 shows a picture of the released PCR device.
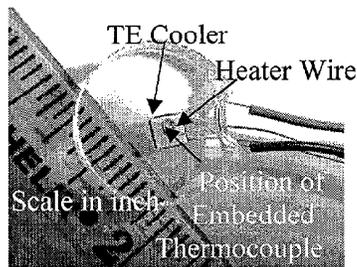

Figure 6: Photograph of micro PCR device

PCR DEVICE CHARACTERIZATION AND RESULTS

Automated temperature control was obtained using a programmed controller (PID loop, MOD30ML, from ABB). The controller was connected to a PC running a custom graphical interface. The temperature feedback from the thermocouple at the bottom of the chamber was fed to the controller which controlled both the heaters and coolers. Heating was obtained using both the TE modules and the coil heaters and cooling was obtained using only the TE module.
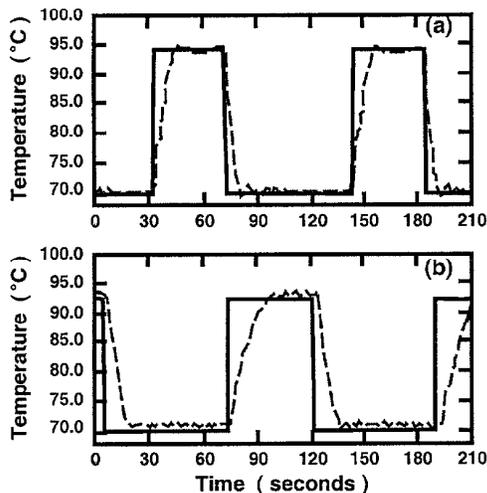

Figure 7: Thermal cycles obtained with the plastic microreaction chamber with a) both heating and TE module and b) only heating module

The heating and cooling rates were observed for a two-step temperature cycle (94°C for 30 seconds and 70°C for 60 seconds) for 24 cycles. When both the heating

coils and the TE module were turned on, a heating rate of 2.4°C/s and a cooling rate of 2.0°C/s were obtained and when the TE module was turned off a heating rate of 1.0°C/s and a cooling rate of 1.6°C/s were obtained. The graph showing the heating and cooling cycles is shown in Figure 7

Escherichia Coli cells were first thermally lysed and then the released genomic DNA segments were amplified. Approximately two-thirds of E. coli bacterial colony was collected and resuspended in 50 μl of sterile H_2O. Then this suspension was diluted 1:10 in a PCR reaction mixture containing 10 mM Tris-HCL (pH 8.3), 50 mM KCL, 1.5 mM $MgCl_2$, 0.0001% gelatin, 250 g/ml bovine serum albumin, 200 μM each deoxynucleotide triphospate, 1.0 μM each primer, 2.5 units/100 μl Ampli-Taq DNA polymerase. 1.5 μl of the mixture is loaded in the reaction chamber and is topped off with liquid wax to prevent evaporation of the sample. Cell lysis was done at 94°C for 1 minute and then thermally cycled at 50°C for 30s and 72°C for 30 seconds for 25 cycles. The primer set used to amplify a 346-bp segment of E. coli lamB gene was 5'-CTG ATC GAA TGG CTG CCA GGC TCC-3' and 5'-CAA CCA GAC GAT AGT TAT CAC GCA-3'. Then the amplified product obtained using this device was compared with the one obtained using a conventional thermal cycler. Figure 8 shows the results of gel electrophoresis on both the amplified products.


Lane 1: DNA Ladder

Lane 2: Plastic device Micro PCR (1.5 uL)

Lane 3: Control PCR reaction in conventional PCR instrument
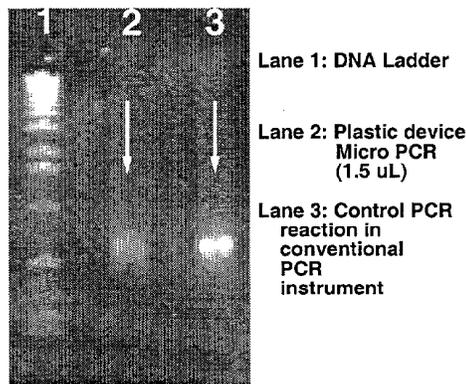
Figure 8: Separations obtained by gel electrophoresis after on-chip and off-chip PCR

CAPILLARY ELECTROPHORESIS DEVICE

Devices for capillary electrophoresis can be easily made using the casting process. A silicon wafer with a photoresist (Microposit SC1827) mask is etched using a deep RIE to get the profiles required to form the channels (50 μm deep and 100 μm wide). The separation channel is 3 cm long and the reservoir dimensions are 5 x 5 mm^2. The channels for sample introduction are 0.5 cm long. Figure 9 shows a fabricated capillary electrophoresis device.

CAPILLARY ELECTROPHORESIS RESULTS

The fabricated capillary electrophoresis devices were used to perform separations using a 0.5%(w/v) hydrox-
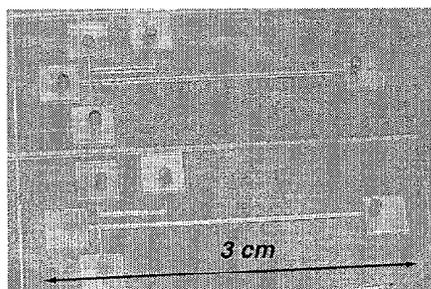
Figure 9: Microscope picture of a fabricated Capillary electrophoresis device with separation channels and reagent reservoirs

yethylcelluose (HEC) (MW 90,000-105,000) sieving matrix. Approximately $3\mu l$ of the electrophoresis buffer (0.2g HEC, 4 ml of 1x TBE buffer, and 36 ml of distilled, DI water) was introduced in the analyte waste reservoir using a stiff needle syringe. After the channels have filled completely due to capillary action, the other reservoirs are filled with the buffer and the analyte reservoir is filled with 0.5x TBE.

Pre-electrophoresis was done by applying 300 V/cm across the separation channel for 10 minutes and keeping the analyte and analyte waste reservoirs at a potential of 0 V. This is done to produce a concentration gradient in the HEC while increacing the concentration of HEC in the column above 0.5%. Approximately 2 μl of a DNA sample labeled with SYBR Green I at an intercalating ratio of greater than 1:5, dye:DNA bp is loaded into the injection reservoir. Samples were then separated under an electric field of 110 V/cm in a distance of 3 cm. Figure 10 shows the DNA plug and separations obtained using the plastic capillary electrophoresis devices. The plastic has very low background fluorescence and does not interfere with the fluorescence of the DNA bands.
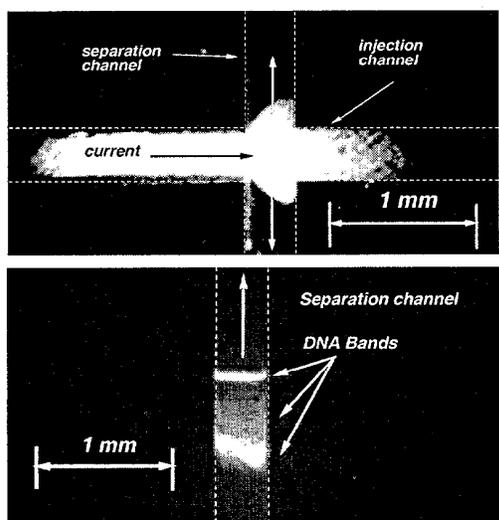

Figure 10: Images of (a) Injection plug and (b) DNA bands obtained by capillary electrophoresis

SUMMARY

We have presented a low cost casting method for the fabrication of microfludic devices. Using this casting method we have developed devices for genetic analysis applications like PCR and capillary electrophoresis. We have also have demonstrated working devices with measured results.

ACKNOWLEDGMENTS

References

[1] A. T. Woolley, D. Hadley, P. Landre, A. J. deMello, R. A. Mathies, and M. A. Northrup, "Functional integration of pcr amplification and capillary electrophoresis in a microfabricated dna analysis device," *Anal. Chem.*, vol. 68, pp. 4081–4086, 1996.

[2] S. C. Jacobson, J. P. Kutter, C. T. Culbertson, and J. M. Ramsey, "Rapid electrophoretic and chromatographic analysis on microchips," *Micro Total Analysis Systems '98, Banff, Canada*, 1998.

[3] M. A. Northrup, B. Benett, D. Hadley, P. Landre, S. Lehewa, J. Richards, and P. Stratton, "A miniature analytical instrument for nucleic acids based on micromachined silicon reaction chambers," *Anal. Chem.*, vol. i70, p. 918922, 1998.

[4] T. D. Boone, H. H. Hooper, and D. S. Soane, "Integrated chemical analysis on plastic microfluidic devices," *Solid State Sensor and Actuator Workshop '98, Hilton Head Island, South Carolina*, 1998.

[5] H. Becker and W. Dietz, "Microfluidic devices for micro-tas applications fabricated by polymer hot embossing," *SPIE Conference '98, Santa Clara, California*, 1998.

[6] E. Kim, Y. Xia, and G. M. Whitesides, "Polymer microstructures formed by molding in capillaries," *Nature*, vol. 376, p. 347, 1996.

[7] R. K. Traeger, "Nonhermeticity of polymeric lid sealant," *IEEE Transactions, Parts, Hybrids and Packaging*, vol. PHP-13, pp. 147–152, 1977.

[8] P. F. Man, C. H. Mastrangelo, M. A. Burns, and D. T. Burke, "Microfabricated capillarity-driven stop valve and sample injector," in *International Workshop on Micro Electro Mechanical Systems (MEMS 98), Heidelberg, Germany*, pp. 45–50, 1998.

[9] J. S. Rossier, M. A. Roberts, R. Ferrigno, and H. H. Girault, "Electrochemical detection in polymer microchannels," *Analytical Chemistry*, vol. 71, pp. 4294–4298, 1999.

MICROFABRICATED PLANAR PRECONCENTRATOR

Ronald P. Manginell, Gregory C. Frye-Mason, Richard J. Kottenstette,
Patrick R. Lewis, and C. Channy Wong
Sandia National Laboratories
Albuquerque, New Mexico, 87185-1425

ABSTRACT

Front-end sampling or preconcentration is an important analytical technique and will be crucial to the success of many microanalytical detector systems. This paper describes a microfabricated planar preconcentrator ideal for integration with microanalytical systems. The device incorporates a surfactant templated sol gel adsorbent layer deposited on a microhotplate to achieve efficient analyte collection, and rapid, efficient thermal desorption. Concentration factors of 100-500 for dimethyl methyl phosphonate (DMMP) have been achieved with this device, while selectivities to interfering compounds greater than a factor of 25 have been demonstrated. Device performance will be compared with conventional preconcentrators, and the effects of system flow rate, flow channel geometry and collection time will be presented. A physical model of adsorption/desorption from the device will be reviewed and compared with experiment, while numerical simulation of flow over the device will be described.

INTRODUCTION

Chemical preconcentration is a critical analytical procedure since it can purify sample mixtures and boost small analyte concentrations within the range of detectability of a given sensor. This technique is well known in the field of gas chromatography (GC) where a preconcentrator (PC) typically consists of a small tube, often 6 mm in diameter by 100 mm long, packed with adsorbent resins between two glass wool plugs. Desired analytes are first collected on the resin by passing a sample mixture through the PC. Then, temperature ramping (~10 °C/sec) is used to thermally desorb the collected analytes as a plug of gas. The preconcentrator is physically located at the front end of a GC column and the plug is fed into the GC for separation into its constituents.

Previously, Kovacs et. al. [1] and Neuhold and Wang et. al. [2] demonstrated *electrochemical* preconcentration in microfabricated, electrochemical-based detection systems. Their techniques, however, are fundamentally different than that described in this paper. We have developed a microfabricated preconcentrator for *gas-phase* microanalytical detection systems similar in function to that described above for conventional GC systems. Here, a surfactant templated (ST) sol gel layer, deposited on a microhotplate, is used as the adsorbent and collects analytes from a sample stream directed over its surface. With surface areas as high as 1000 m^2/g and the ability to tailor surface chemistry and pore size for specific analytes, this class of materials is very attractive as an adsorbent layer. The rapid heating capability of the underlying microhotplate is used to thermally desorb analytes collected in the adsorbent. Heating rates of the order of 10^4 °C/sec represent at least three orders of magnitude improvement over conventional systems. Given this rapid ramp rate, and, given the low steady-state power consumption (100 mW at 200 °C, a typical desorption temperature), power consumption in the microfabricated device is at least three orders of magnitude less than conventional systems. These capabilities make this microfabricated design suitable for portable detection systems.

Though useful as a front-end sampling stage for many microanalytical systems, this device has been used primarily in Sandia's µChemLab™ detection system [3,4]. To achieve the goal of rapid (1-3 min), trace detection of target analytes in a hand-held,

battery-powered, autonomous system, the µChemLab™ relies on three critical microfabricated components: a deep-etched, silicon-based GC separation column [5], a chemically selective SAW array sensor [6], and the preconcentrator described in this paper.

FABRICATION AND PACKAGING

The preconcentrator microhotplate has been fabricated by through-wafer silicon etching (via the Bosch process), stopping on a low-stress silicon nitride membrane layer [7]. Prior to silicon etching, a thin-film Pt heater is patterned on the membrane layer on the opposite side of the wafer from the etch window; typically, 1000 Å of Pt and a 150 Å Ti adhesion layer are used (Figure 1). Through-wafer etching is preferred to front-side etching for the fabrication of preconcentrator microhotplates, since in the latter embodiment analyte could potentially bypass the adsorptive coating by flowing underneath the membrane. All of the devices described in this paper were Bosch etched. However, wet etchants such as KOH and TMAH could have been used as well.

The adsorbent ST sol gel layer was applied to the front surface of the microhotplate after the hotplates were etched and the wafer was diced. In most cases, the adsorbent is spray coated, though spin and dip techniques have been demonstrated.

The majority of the test data was taken with a planar preconcentrator mounted in a 24-pin DIP package. In this case, sample flow was directed over the PC by means of a Pyrex lid with a machined flow channel. In the µChemLab™ system, an electrical and fluidic circuit board packaging scheme is used [4]. PC performance for both packaging schemes is comparable.
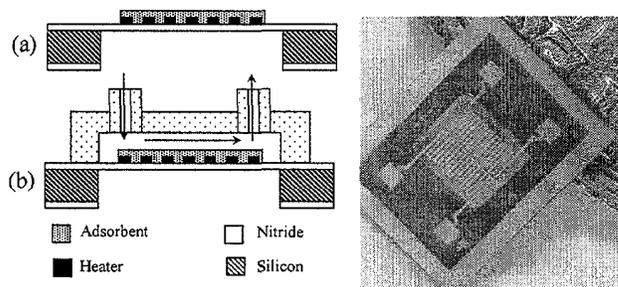

Figure 1. (a) A sectional view of a planar PC. (b) The PC with a glass lid used for testing. At right, a completed device on the edge of a quarter. The microhotplate membrane is 2.5 mm on a side.

DEVICE PERFORMANCE

The operation of the planar preconcentrator can be broken into two basic steps. First, analytes are selectively collected in the adsorbent by passing a carrier gas (nitrogen) and analyte mixture over its surface. In this step the underlying microhotplate remains in the off, or unheated state. Once sufficient analyte has been collected, the rapid heating capability of the microhotplate is exercised to flash desorb the analyte, returning it to the carrier stream as a purified and concentrated plug. Presented data pertains to DIP packaged PCs with spray-coated adsorbent layers encapsulated by glass flow channels.

Figure 2 demonstrates thermal desorption of dimethyl methyl phosphonate (DMMP) into narrow peak widths, only 200 msec full-width at half maximum (FWHM). The sample was collected by passing 5 ppm of DMMP over the PC for 1 minute. Desorption

was achieved by pulsing the microhotplate to 200 °C in 10 msec by the application of a square-wave voltage pulse. Four repetitions of this experiment show reproducibility in peak height and width. Note that the data were renormalized in time to the maximum of one of the curves without loss of generality, since, for a given device, the desorption peak is uniquely determined by the time at which the microhotplate is activated.

The *concentration factor* (*f*) specifies the concentration of analyte desorbed relative to the free stream concentration. As might be expected, *f* is directly proportional to the total collection time (t_c) at a specified flow rate and desorption temperature. For example, for a ST sol gel on a planar PC, collection times of 30 s, 60 s and 90 s yield concentration factors of 80, 180 and 300, respectively, for the analyte DMMP. The fact that large values of *f* can be achieved with short collection times indicates that the planar PC can be used for rapid-response detection systems. Figure 3 illustrates this concept for 50 ppb of DMMP after only 4 sec of collection.

Flow rate also impacts the collection and desorption processes. Generally speaking, collection efficiency is inversely proportional to flow rate. However, desorption peak width is inversely proportional to flow rate due to the relatively larger affect of diffusion processes at low flow. For instance, for a typical PC, a change in flow rate from 6.5 ml/min to 1.9 ml/min broadens the peak width from 50 msec to 250 msec FWHM.

The adsorbent coating can be tailored for specific analytes giving high values of *f* and high selectivity to interferants. Data for the analyte DMMP and two interferants, xylene and methyl ethyl ketone (MEK), are given in Table 1. High selectivity permits low-level detection of analytes in real-world environments where interferant concentrations can be much greater than that of the analyte (Figure 4). Note that the peak width of Figure 4 was broadened by an in-line, conventional capillary GC column.

Table 1. PC concentration factor and selectivity (relative to DMMP) for a 1 minute collection on a ST sol gel coated PC.

Compound	*f*	Selectivity
DMMP	510	1
Xylene	8	64
MEK	18	28

The magnitudes of *f*, selectivity, peak width and flow rate described are relevant to the μChemLab™. First, to achieve adequate separation in the 1 m long microfabricated GC column in a 1-3 minute period with a low-flow miniature pump, the PC must provide pulse widths less than 250 msec. The flow generated by the miniature pump for the connection of the PC, GC and SAW detector is in the range of 1 to 3 ml/min. It is important to emphasize that the rapid heating rate of the PC microhotplate is responsible for the narrow peak widths generated, in spite of the low flow through the system. The presented concentration factors are 10-20 times better than those generated with conventional macroscale preconcentrators.

MODELING AND SIMULATION

Thermal modeling of the preconcentrator microhotplate platform is described in reference [8]. This section summarizes a physical model of adsorption onto, and desorption from the planar preconcentrator, as well as numerical simulation of flow over the device. The goal of the modeling effort is to direct the PC design for maximal analyte collection (within a specified collection time), while minimizing the temporal width of the desorption peak. The relevant design parameters are the adsorbent area, coating material and the flow channel geometry, particularly its height, to satisfy the performance requirements of the μChemLab™.
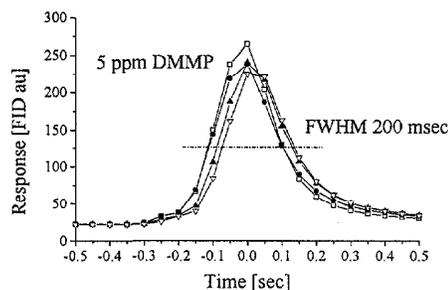

Figure 2. Four repetitions of thermal desorption of DMMP from the planar PC, as detected by a Flame Ionization Detector (FID) in a commercial, benchtop GC system. Flow was ~3 ml/min.
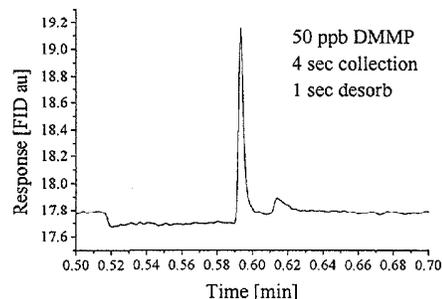

Figure 3. DMMP detection in only 5 sec demonstrating applicability of the PC for rapid response scenarios. The gas flow rate was ~3 ml/min.
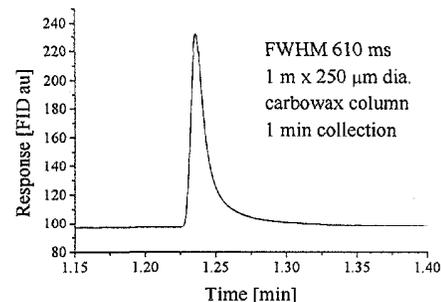

Figure 4. Detection of 50 ppb of DMMP in a 50 ppm background of xylene. Gas flow was ~2 ml/min.

Modeling Adsorption

The "breakthrough curve" gives the time dependence of the sample concentration *eluted* from the preconcentrator during adsorption. It must be emphasized that the breakthrough curve is measured *downstream* from the PC and quantifies the amount of sample that is adsorbed. An analytical model of the breakthrough curve has been developed based on several assumptions: (1) adsorption is irreversible, (2) the adsorption isotherm is linear, (3) the effect of the heat of adsorption is minimal, and (4) the flow in the channel is two-dimensional and laminar, with a parabolic velocity profile. These assumptions simplify the analysis somewhat and lead to an analytical model that illuminates the relative significance of the design parameters. However, the assumptions mandate experimental validation and detailed simulation to support any findings.

The present adsorption model separates the analyte concentration profile in the adsorbent bed into two distinct regions: a saturation zone in which the bed is in equilibrium with

the feed mixture (analyte and carrier gas) and an adsorption zone in which uptake is still occurring. Prior to the development of a saturation zone at the leading edge of the bed, the breakthrough concentration of the analyte takes on a minimum value given by

$$c_{min} = c_0 \cdot \exp(-\alpha \cdot l), \quad with \quad \alpha = \frac{kda}{uh}. \quad (1)$$

The remaining symbols used are defined in Table 2. The time required to establish a saturation zone at the front end of the bed is

$$t_{sat} = q_0 \cdot (1 - \varepsilon)/(\alpha \cdot c_o \cdot u). \quad (2)$$

As the saturation front propagates across the bed, the breakthrough concentration as a function of load time, t, is

$$c = c_o \cdot \exp\left[-\alpha \cdot \left(l - \frac{u \cdot c_o}{q_o \cdot (1 - \varepsilon)} \cdot (t - t_{sat})\right)\right]. \quad (3)$$

A series of tests have been conducted in which feed concentration, ambient temperature, and channel height have been varied for the purposes of understanding the adsorption and desorption processes, and to evaluate the analytical model. Figure 5 compares the predicted and measured breakthrough curves for adsorption. The model predicts that the breakthrough concentration will rapidly fall to c_{min} after adsorption begins at time, $t = 0$. The concentration will remain there until the leading edge of the bed becomes saturated. Then, the concentration will increase as the saturation zone propagates across the bed. Eventually the concentration will return to its initial value when the bed is fully saturated. Observed differences between the predicted and experimental breakthrough curves are mainly attributable to the failure of the model to account for the reversibility of the adsorption process.

Modeling Desorption

During the thermal desorption process, the release of analyte is assumed to be uniform across the bed, but overall changing with time. (There is no saturation front propagating across the bed.) Analyte is released into the environment due to the change in adsorption equilibrium. When the temperature of the bed is raised, desorption occurs and the concentration of analyte in the adjacent gas phase increases rapidly. The concentration gradient is thus the driving force for diffusion, leading to an increase in concentration of eluted analyte. Desorption will proceed until the concentration of analyte in the gas mixture adjacent to the bed has regained equilibrium with the feed concentration. Hence, it is very important that the model account for the transient behavior of the analyte concentration on the surface of the bed throughout the desorption process.

Taking this transient behavior into account, the breakthrough curve acts as follows: desorption is initiated at $t = 0$ when the temperature of the adsorbent is elevated. A sudden sharp spike in the breakthrough concentration is observed, followed by an exponential decay. During desorption the concentration exiting the PC is given by

$$c = A(t) \cdot [1 - \exp(-\alpha \cdot l)] \quad + \quad c_o \exp(-\alpha \cdot l), \quad (4)$$

where

$$p = \frac{ka}{(1 - \varepsilon) \cdot (q_o / c_1)}, \quad (5)$$

$$A(t) = c_1 \cdot \exp(pt) + c_o \cdot [1 - \exp(-pt)]. \quad (6)$$

Eqn. (4) shows that the width of the desorption peak is affected by the geometry of the adsorption bed, the adsorption isotherm at the elevated desorption temperature, and the mass transfer coefficient.

The change in the adsorption isotherm with temperature was determined from detailed analysis of desorption data; this analysis is beyond the scope of this paper, but will be provided in a

subsequent discourse [9]. Using the derived isotherm, the desorption breakthrough was determined. Figure 6 shows reasonable comparison between the prediction and experiment.

Table 2. Model definitions.

Symbol	Definition
c	analyte concentration exiting PC
c_o	analyte concentration in the feed stream
k	gas-phase mass transfer coefficient
a	adsorbent surface area per adsorbent volume
d	adsorbent layer thickness
l	adsorbent layer length
u	average channel flow rate
h	flow channel height
c_1	saturation concentration at desorption temperature
q_o	analyte concentration in the bed
ε	void fraction

Computational Simulation

Figure 1b illustrates the packaging scheme used here. The width of the channel is 2.8 mm, while the adsorbent is about 2.1 mm wide. The height of the flow channel is small, in most cases 150 μm, while for some experiments heights of 50 μm and 100 μm were used. Finally, gas entry and egress to the lid was accomplished by means of glass capillaries with an inner diameter of 500 μm. Based on this geometry and usual flow rates, the concern arose that the gaseous flow may be confined to a narrow region between the inlet and outlet. In this scenario, some of the adsorbent area would not be well exposed to the sample stream, leading to degraded collection efficiency. To address this concern, a three-dimensional computational flow simulation using MPSalsa was performed.

MPSalsa is a computer program developed at Sandia for laminar, low Mach number, three-dimensional incompressible and variable density reacting fluid flows; the code is solved on massively parallel computers, using a Petrov-Galerkin finite element formulation. Figure 7 shows the velocity profile in the chamber predicted by MPSalsa. For this geometry, gaseous flow in the chamber is three-dimensional and the flow across the adsorbent is relatively uniform in the transverse direction. This implies that the capture efficiency should be reasonably good and the assumptions of the analytical model are valid.

Modeling and Simulation Summary

The adsorption and desorption model reviewed here is a useful tool for the optimization of the PC for maximal collection efficiency and minimal desorption peak width. The assumptions required to obtain an analytical solution imply the need for detailed model validation through experiment and simulation. Finally, to numerically model flow over the PC, the MPSalsa code was used. Simulation showed that flow over the adsorbent is relatively uniform for the given geometry and flow conditions.

CONCLUSIONS

A microfabricated planar preconcentrator has been demonstrated that is ideal for use as a front end in microanalytical and portable systems. The combination of microhotplate technology with high-uptake, tunable adsorbents results in concentration factors and selectivity superior to conventional macroscale devices in many instances, while yielding extremely narrow desorption pulse widths at low power.

An analytical model of adsorption and desorption has been presented that can be used to perform parametric analysis and to optimize the preconcentrator design. The model requires reliable information regarding the characteristics of the adsorbent bed that can be extracted from adsorption/desorption studies. Future correspondence will concentrate on the details of the model and parameter extraction.
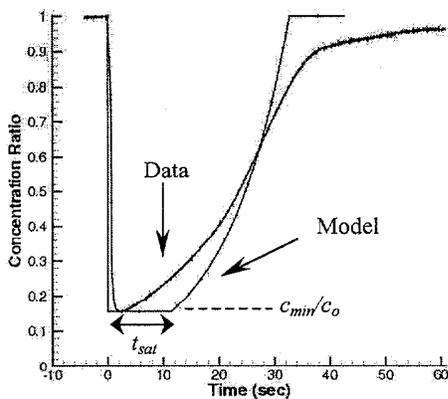
Figure 5. *A comparison of the predicted and measured adsorption breakthrough curves. (100 ppm DMMP at 3.3 ml/min.)*
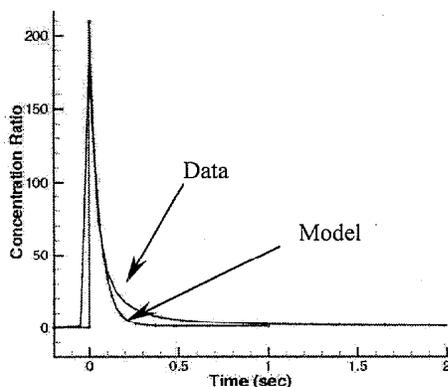

Figure 6. *Comparison between the measured and predicted desorption breakthrough curves (100 ppm DMMP at 3.3 ml/min.).*
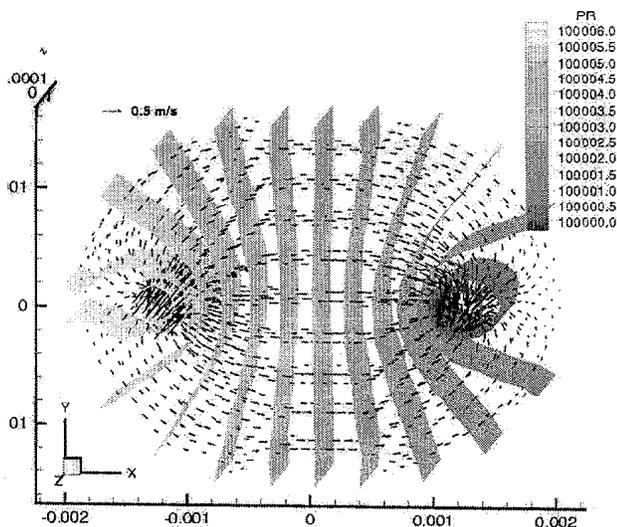

Figure 7. *MPSalsa prediction of the distributions of velocity and pressure ("PR" in Pa) in the test flow chamber (150 μm height) of the PC. Flow proceeds from the left to the right.*

ACKNOWLEDGMENTS

REFERENCES

1. G.T.A. Kovaks, C. W. Storment, and S. P. Kounaves, "Microfabricated Heavy-Metal Ion Sensor", *Sensors and Actuators B, 23*, 41 (1995).

2. C. G. Neuhold, J. Wang, X. H. Cai, and K. Kalcher, "Screen-Printed Electrodes for Nitrite Based On Anion-Exhanger-Doped Carbon Inks", *Analyst, 120*, 2377 (1995).

3. G.C. Frye-Mason, R. J. Kottenstette, E.J. Heller, C. M. Matzke, S.A. Casalnuovo, P.R. Lewis, R. P. Manginell, W. K. Schubert, V. M. Hietala, and R. J. Shul. "Integrated Chemical Analysis System for Gas Phase CW Agent Detection", *Proceedings of the Micro Total Analysis Systems (μ-TAS) Workshop*, Banff, 10/12-16/98, Kluwer, (1998) pp. 477-481.

4. G. Frye-Mason, R. Kottenstette, P. Lewis, E. Heller, R. Manginell, D. Adkins, G. Dulleck, D. Martinez, D. Sasaki, C. Mowry, C. Matzke, and L. Anderson, "Hand-Held Miniature Chemical Analysis System (μChemlab) for Detection Of Trace Concentrations of Gas Phase Analytes", *Proceedings of the μ-TAS '00 Workshop*, Enschede, Netherlands, 5/14-18/00. (In press.)

5. C.M. Matzke, S.A. Casalnuovo, G.C. Frye-Mason, R. Kottenstette, R.P. Manginell, D.Y. Sasaki, and C. C. Wong, "Integrated Silicon Gas Chromatographic Micro-Channel", *Proceedings of The SPIE Micromachining and Microfabrication Conference*, Santa Clara, 9/20-22/98, SPIE Vol. 3511 (1998) pp. 262-268

6. E. Heller, V. Hietala, R. Kottenstette, R. Manginell, C. Matzke, P. Lewis, S. Casalnuovo, G. Frye-Mason, "An Integrated Surface Acoustic Wave-Based Chemical Microsensor Array for Gas-Phase Chemical Analysis Microsystems", *Proceedings of the Electrochemical Society*, Honolulu, Hawaii, 10/17-27/99, ECS Vol. 99-23, (1999) pp. 138-142.

7. R. P. Manginell, G. C. Frye, R. J. Shul and C. G. Willison, "Microfabrication of Membrane-Based Devices By HARSE and Combined HARSE/Wet Etching", *Proceedings of The SPIE Micromachining and Microfabrication Conference*, Santa Clara, 9/20-22/98, SPIE Vol. 3511 (1998) pp. 269-276.

8. R. P. Manginell, D. A. Rosato, D. A. Benson, and G. C. Frye-Mason, "Finite Element Modeling of a Microhotplate for Microfluidic Applications", *Proceedings of the Second International Conference on Modeling and Simulation of Microsystems*, San Juan, Puerto Rico, 4/19-21/99, Computational Publications, Cambridge (1999) pp. 663-666.

9. To be published in the *ASME 2000 International Mechanical Engineering Congress & Exposition*, Orlando, FL, 11/5-10/2000.

DISTRIBUTION AND MIXING OF REAGENTS ON MULTICHANNEL PLASTIC CHIPS

Mingqi Zhao, Travis D. Boone, X. Charmaine Qiu, and Antonio J. Ricco*

ACLARA BioSciences, Inc.

1288 Pear Avenue, Mountain View, CA 94043

Fax: (650) 210-1210, email: aricco@aclara.com

ABSTRACT

We report several methods to monitor and improve reagent distribution and mixing on an acrylic card designed to demonstrate multiplexed distribution. Results show that the even distribution of a reagent from a common central well to multiple channels is strongly dependent on the chemical state of the microchannel surfaces. We have characterized the performance in acrylic microchannels treated by solution, polymer adsorption, and chemical vapor-phase processes. We also describe several pumping systems built on our multiplex distribution cards that

Figure 1. Fluorescent image of a 64-channel HTS assay card designed and produced at ACLARA. The key features of one channel are magnified in the center, showing sample injection and separation of the fluorescent marker ("Control"), the substrate, and the product formed by the action of a kinase enzyme on the substrate.

utilize electroosmotic flow as the motivating mechanism. By coating the appropriate channels with charged species, these systems have been demonstrated as a method of reagent distribution.

INTRODUCTION

Microfluidic devices integrating mixing, reaction, and separation functions have drawn great attention in the last decade because of their potential application to genetic analysis, pharmaceutical development, and medical diagnosis. Two critical

technical problems are being addressed to hasten commercialization of these devices: (1) fabrication of plastic microchips by an economical method of mass-production; (2) reduction in the quantities of reagents used. Specifically, we have developed processes to fabricate microfluidic devices using inexpensive plastics and economical molding techniques. In most cases, the microfluidic chips require sub-microliter sample and reagent volumes, even for hour-long incubation times [1-3].

Disposable plastic microchips offer additional advantages for microanalytical systems [1-4]. For example, they minimize carry-over and resultant cross-contamination, which is important for high-throughput screening (HTS) of pharmaceutical candidate compounds. In HTS, the action of an enzyme upon a substrate is quantified in the presence of each member of a "library" of hundreds of thousands of chemical compounds, the goal being to discover which library compounds significantly retard the enzyme/substrate reaction. On a single 96-channel disposable LabCard™ chip developed at ACLARA, we can accomplish up to 96 separate assays in highly parallel fashion without cross-contamination.

If each assay is conducted in an independent fluidic "microcircuit," as shown in Figure 1, all reagents (enzyme, substrate, test compound) must be dispensed individually into fluid reservoirs, requiring on the order of 100 nL of each reagent. To in particular—reagent distribution from a common reservoir shows promise. This strategy requires uniform distribution of reagents


Figure 2. Schematic of the repetitive unit of a multichannel chip Reagent distribution, mixing, mixture incubation, sample injection and separation are integrated on one chip.

from the common reservoir to each of several channels, as well as efficient mixing of the enzyme and substrates on the distribution chip. The first part of this report describes several methods of distributing reagents evenly in microfluidic systems, and the

second part describes methods to monitor and control the mixing of the reagents on the chip.

EXPERIMENTAL DETAILS

Chip Design and Fabrication

The multichannel chip used in this report has the repetitive unit shown in Figure 2. Reservoirs **1** and **2** contain enzyme and substrate, respectively, and buffers are added to reservoirs **3**, **4**, and **5** prior to analysis.
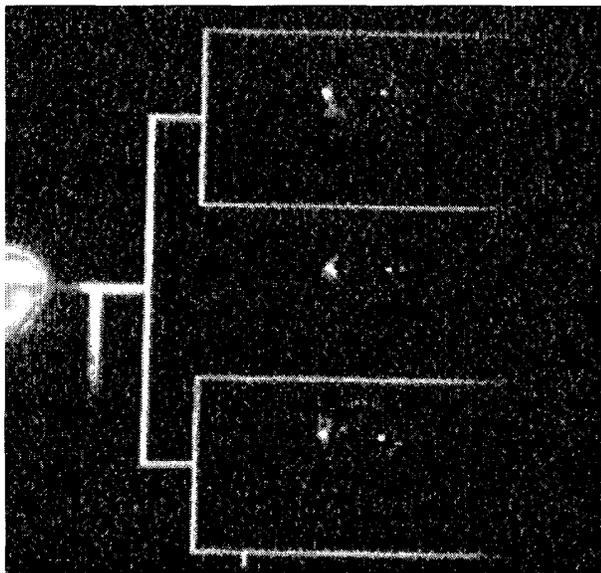

Figure 3. Electrokinetic distribution of fluorescein from a central well to 4 microchannels on a multichannel LabCard^{TM} plate.

The card is fabricated by ACLARA by compression molding of plastic resin using a micromachined Si "master" wafer as the template and completed by lamination with a polymer film [3]. Distribution and mixing of reagents are accomplished by electrokinetic motivation of the liquids [1-6]. Voltages applied to electrodes in the reservoirs are programmed for sample injection, mixing, and separation.

Surface Modification of Microchannels

Several surface treatment methods have been used to control the surface chemistry of the microchannels. Plasma treatment, chemical reaction, or adsorption of OH⁻ is used to create fixed charge on acrylic surfaces. A surface bearing fixed charge of a given sign also can be used to bind charged polymers from solution to create an oppositely charged surface.

RESULTS AND DISCUSSION

Distribution of Reagents

The surface charge of microchannels in microfluidic chips determines the direction and velocity of electroosmotic flow (EOF) when a potential is applied; this effect is superimposed upon electrophoretically induced motion. Therefore, uniform distribution of reagent from a central well to multiple channels is strongly dependent on the chemical state of the microchannel

surfaces. We have characterized the performance in acrylic microchannels that are untreated as well as those chemically treated by solution, polymer adsorption, and chemical vapor-phase plasma processes.

We use fluorescein (absorbance max.: 490 nm, emission
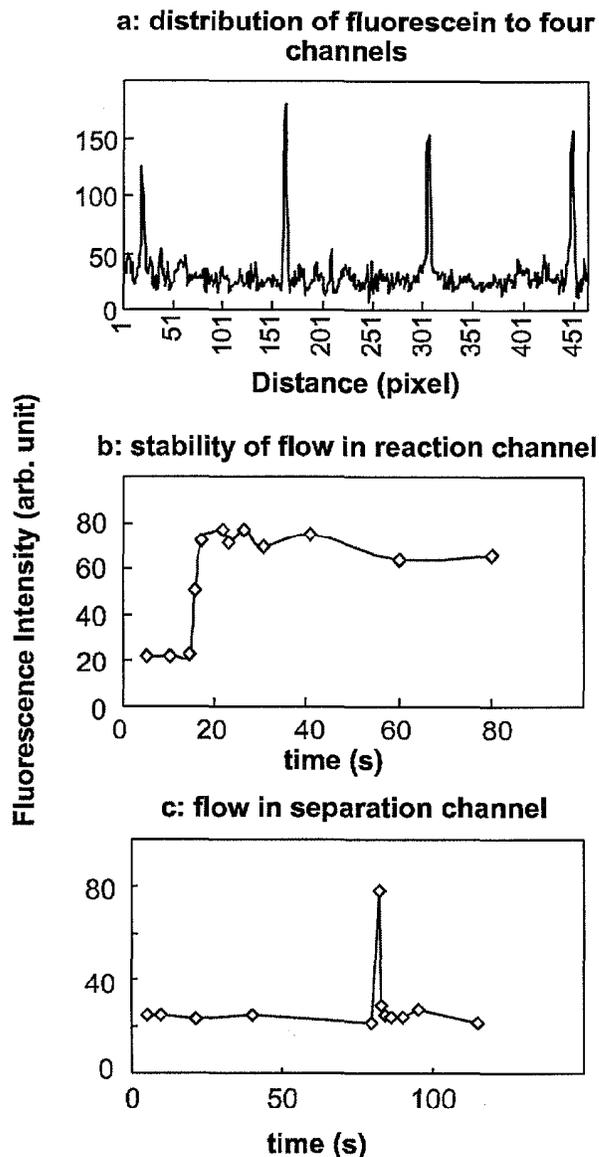
Figure 4. Digital data transferred from images recorded on CCD camera, showing the distribution and separation using a 4-channel LabCard^{TM} device as shown in Figure 2: (a) flux of distributed 1 mM fluorescein in the 4 reaction channels at steady state; (b) time course of fluorescence signal in one channel as 0.5 mM fluorescein is distributed to it and several identical channels; (c) 0.5 mM fluorescein plug in a separation channel following electrokinetic injection.

max.: 514 nm) as a marker to reveal the distribution of reagent within the channels. The fluorescent images are obtained using a home-built imaging system. Although distribution of fluorescein

is not perfectly uniform with untreated acrylic channels, distribution uniformity is improved significantly using aqueous OH⁻ pretreatment, and the most stable, uniform distribution is achieved when the channel is coated with a charged polymer (Figure 3). Even distribution results are also observed for chips laminated with films pretreated with a gas-plasma discharge and having channels coated with positively charged polymers. If the channels are further modified with negatively charged polymers atop the positively charged polymers, the dye does not distribute as evenly from the central well 1 to the channels.

We developed a method to obtain quantitative information about reagent distribution on multichannel chips. In this method, images are recorded using a CCD camera and are then

converted to digital format using XCAP (EPIX, Buffalo Grove, IL) and Image-Pro Plus software (Media Cybernetics, Silver Spring, MD). Figure 4a shows the flux of fluorescent molecules in each channel obtained by converting fluorescence intensity to a numerical value along a line perpendicular to the direction of flow through the four channels on the multiplexed distribution card.

The results shown in Figure 4a indicate that the amount of fluorescein in each channel varies by no more than 20%. By monitoring fluorescence intensity at a fixed position as a function of time, we have found the flow in the reaction channels to be quite stable (Figure 4b). Electrokinetically driven injection of the fluorescein from the reaction channel produces the expected
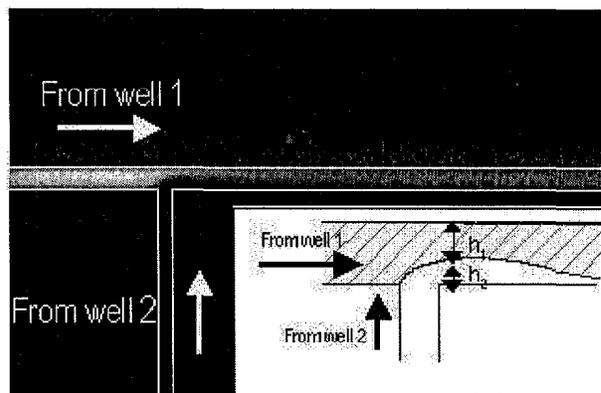

Figure 6. Mixing of reagents at a T-junction from two channels in a plastic fluidic card. The ratio of the mixing of reagents, h1/h2, is almost equal to the ratio of the currents passing through the channels, Ia/Ib.

Figure 5. Electrochemical data showing that electroosmotic transport in the multiplexed distribution chip delivers reagents from the central well through the microchannels and to specific wells. The cyclic voltammetry of $Fe(CN)_6^{3-}$ (a) and $Ru(NH_3)_6^{3+}$ (b) are measured in wells 3 and 4, 70 s after initiating reagent distribution. 5 mM $Fe(CN)_6^{3-}$ and 5 mM $Ru(NH_3)_6^{3+}$ are added to the central well in the case of (a) and (b) respectively. A three-electrode bundle is placed in the wells 3 or 4, which are 1.5 mm in diameter. The working and counter electrodes are Au and Pt wires 0.25 mm in diameter. The reference electrode is Ag/AgCl made from 0.25 mm-diameter Ag wire.

discrete fluorescent plug in the separation channel (Figure 4c).

Electrochemical measurements provide a method to monitor the distribution of electroactive species through microchannels to specific reservoirs, as shown by the results in Figure 5. If 5 mM $K_3Fe(CN)_6$ solution is added to the central well 1 and electroosmotically transported toward well 3 for 70 s, a Faradaic current arising from of $Fe(CN)_6^{3-}$ reduction is subsequently observed in well 3 when it is probed electrochemically, no similar current being observed in well 4. Similarly, Figure 5b shows the electrochemical detection of $Ru(NH_3)_6Cl_3$ that was electroosmotically transported from well 1 to well 4. The current was measured in both cases using a cluster of three electrodes bound together in such a way that the cross section was about 1 mm.

Mixing of Reagents

For chips configured as shown in Figure 2, the reagent from the central well must be mixed and reacted with reagents from the side wells in the reaction channel prior to separation. To efficiently use expensive reagents such as enzymes, it is important to tightly control the ratio of the two reagents in the reaction channel. The ratio of mixing of reagents from two wells can be estimated by measuring the ratio of the widths of the two fluids in the channel at the mixing point using data from the CCD camera (Figure 6). As shown in Table 1, the ratio of the heights (h_1/h_2) of the intensity peaks for the two channels agrees with the ratio of the current passing through the channels. Therefore, by using the current as a measurement parameter, it is possible to control the

ratio of mixing of two reagents by adjusting the voltages applied to the driving electrodes.

Table 1. Relationship between applied voltages applied to three reservoirs, ratio of currents in the two corresponding channels, and the height ratio of the fluid in the channel measured from the CCD images. *

V_1, V_2, V_3 (Volts)**	I_a/I_b ***	h_2/h_1
0, 0, 1000	2.9	~4.5
50, 0, 1000	1.9	1.7
100, 0, 1000	1.2	0.9
150, 0, 1000	0.75	-
200, 0, 1000	0.42	0.46
250, 0, 1000	0.16	-
300, 0, 1000	0.05	~0

*1 mM fluorescein solution is added to well **1** and buffer is added to well **2**. The channel surfaces are positively charged.
V_1:V_2:V_3 is the the voltages of the electrodes in wells **1, **2**, and **3**.
***I_a, I_b are the current from well **1** to **3** and from well **2** to **3**, respectively.

CONCLUSIONS

We have shown that a plastic multiplexed distribution card can be fabricated and that uniform distribution of reagents can be achieved by tailoring the properties of the channel surfaces using polymers and/or the plasma treatment of the laminated films. Electroosmotic flow has been demonstrated on multichannel chips as a means to distribute and mix reagents without subjecting them directly to high electric fields. The mixing of reagents from two wells can be controlled by adjusting the current passing through the channels, which is achieved by tuning the voltages applied to the driving electrodes.

REFERENCES

1. R. M. McCormick, R. J. Nelson, M. G. Alonso-Amigo, D. J. Benvegnu, and H. H. Hooper, "Microchannel Electrophoretic Separations of DNA in Injection-Molded Plastic Substrates", *Anal. Chem.*, *69*, 2626 (1997).

2. C. S. Effenhauser, G. J. M. Bruin, and A. Paulus, "Integrated Chip-Based Capillary Electrophoresis", *Electrophoresis*, *18*, 2203 (1997).

3. T. D. Boone, H. H. Hooper, and D. S. Soane, "Integrated Chemical Analysis on Plastic Microfluidic Devices", *Technical Digest of the 1998 Solid-State Sensor and Actuator Workshop*, Transducers Research Foundation, Cleveland (1998); pp. 87-92.

4. M. A. Roberts, J. S. Rossier, P. Bercier, and H. Girault, "UV Laser Machined Polymer Substrates for the Development of Microdiagnostic Systems" *Anal. Chem.*, *69*, 2035 (1997).

5. A. Manz, J. C. Fettinger, E. Verpoorte, H. Lüdi, H. M. Widmer, D. J. Harrison, "Micromachining of Monocrystalline Silicon and Glass for Chemical Analysis Systems - A Look into Next Century's Technology or just a Fashionable Craze?" *Trends Anal. Chem.*, *10*, 144 (1991).

6. D. J. Harrison, K. Fluri, K. Seiler, Z. Fan, C. S. Effenhauser, and A. Manz, "Micromachining a Miniaturized Capillary Electrophoresis-Based Chemical Analysis System on a Chip", *Science*, *261*, 895 (1993).

CHAOTIC MIXING IN MICROFLUIDIC SYSTEMS

Mark A. Stremler, Michael G. Olsen, and Byung H. Jo
The Beckman Institute for Advanced Science and Technology
University of Illinois at Urbana–Champaign
Urbana, IL 61801

Ronald J. Adrian and Hassan Aref
Department of Theoretical and Applied Mechanics
University of Illinois at Urbana–Champaign
Urbana, IL 61801-2983

David J. Beebe
Department of Biomedical Engineering
University of Wisconsin at Madison
Madison, WI 53706-1608

ABSTRACT

Obtaining rapid mixing in microfluidic systems is a problem that must continue to be addressed if microelectromechanical systems are to attain their full potential in commercial markets. We present the paradigm of 'designing for chaos' as a general framework for enhancing mixing in microfluidic applications. This approach is based on a fundamental understanding of the kinematics underlying the mixing process, freeing the MEMS researcher to work with design guidelines instead of empirically determined physical configurations. We have applied this strategy in designing a passive in-line micromixer. Computational and experimental analyses demonstrate the effectiveness of the resulting design in generating chaos in the flow and hence enhancing mixing.

INTRODUCTION

The difficulty in mixing fluids on the microscale lies in the small size of the devices. Despite the small length scales involved, mixing solely by molecular diffusion is too slow for many applications, as the relevant time scales can also be extremely small. Flows at this scale are predominantly laminar, so the efficient mixing obtained in turbulent flows is not practically attainable. Furthermore, fabrication limitations often preclude the use of standard macroscale mixing techniques such as mechanical stirring.

Fortunately, mixing in such flows can be enhanced through 'chaotic advection' [1-3], in which passive fluid particles advected by a periodic, laminar velocity field exhibit chaotic trajectories. Relative to integrable (non-chaotic) advection, chaotic advection enhances stretching and folding of material interfaces. This deformation of fluid-fluid boundaries increases the interfacial area across which diffusion occurs, which increases the mean values of the gradients driving diffusion and leads to more rapid mixing.

A number of tools are available for detecting chaotic advection and measuring its extent, making it possible for a microscale mixer to be 'designed for chaos'. First, a general configuration is chosen based on design guidelines suggested by the theory of chaotic advection. For example, the active mixer of Evans, Liepmann and Pisano [4] is derived from the pulsed source-sink system investigated by Jones and Aref [5]. A mixer is then designed from the general configuration by taking into account the specific application and the fabrication constraints. The resulting device is modeled numerically, and diagnostics are computed to verify the occurrence of chaotic advection. Iterations can then be made on the design to

optimize the amount of chaos present. Finally, the device is tested experimentally to verify that mixing is enhanced. The use of this paradigm for designing a passive mixer is illustrated in the following sections.

DEVICE DESIGN

Passive mixing in a single microchannel is attractive for a number of reasons. First of all, a passive mixing scheme, which relies solely on a constant flow source, is generally more robust and easier to implement than an active mixing scheme, which relies on exerting some form of control over the flow field through such means as moving parts or varying pressure gradients. Secondly, a single channel maintains a relatively constant cross-section, which results in lower strain rates than if the flow is split into multiple smaller streams or forced through small orifices. Thus the existing work on chaotic advection is investigated to determine the guidelines for designing such a mixer.

It has been shown with a simple analytical model that flow in a 'twisted pipe' can contain significant amounts of chaotic advection [6]. The basic configuration consists of a sequence of pipe bends, with successive bends oriented along different planes. This result has been applied successfully to heat transfer enhancement on the macroscale [7-9], and the goal is to now implement this design on the microscale.

Previous studies of the 'twisted pipe' have only considered flow in circular pipes arranged in curved three-dimensional configurations. Fabrication limitations lead us to consider instead the related orthogonal geometry shown in figure 1, which can be constructed using a variety of techniques. The channels used in this study are fabricated of polydimethylsiloxane (PDMS) using a micro-molding technique [10]; a variation on the design in figure 1 has also been fabricated and tested in silicon [11].
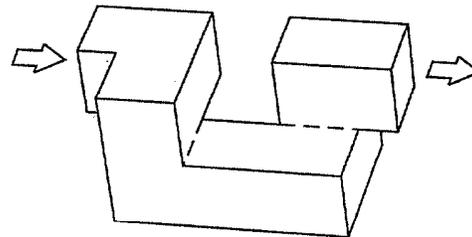

Figure 1. One segment of the '3D serpentine' mixer. The fabricated mixer consists of 11 segments placed end-to-end, and the channel cross-section is 300μm x 300μm.

An alternative configuration is the 'zig-zag' channel investigated by Branebjerg, Fabius, and Gravesen [12]. However, theoretical results for the 'twisted pipe' suggest that this channel design will generate little or no chaotic advection. We investigate the similar 'square-wave' geometry shown in figure 2 and compare its mixing capabilities with the 3D serpentine channel to illustrate the effectiveness of producing chaos in the flow.
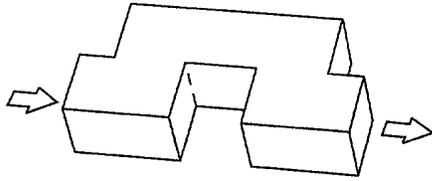

Figure 2. One segment of the 'square-wave' mixer. The fabricated mixer consists of 14 segments placed end-to-end, and the channel cross-section is again 300μm x 300μm.

NUMERICAL MODELING

While theory can suggest or predict the occurrence of chaotic advection, the amount of chaos present can only be determined for specific examples. A standard tool for detecting chaos in spatially or temporally periodic systems is the Poincaré section [3]. A schematic of how this section is generated for spatially periodic systems is shown in figure 3. The Poincaré section consists of patterns of points formed by passive particles advected through the periodic flow field. Chaotic advection is present in those portions of the section where the resulting pattern is random.
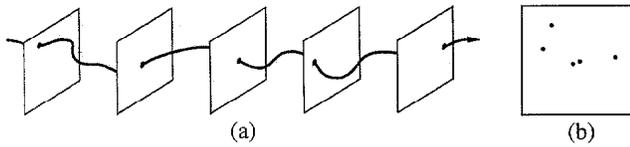
Figure 3. Generation of a Poincaré section in a spatially periodic system. (a) A passive particle is advected by the periodic velocity field and passes through a series of periodic planes in the system. For the mixers considered here the planes are located at the exit (or entrance) of each mixing segment. The position at which the particle passes through each plane is recorded. (b) All of the positions are transferred to a single plane, which is the Poincaré section. Regular patterns in the Poincaré section indicate integrable (non-chaotic) behavior, while random patterns indicate chaotic behavior.

In order to generate Poincaré sections in these channels, the velocity fields are first computed on a grid using CFD-ACE, a commercial finite-volume code from CFD Research Corporation. The channels are modeled as infinitely long by enforcing periodic boundary conditions at the inlet and exit of a single mixer segment. An important parameter in these flows is the Reynolds number, Re = UL/ν, which gives the relative importance of inertia and viscosity in the system in terms of the mean channel velocity U, the kinematic viscosity ν of the fluid, and the width L of the channel cross-section. For water-based flows in these channels, Re=10 corresponds to a flow rate

of approximately 0.2 mL/min. Velocity fields are computed at various Reynolds numbers by varying the specified pressure drop across one segment of the mixer. Particle traces and the corresponding Poincaré section are then generated by integrating numerically through the discrete velocity field using an adaptive Runga-Kutta scheme. The velocity at any point within the channel is determined by polynomial interpolation.
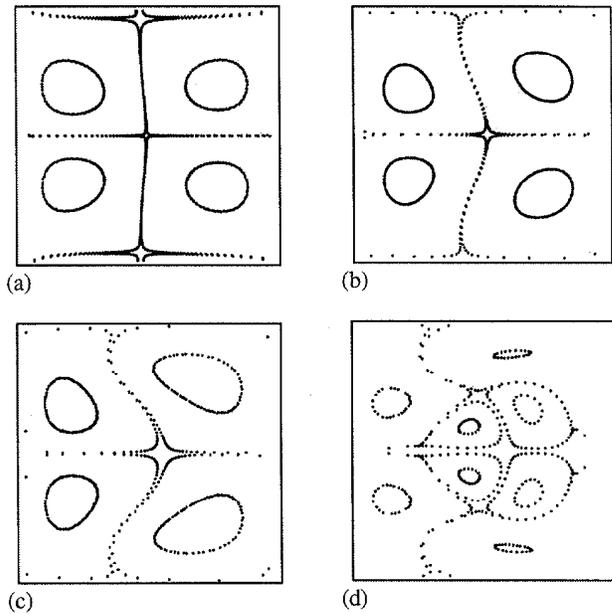

Figure 4. Poincaré maps in the square-wave mixer at (a) Re=10, (b) Re=20, (c) Re=30, and (d) Re=50. Each closed curve shown is generated by tracing a single particle. Those curves passing very close to the boundary are not shown as closed due to numerical error that occurs in following particles near the wall of the channel. While the complexity of the advection does increase with increasing Reynolds number, no chaos is present.
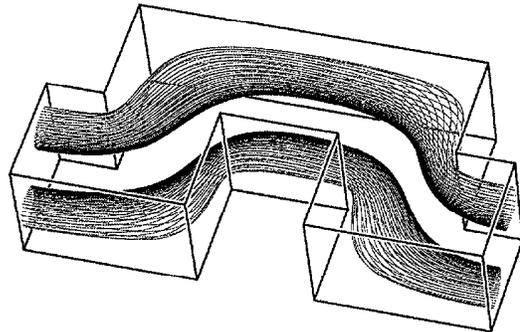

Figure 5. Particle traces in the square-wave mixer at a Reynolds number of 10. The lower tube corresponds to initial conditions lying on the lower left 'circle' in figure 4(a), and the upper tube to the upper right 'circle' in figure 4(a).

Poincaré sections in the square-wave channel for Reynolds numbers from 10 to 50 are shown in figure 4. In each case no chaotic advection appears, as is expected from the analytical results [6]. The reason a non-chaotic flow is expected to mix poorly is illustrated in figure 5. Closed curves in the Poincaré section correspond to 'stream tubes' in the flow that are barriers

across which fluid particles cannot move. Thus a mechanism such as diffusion is required to mix between the two tubes shown in figure 5.

Poincaré sections in the 3D serpentine channel for Reynolds numbers from 1 to 20 are shown in figure 6. In contrast to the square-wave mixer, chaotic advection does occur in the flow for the Reynolds numbers considered, with significant amounts of chaos present already at Re=10. Thus it is expected that the 3D serpentine mixer will perform significantly better than the square-wave mixer for moderate Reynolds numbers (roughly for Re≥10).

(a)

(b)

(c)

(d)
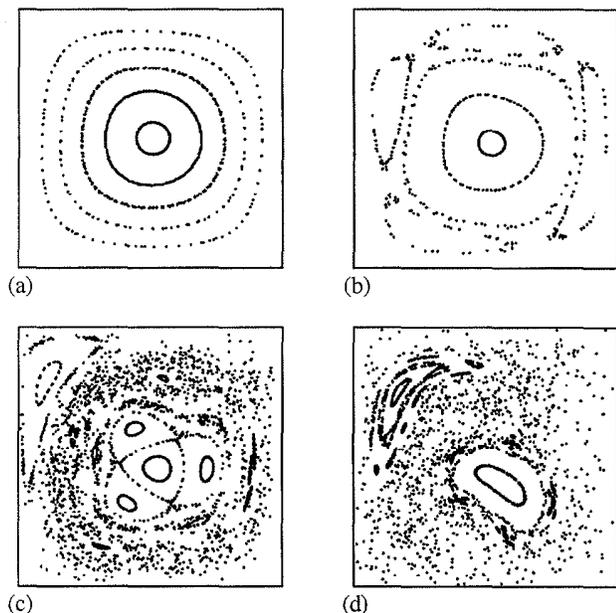
Figure 6. Poincaré maps in the 3D serpentine mixer at (a) Re=1, (b) Re=5, (c) Re=10, and (d) Re=20. At low Reynolds numbers the Poincaré sections are dominated by regular closed curves, and the advection is predominantly non-chaotic. However, for Re≥10 the Poincaré sections contain large regions of chaos. In (c) and (d) the particle patterns forming the chaotic regions are generated by following only 1 or 2 particles.

EXPERIMENTAL COMPARISONS

Mixing in the channels is evaluated experimentally by observing the color change of a pH indicator as it flows through the channel. Two separate fluid streams flow into each channel through a T-junction. One stream contains phenolphthalein dissolved in a mix of 50% ethyl alcohol and 50% de-ionized water with a concentration of 0.016 M. The second stream contains 98.3% sodium hydroxide pellets also dissolved in a 50/50 mix of alcohol and water, giving a pH of approximately 13.

The phenolphthalein and sodium hydroxide streams are introduced into the mixer from two water columns. After the two colorless streams come in contact, both sodium hydroxide and phenolphthalein begin to diffuse. The reaction time of the phenolphthalein is negligible, although the ratio of reacted to unreacted phenolphthalein at any point within the device depends on the local pH value in a non-trivial way [13]. Thus, the amount of reacted phenolphthalein produced depends on the

diffusivities and initial concentrations of both phenolphthalein and sodium hydroxide.

(a)

(b)
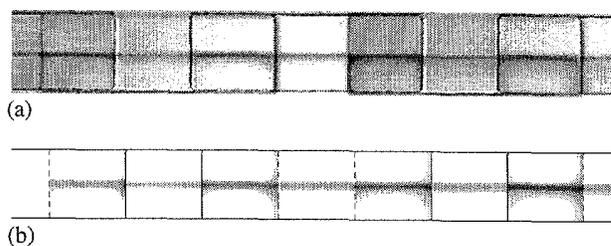
Figure 7. (a) Experimental results and (b) computational results for diffusion in the first two segments of the square-wave mixer at a Reynolds number of approximately 10. Flow is from left to right. The view is from the front of the mixer as oriented in figure 2. At the mixer inlet the top half of the channel contains the NaOH stream and the bottom half the phenolphthalein stream. The reacted phenolphthalein appears as a dark product.


(a)

(b)
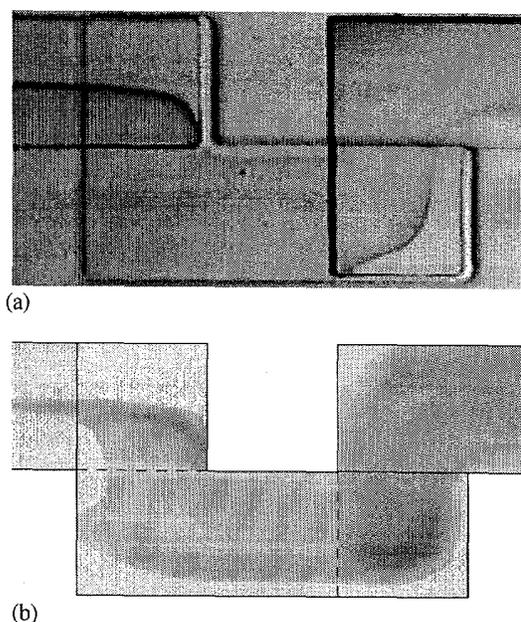
Figure 8. (a) Experimental results and (b) computational results for diffusion in the first segment of the 3D serpentine mixer at a Reynolds number of approximately 10. Flow is from left to right. The view is from the front of the mixer as oriented in figure 1.

The mixing ability of the device is determined qualitatively by observing the color change of the phenolphthalein during the mixing process. Images of the reacted phenolphthalein are captured through an Olympus BX60 microscope at 4X magnification with a Sony 8-bit CCD camera. The microscope objective used has a depth-of-focus of 175 μm. Although by this definition the entire channel depth is not in perfect focus, blurring of features in the flow is negligible. Illumination comes from a halogen light source mounted behind the mixer. Since the halogen lamp produces incoherent light, the light waves scattered by two molecules are assumed to have additive intensities when the images of the molecules overlap. Thus the

intensity of red observed in the channel is proportional to the amount of reacted phenolphthalein present.

As shown in figures 7 and 8, there is good qualitative agreement between the experimental results and a simple computational diffusion model, which demonstrates the validity of the numerical model and lends credibility to the numerical diagnosis of chaos presented in figures 4 and 6.


(a)

(b)
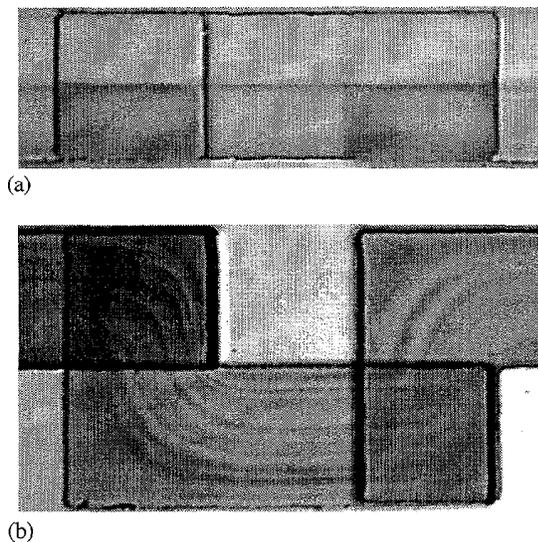
Figure 9. Comparison of the mixing results at the last mixing segment in (a) the square-wave mixer and (b) the 3D serpentine mixer. The fluid has traveled approximately the same distance in each mixer at the segments shown. Stretching and folding of the material interface is apparent in (b).

The mixing capabilities of these two channels are compared experimentally in figure 9. In the square-wave mixer product is observed only in the bottom half of the channel, and much of the product is concentrated at the centerline. In contrast, the 3D serpentine mixer produces more product, and this product is distributed throughout the channel cross-section. Thus, designing for chaos has resulted in a mixer that performs better than one that does not produce chaotic advection.

CONCLUSIONS

The paradigm of designing for chaos in microfluidic systems gives the MEMS researcher the freedom to work within a flexible design framework instead of relying solely on the use of empirically determined devices. The validity of this approach has been demonstrated in the design of a passive, in-line micromixer. The resulting device enhances mixing significantly relative to a similar in-line mixer that does not produce chaotic advection in the flow. This approach of 'designing for chaos' is not limited to passive in-line devices. Indeed, there are a variety of mixers that can be developed using this framework.

ACKNOWLEDGEMENTS

REFERENCES

1. H. Aref, "Stirring by chaotic advection", *J. Fluid Mech.*, *143*, 1-21 (1984).

2. H. Aref, "Chaotic advection of fluid particles", *Phil. Trans. R. Soc. Lond.* A, *333*, 273-288 (1990).

3. J. M. Ottino, *The Kinematics of Mixing: Stretching, Chaos, and Transport*, Cambridge University Press (1989).

4. J. Evans, D. Liepmann, and A. P. Pisano, "Planar laminar mixer", *Proc. IEEE Workshop Micro Electro Mech. Sys. (MEMS '97)*, Nagoya, Japan (1997), pp. 96-101.

5. S. W. Jones and H. Aref, "Chaotic advection in pulsed source-sink systems", *Phys. Fluids, 31*, 469-485 (1988).

6. S. W. Jones, O. M. Thomas, and H. Aref, "Chaotic advection by laminar flow in a twisted pipe", *J. Fluid Mech.*, *209*, 335-357 (1989).

7. N. Acharya, M. Sen, and H.-C. Chang, "Heat transfer enhancement in coiled tubes by chaotic mixing", *Int. J. Heat Mass Trans.*, *35*, 2475-2489 (1992).

8. D. Sawyers, M. Sen, and H.-C. Chang, "Effect of chaotic interfacial stretching on bimolecular chemical reaction in helical-coil reactors", *Chem. Eng. J.*, *64*, 129-139 (1996).

9. A. Mokrani, C. Castelain, and H. Peerhossaini, "The effect of chaotic advection on heat transfer", *Int. J. Heat Mass Trans.*, *40*, 3089-3104 (1997).

10. B. H. Jo, L. VanLergerghe, K. Motsegood, and D. J. Beebe, "Fabrication of three-dimensional microfluidic systems by stacking molded PDMS layers", *Proc. SPIE Symp. Micromachining Microfabrication*, Santa Clara, CA (1999), pp. 222-229.

11. R. H. Liu, M. A. Stremler, K. V. Sharp, M. G. Olsen, J. G. Santiago, R. J. Adrian, H. Aref and D. J. Beebe, "Passive mixing in a three-dimensional serpentine microchannel", *J. Microelectomechanical Sys.*, (to appear).

12. J. Branebjerg, B. Fabius, and P. Gravesen, "Application of miniature analyzers: from microfluidic components to μTAS", *Proc. Micro Total Anal. Sys. Workshop (μTAS '94)*, Enschede, The Netherlands (1994), pp. 141-151.

13. S. Zhang, S. P. Schneider, and S. H. Collicott, "Quantitative molecular-mixing measurments using digital processing of adsorption images", *Exp. in Fluids, 19*, 319-327 (1995).

A NOVEL DESIGN ON A CD DISC FOR 2-POINT CALIBRATION MEASUREMENT

Marc J. Madou[1], Yumin Lu[1], Siyi Lai[2], Yije Juang[2], L. James Lee[2], Sylvia Daunert[3]

[1]Department of Materials Science and Engineering, The Ohio State University, Columbus, OH 43210
[2]Department of Chemical Engineering, The Ohio State University, Columbus, OH 43210
[3]Department of Chemistry, University of Kentucky, Lexington, KY 50506

ABSTRACT

Today the Si/polymeric medical diagnostic sensors that are on the market only feature a 1-point calibration [1]. Such a measurement results in less accurate sensing and more in-factory sensor rejection. The 2-point calibration fluidic method introduced here will alleviate some of the shortcomings of such current miniature analytical systems. Our fluidic platform is a multi-purpose micro analytical laboratory on a compact disc (CD) [2, 3]. This system is based on the centrifugal force, in which fluidic flow can be controlled by the spinning rate of the CD and thus a whole range of fluidic functions including valving, mixing, metering, splitting, and separation can be implemented. Furthermore, optical detection such as absorption and fluorescence can be incorporated into the CD control unit to obtain signals from pre-specified positions on the disc.

Keywords: microfluidics, centrifuge, SU-8, photolithography.

INTRODUCITON

The lab CD platform technology we are developing features fluidic structures micromachined into a plastic CD. In the manufacture, we rely on traditional mechanical machining and on lithography based techniques. The centrifugal force due to the rotation of the CD provides the pumping force for the release and flow of reagents and analytes. No external pumps and valves are needed to control the fluids. We used computer numerical control (CNC) milling to mechanically manufacture the 2-point calibration system on a CD. Although CNC machining may not achieve satisfactory absolute tolerance, its accessibility and rapid prototyping capability enables us to conduct various investigations concerning microfluidic dynamics faster.
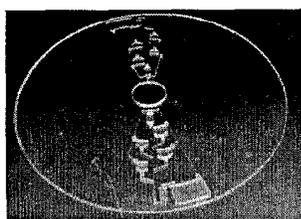
Figure 1. (a) Two 2-point calibration units on a CD
(b) Functional diagram of a single 2-point calibration unit.
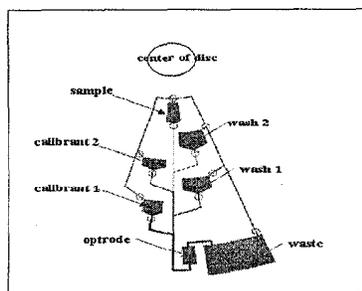
For smaller fluidic features (e.g. channels with $\phi < 80$ μm) and mass production, we are making the CD with a combination of photolithography and replication tools. In this approach, a thick layer of SU-8 (SU-8-100, MicroChem, Inc.) is spin-coated on a metal covered plastic CD, followed by photolithography to pattern the microfluidic structures. After development, nickel is electroplated to make a metal-insert mold which can be used for hot embossing or injection molding.

In our 2-point calibration design (Fig 1), the liquid flows in the order of calibrant 1, wash1, calibrant 2, wash 2, and sample to the optrode chamber by increasing the rotation speed. The same optrode chamber is used for measuring calibrants and sample to eliminate artificial system errors common with devices using separate chambers for measuring sample and calibrants.

For optical detection purposes, chromoionophores are incorporated into small chambers close to the rim of the CD. To demonstrate the functionality of the 2-point calibration system, we deposited a potassium ion selective membrane in the optrode chamber using a micro-delivery set-up (Nano-plotter, from GeSiM) [4]. The membrane is composed of Poly(vinyl chloride) (PVC), valinomycin (ionophore), dioctyl sebacate (plasticizer), potassium tetrakis(4-chlorophenyl) borate and chromoionophore for detection at 640-nm wavelength. In order to identify the proper polymeric material with the lowest protein adsorption and minimal optical interference with the optical measurement, the protein adsorption characteristics and optical absorption properties of seven candidate polymers for the CD platform were evaluated.

THEORY

Fluid Propulsion

In order to sequentially valve fluids through a monotonic increase of rotational rate, we designed a structure with progressively higher "burst" frequencies. The burst frequency is the frequency at which the fluidics are released from their reservoirs.
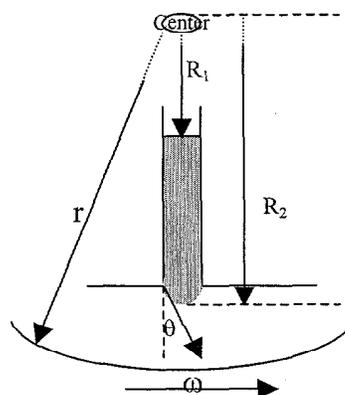

Figure 2. Schematic illustration of fluid propulsion in centrifugal microfluidics.

In the LabCd platform, the centrifugal force provides the pumping force, while the capillary force at junctions inhibits flow. This constitutes a RPM-dependent capillary "valve" controlling the release and flow of fluids. The two opposing forces can be described as follows (see Figure 2).

The pumping force (P_c) due to centrifugal force is given by:

$$\frac{dP_c}{dr} = \rho\omega^2 r \qquad (1)$$

Here ρ is the density of the liquid, ω is the angular velocity of the CD platform, r is the distance between a liquid element and the center of the CD,

Integration of equation (1) from $r=R_1$ to $r=R_2$ gives:

$$\Delta P_c = \rho\omega^2 (R_2 - R_1)\frac{R_2 + R_1}{2} = \rho\omega^2 \Delta R \cdot \overline{R} \qquad (2)$$

The capillary force (P_s) due to interface tension is given by:

$$\Delta P_s = \frac{\gamma \cos\theta \cdot C}{A} \qquad (3)$$

Here, γ is the interface tension, θ is the contact angle, A is the cross-section area of the capillary and C is the associated contact line length.

The liquid will be released when ΔP_c is greater than ΔP_s, consequently the burst frequency calculated from equations 1 and 3 is given by:

$$f_b \geq (\frac{\gamma \cos\theta \cdot C}{\pi^2 \rho \cdot \overline{R} \cdot \Delta R \cdot 4A})^{\frac{1}{2}} = (\frac{\gamma \cos\theta}{\pi^2 \rho \cdot \overline{R} \cdot \Delta R \cdot d_H})^{\frac{1}{2}} \qquad (4)$$

Here d_H is so called hydrodynamic diameter of the channel.

2-point calibration

By appropriately choosing the channel diameter and reservoir dimensions and locations, we designed a 2-point calibration system, which is capable of 5 sequential fluid movement steps. The liquid flows in the order of calibrant 1, wash1, calibrant 2, wash 2, and sample to the optrode chamber by increasing the rotating speed.

Table 1 shows the 2-point calibration design parameters for the CNC machined CD. Figure 3 shows the calculated burst frequencies of the reservoirs in the 2-point calibration platform.

A typical 2-point calibration curve is shown in Figure 4. The slope (sensitivity) is determined by the measurements of calibrant 1 and calibrant 2. The ion concentration of the unknown sample is deduced from this calibration curve.

Table 1: Parameters of CNC machined CD platform

| Reservoir	Thickness	Volume	Burst RPM	Observed bust RPM
Sample	800μm	10.47μL	1222	1126
Wash 2	800μm	18.61μL	1032	1026
Calibrant 2	800μm	9.56μL	832	822
Wash 1	800μm	18.03μL	623	622
Calibrant 1	800μm	11.79μL	427	524
Optrode	500μm	6.81μL	N/A	N/A
Waste	800μm	84.35μL	N/A	N/A
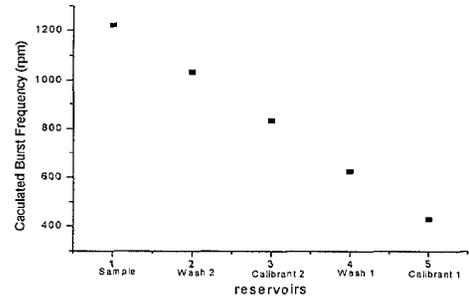
*Figure 3. A plot of calculated burst frequencies.*
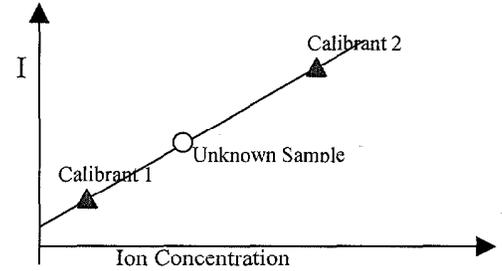
Figure 4. A typical 2-point calibration curve.

EXPERIMENTAL

Computer Numerical Control (CNC) machined CD

Numerical control machining is a method of automatically operating a machine. It has been used in industry for over 40 years [5]. CNC machining features an added on-board computer, which enables programmed control of the machine. The typical cutting tools used in CNC machining are: milling, drills, multifacet drills, boring, tapping, angle milling, etc [5]. Two CNC machined 2-point calibration units in a polycarbonate disc are showed in Figure 1.

Although CNC machining has advantages such as easy accessibility and rapid prototyping capability, it has several limitations in this application. For example, CNC milling can hardly fabricate samples with feature size less than 80 microns or make smooth turns. Another concern is the aspect ratio (the ratio of depth/width). The aspect ratio achieved by CNC machining is typically not high (0.5 in our case).

Considering those limitations of CNC machining, we are exploring alternative manufacturing methodologies. Photolithography with deep-UV resist followed by electroplating and molding is our preferred method for LabCD platform fabrication.

SU-8 BASED CD

Thick photoresist based lithography is widely used in micromachining. It has several advantages, for example high aspect ratio features are available (over 20 reported) [6] and very small feature sizes can be obtained [7, 8]. Furthermore, it is much easier to fabricate any desired pattern compared to CNC machining. For example, the shape of the reservoir can be easily changed from square to circular or drop-like shape with lithography.

We used SU-8, which is a negative resin-based photoresist, in this application. Firstly a thick layer of SU-8 was coated on a metal-covered plastic CD. Metal (Ti/Au) is used to protect the CD's surface from the SU-8 developer. To decrease the thermal stress in the SU-8, the sample was soft-baked at 70°C for 1 hour and then at 95°C for 3 hours. The Exposure step was performed on a PLA-501 FA Canon mask aligner. The post-bake was carried out at 60°C for 20 minutes and then at 90°C for 30 minutes. Again, this was to minimize the thermal stress.

After development in PMGA SU-8 developer (MicroChem, Inc.), microfluidic structures such as channels and reservoirs shown in Figure 5 resulted. The width of the channel in Figure 5b is 50 μm and the height is 200 μm (aspect ratio = 4).

The mechanical and optical properties of SU-8 structures need further investigation. However, microfluidic components fabricated in SU-8 may in some instances be directly used in platform evaluation.

(a)

(b)
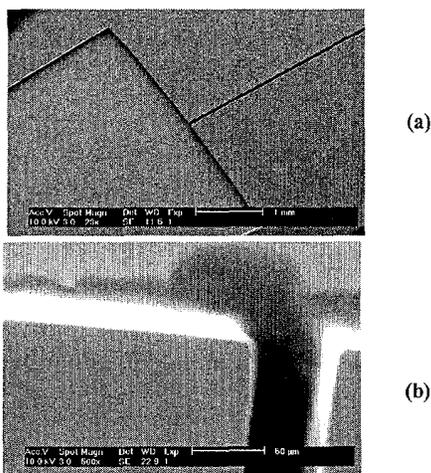
Figure 5(a) (b). SEM of the microchannel and reservoir manufactured by photolithography in SU-8.

PLASTIC REPLICATION

Firstly, an SU-8 structure is fabricated by photolithography as described in the previous section. The resulting structure features the same pattern as the desired final product. Electroplating is then carried out to replicate a metal mold (typically nickel) from the SU-8 structure. Such mold can be used in hot embossing or injection molding for mass production of the CD platform [9].

Polycarbonate (PC), Optical Quality Polycarbonate (OQPC), Polyurethane (PU), Polydimethyl Siloxane (PDMS), Polystyrene (PS), Polyvinyl butyral (PVB), and Polymethylmethacrylate (PMMA) were used in this application.

In the fabrication of regular PC, OQPC, PMMA and PS discs, the hot embossing process was applied and the experimental setup is shown in Figure 6. The mold was first brought in contact with the polymer substrate. Both of them were heated up above the glass transition temperature (Tg), followed by hot embossing process for a few seconds. After the mold and polymer substrate were cooled down below Tg, the mold was then separated from the polymer substrate with the microstructures fully transferred into it. For rubber-like polymers such as PU and PVB, same procedures were applied except the heating-temperature was 105°C. High temperature was adopted in order to prevent the rubbery material from springing back after embossing. As to PDMS, a casting method was employed, which is shown in Figure 7. A holding ring was made to hold the resin and the mold. After the resin is poured

into the holding ring, it cured slowly and became rubbery PDMS with pattern replicated onto it, which can be easily peeled off from the mold.
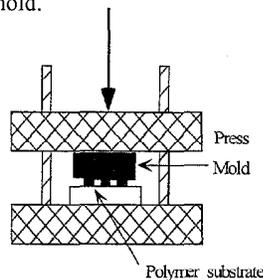

Figure 6. Schematic of hot embossing.
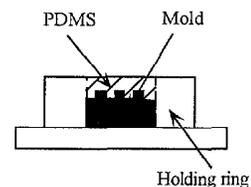
Figure 7. Schematic of casting.

Optical Sensor incorporated CD

We deposited a potassium ion selective membrane in the optrode chamber for the measurement of the ion concentration in the calibrantas and unknown sample. The deposition is performed by a GeSiM Micro pipetting system (Nano-plotter). Table 2 represents the composition of the membrane.

When the testing solution contacts the membrane, the potassium ion is extracted into the membrane and bound with the valinomycin. To ensure the electrical neutrality within the membrane, the chromoionophore loses a hydrogen ion which goes into the buffered sample solution. The loss of the hydrogen ion causes a change in the optical properties of the chromoionophore (absorption, fluorescence, etc.), which is what we measure. The magnitude of this change depends on the ion concentration of the testing solution. Therefore, we can obtain the ion concentration by measuring the change in the optical properties of the chromoionophore.

Table 2. K^+ selective membrane composition

Dioctyl sebacate	65 mg
PVC	33 mg
valinomycin	2 mg
Chromoionophore (ETH5350)	50 mole% relative to the ionophore
Potassium tetrakis(4-chlorophenyl)borate	50 mole% relative to the ionophore

RESULTS AND DISCUSSIONS

Protein adsorption characteristics of polymer materials for CD platform

Since the proteins might be used in the CD platform, protein adsorption characteristics of the CD substrate were studied to identify a proper polymeric material for this application. Seven polymeric materials were evaluated. They are: Polycarbonate, OQPC, Polyurethane, PDMS, Polystyrene, PVB, and PMMA.

It was found that PC and PMMA have comparable protein adsorption characteristics to that of PS. PVB was not suitable for this application since it is too sensitive to protein and as a consequence, its optical properties are greatly affected upon the exposure to the protein.

Burst frequencies of the 2-point calibration platform

The CNC machined polycarbonate disc was covered with another PC plate and tested for the burst frequencies of the reservoirs. The effect of the cover material's properties (e.g. stiffness, interface tension, etc) on the burst frequency was

observed. The CD covering is an important issue when testing the burst frequency because the covering material may affect the contact angle and the interface tension between the solution and substrate. When testing the burst frequency of different chambers, two methods of CD covering were tried: soft cover (cover the CD with the Scotch tape) and hard cover (cover the CD with a polycarbonate plate). Since the Scotch tape is flexible, it can be bulged up when pressure is applied, higher burst frequencies than the calculated values were expected and the experimental results confirmed it.

Table 1 shows the test results on a polycarbonate covered CD. Considering the tolerance of CNC machining and the measuring deviation, the results are quite consistent with the calculated frequencies.

CONCLUSION

A 2-point calibration measurement platform based on the centrifugal pumping force and RPM-dependant "valves" are developed. Three fabrication methodologies: CNC machining, SU-8 photolithography and polymer replication are discussed and contrasted. The burst frequencies of the CNC machined CD were tested and compared to the calculated values. Optical sensor is incorporated into this platform for detection of ion concentration. The protein adsorption characteristics of seven polymeric materials were evaluated. By using the micromachining technology, the fluidic structure's size is decreased to micron scale. Thus the sample volume utilized in the LabCD platform is reduced to micro/naro liter. Multi analysis and high throughput screening can be achieved by different designs in the LabCD.

REFERENCES

1. Jacobs, E., Vadasdi, E., Sarkozi, L., Coman, N. " Analytical Evaluation of i-STAT portable Clinical Analyzer and Use by Nonlaboratory Health-Care professionals". *Clin. Chem.* 1993, 39(6), 1069.
2. Duffy, D. C., Gills, H. L., Lin, Joe, Sheppard, N. F., kellogg, G. "Microfabricated Centrifugal Microfluidic Systems: Characterization and Multiple Enzymatic Assays" *J. Anal. Chem.* 1999, 71, 4669.
2. Madou, M. J., Kellogg, G. J. "The LabCD^{TM}: A centrifuge-based microfluidic platform for diagnostics". *SPIE.* San Jose, CA (1998).
3. GeSiM mbH, Bautzener Landstrasse 45, D-01454 Grosserkmannsdorf Germany.
5. Valentino, J. V., Goldenberg J. "Introduction to computer numerical control" Prentice Hall, 1993.
6. Report on NANO^{TM} SU-8 resists for MEMS/microsystem application by MicroChem Corp. Nov. 1999.
7. K. Y. Lee, N. Labianca and S.A. Rishton " Micromaching Application of a High Resolution Ultrathick Photoresist" *J. of Vacuum Science and Technology B*, 13(6): 3012-3016 (1995).
8. H. Lorenz, M. Despont, N. Fahni, N. Labiaca, " SU-8: A low-cost negative resist for MEMS" *SUSS report*, 10:1-3, Third/Fourth Quarter (1996).
9. H. Schift, R. W. Jaszewski, C. David, and J. Gobrecht, "Microelectronic Engineering", 46, 121 (1999).

Molecular Tagging Velocimetry for Microfluid Applications

Zilan Shen, Hyoun Park, Jia Ming Chen, Yongchi Tian, and Ann-Marie Lanzillotto
Sarnoff Corporation, 201 Washington Road,
Princeton, NJ 08540

ABSTRACT

We are developing a high-resolution, molecular tagging velocimetry technique that will accelerate the design and optimization of micro-scale devices such as a microtissue array based biosensor. The complex aerodynamic and hydrodynamic flows around and within the devices can be analyzed with capabilities beyond that of traditional measurement techniques. This non-invasive measurement technique will provide highly detailed information on the entire flow field, such as 3-D velocity, 3-D vorticity, 3-D pressure fields, shear stress and flow rate distribution patterns within micro-scaled devices.

INTRODUCTION

Microfluidic technology exploits sample size reduction and parallel processing for sensor, micro-pumping and actuation applications. These applications require precise fluidic delivery and management that can only be achieved with a correct knowledge of the flow field within individual channels and valves [1]. Further more, some of the applications will also benefit from accurate measurement of the pressure and force field related to the flow media [2]. For example, viscous damping and pressure field information are crucial in evaluating the squeeze film damping effects on the dynamic responses of MEMS devices, such as resonance frequency and the Q-number of an actuator [3].

In the past, traditional techniques have not provided detailed flow information on the micro scale. Point-by-point measurement methods, including laser Doppler velocimetry (LDV) and hot wire techniques, are limited to steady or periodic flows. When used for micro-scale application, hot-film techniques are highly invasive and can create substantial flow disturbances. Similarly, the typical sample volume size for LDV has been on the order of hundreds of microns, making it very difficult to obtain near-wall measurement due to the backscatter of the laser signal from boundaries. Furthermore, these techniques are not suitable for very low velocity flow measurements. The hot-wire sensor measures the heat transfer from the environment, and therefore suffers from the "free convection" effect. Similarly, since LDV measures the Doppler frequency shift due to the movement of the scattering center (particle), more complicated heterodyne detection is needed to differentiate a very small frequency change, on the order of ten to hundred Hz, from the basic laser frequency of 10^{14} Hz.

Full-field methods, such as particle tracking or image velocimetry (PTV & PIV), have had only limited success at characterizing 3-D flows. These techniques use particles that are easily trapped within microfluidic devices. These particle techniques also suffer from heavy computational demands and are prone to errors related to correlating particles from one image to another for complex flows with substantial velocity through the plane of illumination. Furthermore, due to inertial effects, particles are often ejected out of high vorticity regions where information cannot be obtained. Additionally, objects or obstacles within or near regions of interest within the flow field can also deplete the area of the seeded particles, making accurate flow measurements difficult.

This non-invasive, high-resolution, molecular tagging velocimetry technique that we are developing obtains direct velocity field measurements of complex aerodynamic and hydrodynamic flows. The technique uses fluorescent or colored molecules as flow tracers, does not have particle seeding requirements, and is well suited to capture the full flow field instantaneously, in three dimensions. Velocity gradients are obtained from the detailed velocity field measurements. From this gradient field, the pressure field of the entire region of interest can be derived from the Navier-Stokes equation using a local modeling technique.

METHOD AND THEORY

The molecular tagging technique uses a laser to tag fluid elements by changing the electronic states of the tagged molecules. These tagged molecules are then tracked in time. There are two different molecular tagging methods: laser-induced fluorescence/phosphorescence [4] and the photochromic [5] method. When laser-induced fluorescence/phosphorescence molecules are used, the tagging trace is visualized by photon emissions from the trace itself. These emissions are induced by the tagging laser or by the secondary illumination. For the photochromic method, the tagging trace is visualized by the selective absorption of the backlit illumination. These traces are typically observed as a change in color, usually involving a conversion from a colorless to a colored form (e.g., from clear to dark blue). In both cases, a tagging laser is used either to excite the tagging molecule to a long lifetime, higher energy state or to induce a photolysis reaction that activates (or "uncages") the tagging molecules to a light sensitive dye [6]. Typically, the liquid working concentrations are on the order of ten parts per million (ppm), with the excited molecule having a half-life ranging from microseconds to tens of seconds.

After the fluid is tagged, the displacement of the tagging molecules is typically captured by a CCD. For photochromic and caged molecule methods, another short-duration light source is used to capture the photochromic image or to illuminate the released active dye for fluorescent emission. However, past studies typically used a xenon flash lamp to capture images of the colored tagging molecules. This light source, with its wide band of wavelengths, has limited the optical response range of the photochromic tagging molecules to be mainly in visible. To overcome this limitation, we have replaced the flash lamp with a tunable, pulsed laser source (with a wavelength range of 200 – 1000 nm). The monochromatic and tunable output from the laser allows us to consider molecules that were previously not visible and are only observable in the UV or infrared regions.

In order to analyze a flow field, typically, laser pulses are redirected by beam splitters and lens arrays to tag fluid molecules so that sets of fine lines can be observed within the region of interest. By comparing two consecutive images of these trace

patterns, at known points in time, the displacement field can be measured and used to estimate the instantaneous velocity field, $\boldsymbol{u}$. Using a local interpolation technique [7], field gradients, vorticity, pressure fields, and all flow variables throughout the field can be determined, with an accuracy of < 5% of the peak velocities. The vorticity field, ω, is defined as the curl of the velocity,

$$\omega = \nabla \times \boldsymbol{u}. \tag{1}$$

The wall shear stress, τ_w, can be determined everywhere within the region of interest as functions of the dynamic viscosity of the fluid, μ, and the gradient of the tangential velocity along the normal direction of the wall, n,

$$\tau_w = \mu \frac{\partial u_l}{\partial n}\bigg|_{wall}. \tag{2}$$

Furthermore, the first and second velocity gradient fields will be used to estimate the pressure field, P. This will be accomplished by solving the Navier-Stokes equations for steady, incompressible, 3-D flow,

$$\nabla \cdot \rho \boldsymbol{u} = 0 \tag{3a}$$
$$\partial \rho \boldsymbol{u}/\partial t + (\boldsymbol{u} \cdot \nabla) \rho \boldsymbol{u} = -\nabla P + \mu \cdot \nabla^2 \boldsymbol{u} + \rho g, \tag{3b}$$

and integrating the resulting pressure gradient field. Here, ρ is the fluid density and g is the gravitational constant. Equation (3a) is the mass continuity condition, and equation (3b) states that the force in the flow direction equals to momentum change of the flow plus the viscous dissipation and the gravitational force.

EXPERIMENTAL DETAILS
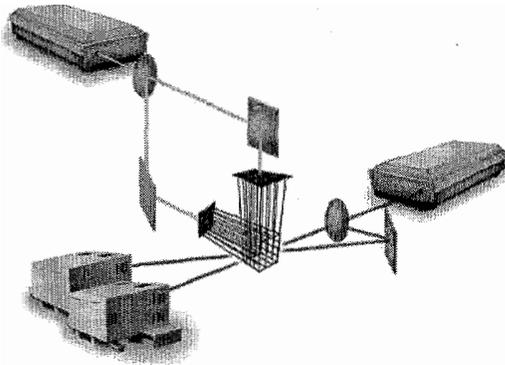
Figure 1. Experimental Set-up for 2-D Molecular Tagging Velocimetry. General Method: 1. Tagging laser generates a short UV pulse which is split and focused into region of interest. 2. Along each line of the resulting grid, UV-sensitive indicators are instantly activated. 3. A second laser, tuned specifically for the activated species, is used to capture the grid pattern at set time intervals. 4. Two views of the 2-D grid are acquired by placing a camera at two positions to stereoscopically capture full 3-D movement. 5. Every aspect of the flow field can be obtained including: instantaneous 3-D velocity, vorticity, shear stress and pressure fields.

The experimental setup for achieving 2D measurement is shown schematically in Figure 1. A Nd:Yag laser is used to provide the UV tagging pulses. The maximum pulse energy is 1000 mJ/pulse, and the pulse width is 10 ± 2 ns. The beam diameter is about 1 cm with operating wavelengths of 266nm and 355nm. The output pulse is split into two beams propagating along two orthogonal directions by a beam splitter, and redirected to the

focusing lens arrays by a pair of dielectric coated mirrors. The laser beams focused by these lens arrays will intersect within the region of interest to produce a two-dimensional 10 x 10 grid pattern with 100 nodes of intersection. The second laser is a tunable laser light source with a wavelength range of 200-1000 nm. This laser serves as a strobe, and can be tuned to the absorption peak of the excited tagging molecules to capture optimal photochromic change, or to the excitation wavelength of "uncaged" dyes to produce fluorescence emission. A high resolution, 1300 x 1300 pixel CCD is used to capture the motion of the grid pattern. A computer and two digital delay generators are used to synchronize the components of the experimental system, including the lasers and the cameras. The computer is also used to record the raw image data from the CCD camera, process the images, and analyze the data to obtain relevant flow field information.

Two photochromic fluid solutions are tested here. The first fluid contains 50 ppm of photochromic dye (1,3,3,-trimethyl-6-nitro-indoline-2-spiro-2-2-benzopyrane) dissolved in deodorized kerosene (Shell-Sol 715), the second is an aqueous fluid that contains 100 ppm of a water-based, light sensitive dye.

Two test devices are reported here. A simple, 2 mm x 1 cm x 15 cm, rectangular device is used to demonstrate a 2D grid. A blank microtissue array biosensor is used to demonstrate an actual microfluid application. The biosensor was developed in collaboration with researchers at MIT. It is made of UV transparent acrylic and houses a stainless steel scaffold patterned with an array of 4 x 10 microchannels. Each microchannel has a 300 μm square cross-section and a 500 μm pitch. A schematic of the biosensor is shown in Figure 2. There is one flow inlet and one flow outlet above the scaffold, and an additional flow outlet beneath the scaffold. The inlet flow enters the biosensor and flows into the region above the scaffold. Part of the inlet flow will continue laterally and exit the biosensor from the top outlet, and the remainder of the flow will pass through microchannels and exit the biosensor via the bottom outlet.

| Top View	Side View	Bottom View

Scaffold   **Microchannels**
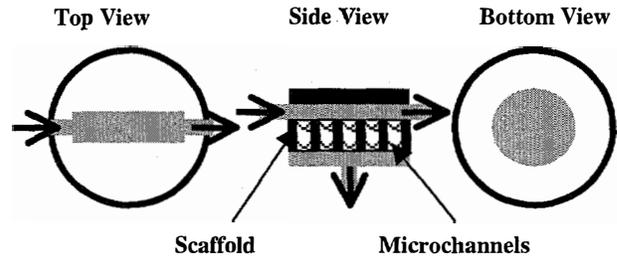
Figure 2 Schematics of a biosensor. The figure shows top, side and bottom views of the device. The arrows indicate inlet and outlets. The top inlet and outlet are horizontal, and the bottom outlet is vertical. The patterned area in the center of the side view is the scaffold, where lighter regions represent microchannels. The top chamber above the scaffold is 2 mm wide, 6 mm long, and 3 mm deep. The diameter of the bottom chamber is 10 mm, and the depth is about 6 mm.

As is shown in the side view, the heart of the biosensor is an array of capillary bed-sized units of "tissue beds" made by self-assembly of dissociated cells from native tissue within the confines of small through-channels created in a silicon or stainless chip by deep-trench etching. Direct measurements of the flow fields through each channel will allow us to design the biosensor for optimal performance.

The measurement will indicate if there is preferential or unequal flow distribution along the microchannels, and if there are any "dead zones", where the flow is stagnant and trapped by the local flow geometry. Figure 3 shows how the experiment is conducted to measure the flow rate of each microchannels. A laser line trace is placed right below the scaffold, which holds the tissue array microchannels. The position of the trace at a given delay time, t_1-t_0, is captured by a CCD.


(a)

(b) **(c)**
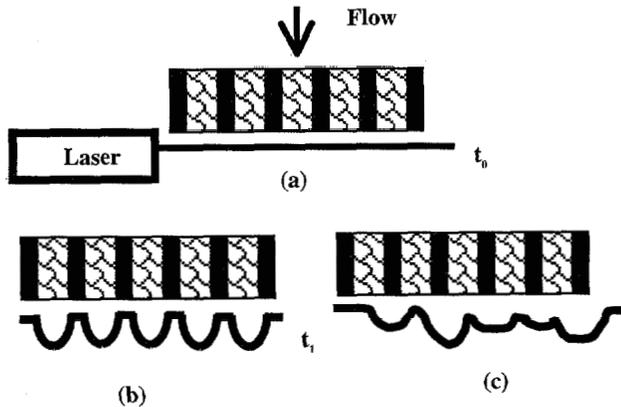
Figure 3 *Schematic demonstration of the flow measurement. (a) A laser line trace is placed directly below the tissue array under a series of microchannels, and is captured at time t_0 immediately after the tagging pulse. (b) The 2nd trace is captured at a later time t_1 showing displacement from the initial trace. (c) A case where the flow distribution is unequal.*

RESULTS AND DISCUSSIONS

The kerosene based tagging solution was first tested in a 2 mm x 1 cm x 15 cm rectangular device. A 2-D photochromic grid was generated and is shown in Figure 4. The line-traces were approximately 100 μm wide, and inter-trace space spacing is 1 mm. The flow is from the top-right to the bottom-left, the Reynolds number is less than 1.
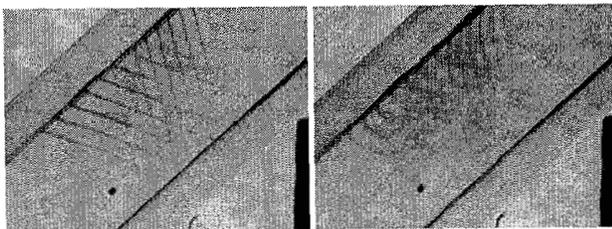
Figure 4. Two-dimensional photochromic grid pattern in a 1 cm x 2 mm rectangular cell. (a) immediately after tagging laser pulse and (b) 1s afterwards. Flow is from top-right to bottom-left. The spacing of the line-traces is 1 mm, and the line-trace width is 0.1 mm. The working fluid is kerosene with 0.005% UV-sensitive dye.

A similar result is obtained using a water-based fluid. This is of particular interest since a lot of biological applications require the use of biocompatible, aqueous fluids. In Figure 5, the optical response of the water-based solution is demonstrated by an array of parallel, 100 μm wide traces with a 1 mm inter-trace spacing.
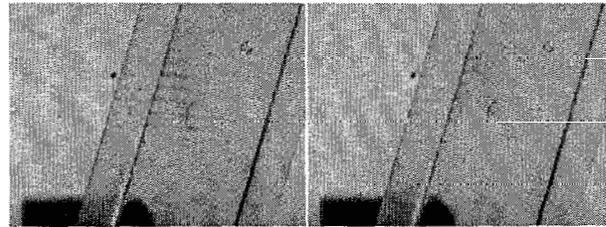
Figure 5. Photochromic traces in a 1 cm x 2 mm rectangular cell. (a) immediately after tagging laser pulse and (b) 1s afterwards. Flow is from bottom-left to top-right. The spacing of the line-traces is 1 mm, and the line-trace width is 0.1 mm. The working fluid contains a UV sensitive, water-based dye.

In our most recent work, an array of photochromic line-traces were generated directly beneath the microchannels of the scaffold with a model of the biosensor. The working fluid is kerosene with 0.005% UV-sensitive dye. A CCD image of the side view of the biosensor is shown in Figure 6(a). The flexible tubing at the top of the picture is the inlet, and the L-shaped channel located at the bottom is the bottom outlet of the biosensor. Another top outlet is on the opposite side of the device, and therefore cannot be seen. The dark thin plate across the image is the stainless steel scaffold with the microchannel array. The scaffold also divides the biosensor into a 2 mm x 6 mm top chamber, and a 10 mm x 10 mm bottom chamber. The flow is from top to bottom, and the Reynolds number is approximately 0.3. The displaced tagging traces 5 seconds after the tagging pulse is shown in Figure 6(b). Again, the initial spacing of the line-traces is 1 mm, and the line-trace width is approximately 0.1 mm.
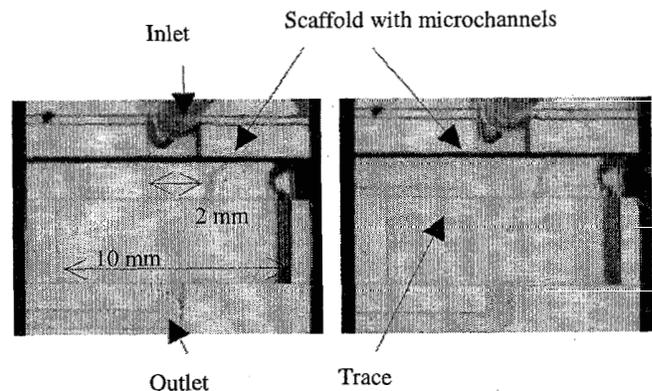
Figure 6. Photochromic traces in a microtissue array biosensor. (a) side view of the testing device. The top inlet feeds the tagging fluid into a 2 mm wide, 6 mm long top chamber; while the bottom outlet is connected to a 10 mm x 10 mm bottom chamber. The two chambers are separated by a stainless steel scaffold with an array of 4 x 10 microchannels. A top outlet is on the other side of the device and can not be seen from this image. (b) image of photochromic traces taken 5 seconds after the tagging pulse. The Reynolds number is approximately 0.3. The initial spacing of the line-traces is 1 mm, and the line-trace width is 0.1 mm. The working fluid is kerosene with 0.005% UV-sensitive dye.

After digital subtraction, the image data from Figure 6 is shown in Figure 7. Note that there are several interesting features. The separate flows out of the four microchannels can be clearly

seen in this image. Due to an effectively no slip boundary condition at the wall of the top chamber, the flow rate out of the two microchannels on either side of the two center channels is slightly less than the flow rate out of the two center channels. Due to viscous action, the flow profile takes on a more familiar parabola-like shape as the flow travels away from the microchannels. Note that the curved profile becomes wider as the flow gets farther away, showing that the non-zero velocity flow profile expands as it enters the larger bottom chamber. The non-zero flow profile converges as it approaches the narrow outlet at the bottom, while the maximum flow speed at the center profile increases substantially. In addition, note that the trace lines on the side barely move, indicating regions that are "dead zones" where elements in the flow media can be trapped due to extremely long residence times. These dead zones can alter the outcome of the biosensor measurement and compromise it's performance and reliability, as the trapped material may lead to contamination or false readings.
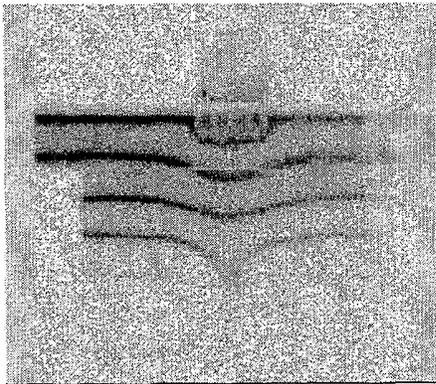
Figure 7. A digitally subtracted image of the photochromic traces in a microtissue array biosensor. There are 4 obvious trace lines shown here. The center of top trace displays the flow profile as it exits the 4 microchannels. The 2nd and 3rd trace shows that the profile is expanded wider and converted to a parabolic shape due to the viscous action. The bottom trace shows an increase of the maximum flow speed at it exits through the narrow outlet.

Only 2D tagging result were reported here. However, with minor additions to the experimental configuration of the measurement system, instantaneous 3-D measurements will be obtained during future work. The 3-D grid will be generated to cover volumes of interest with side dimensions ranging from 5mm to several centimeters. The traces can be produced with a thickness below 50 microns for grid patterns. The inter-trace spacing is often 500 μm to 1 mm or less, with an effective resolution that approaches tens of microns using local modeling and image analysis techniques. For single trace applications the trace thickness can approach several microns. The ultimate resolution of the system will also be enhanced by state-of-the art image processing techniques including edge detection and active contour methods. In terms of temporal resolution, the initial and final images are taken from microseconds to seconds apart. By stereoscopically capturing the displacement of this grid with two CCD cameras, or one camera placed at two different locations, the 3-D velocity field can be obtained.

CONCLUSIONS

Currently, we are in the process of testing various water-soluble and air-based molecular tagging candidates and are developing the flow analysis techniques for vector interpolation and local modeling. Such detailed information is valuable in any effort to develop applications where the fluid flow characteristic have a vital role in effective operation, such as microchip-based biosensor, genetic analysis instruments, micro-pumping and actuation. It may also lead to many product innovations in the fields of drug inhalers, artificial heart valves, and coronary and vascular bypass grafts.

ACKNOWLEDGMENTS

REFERENCES

1. P. Gravesen, J. Branebjerg, and O. Sondergard, "Microfluidics – A Review", *Journal of Micromechanics, 3,* 168 (1993).

2. M. T. A. Saif and N. C. MacDonald, "Measurement of Forces and Spring Constants of Mocroinstruments", *Review of Scientific Instruments, 69,* 1410 (1998)

3. Y. Cho, A. P. Pisano, and R. T. Howe, "Viscous Damping Model for Laterally Oscillationg Microstructures", *Journal of Microelectromechanical Systems, 3,* 81 (1994)

4. B. Stier and M. M. Koochesfahani, "Molecular Tagging Velocimetry Measurements in Gas Phase Flows", *Experiments in Fluids, 26,* 297 (1999)

5. R. E. Falco and C. C. Chu, "Measurement of Two-dimensional Fluid Dynamic Quantities Using a Photochromic Grid Tracing Technique", *SPIE, 814, Photomechanics and Speckle Metrology,* 706, 1987

6. J. E. Guilker, K. R. Gee, P. A. McMurtry, J. C. Klewicki, "Use of Caged Fluorescent Dyes For The Study of Turbulent Passive Scalar Mixing", *Experiments in Fluids, 21,* 237 (1996)

7. H. Park, J. A. Moore, O. Trass, M. Ojha "Laser photochromic velocimetry estimation of the vorticity and pressure field, two-dimensional flow in a curved vessel" *Experiments in Fluids 26,* 55 (1999)

ACTIVE MICRONEEDLES WITH INTEGRATED FUNCTIONALITY

John Brazzle, Dan Bartholomeusz, Rupert Davies, and Joseph Andrade
[1]Department of Bioengineering, University of Utah
Salt Lake City, Utah 84112

Richard A. Van Wagenen[1,2]
[2]Protein Solutions
Salt Lake City, Utah 84112

A. Bruno Frazier[1,3]
[3]School of ECE, Georgia Institute of Technology
Atlanta, Georgia 30332

ABSTRACT

The focus of this work is the design, fabrication, and characterization of a new class of biomedical micro systems, the active microneedle. Active microneedle systems represent an advancement over current microneedle technologies through the integration of additional functionality (e.g. biochemical sensing, mechanical, etc.). The active microneedles described in this paper include the following additional functionality: integration of multiple inlet/output ports, multiple lumens (flow channels), and bioluminescence based biosensors for monitoring metabolic levels (e.g. creatine and glucose).

INTRODUCTION

Both hollow and solid microneedles and microneedle arrays have been demonstrated by a number of research groups over the past few years. Initially, only microneedles made of solid silicon had been realized [1]. Najafi and Wise developed these early microneedles that were used as neural microprobes. Recently, solid silicon microneedles fabricated by Henry et. al. were formed using reactive ion etching [2]. The resulting planar needle array consisted of 20×20 needles for increasing the skin permeability to drugs. Lin, Pisano, and Muller were first to present hollow silicon processed microneedles [3,4]. These microneedles were fabricated on silicon substrate and used silicon nitride to fully enclose the channels. Later, Chen and Wise [5] presented hollow silicon microneedles for neural recording and drug delivery. These two approaches utilize a sacrificial processing technique that consumes the substrate on the devices are built. In contrast, Talbot and Pisano [6] have incorporated a polysilicon micromolding process to develop hollow microneedles. These hollow microneedles are fabricated by micromolding polysilicon (polymolding) in silicon micromolds.

Previously, we reported micromachined needles and needle arrays of hollow metallic needles for minimally invasive drug delivery and biofluid extraction [7,8]. These microneedles have the capability to deliver drugs subcutaneously while while minimizing the pain inflicted to patients. Recent advancements in the design and fabrication of hollow metallic microneedles has led to improved fluid handling capabilities, as well as, increased functionality. Recently, McAllister et al. [9] have reported hollow NiFe microneedle arrays.

The paper presents the design, fabrication, and characterization of a hollow metallic active microneedle. Figure 1 is a schematic representation of the described microneedle with an integrated biosensor. The active microneedle includes design features such as tapered needle tips; multiple output ports on the back and front of each needle; multiple lumens and multiple input ports, Figure 2, and bioluminescence based biosensors for monitoring metabolic levels.

FABRICATION AND PACKAGING

The micromachined multiple output port needles are fabricated using extensions of previously reported micromachining technologies [10,11,12]. The fabrication process is low temperature and is compatible with integrated circuit (IC) technology as a post process. Initially, a metal system of adhesion layers and an electroplating seed layer are electron beam evaporated onto a silicon wafer. Using standard thick photoresist micromolding techniques, this metal layer is patterned and 20μm of palladium is electroplated to form the solid layer or bottom wall of the microneedle [13]. The use of palladium as a structural material provides high mechanical strength and durability, as well as, biocompatibility for use in biomedical applications [14,15,16]. Next, 20μm of commercially available thick photoresist, AZ4620, is deposited and patterned into sacrificial structures. These sacrificial structures serve to precisely define the inner dimension of each microneedle. 800Å of gold is sputter deposited onto the sacrificial photoresist structures to act as an electroplating seed layer for top shell electroplating. Next, a 20μm layer of palladium is electroplated to form the top and side walls (hollow layer) of the microneedle. The sacrificial thick photoresist is then removed
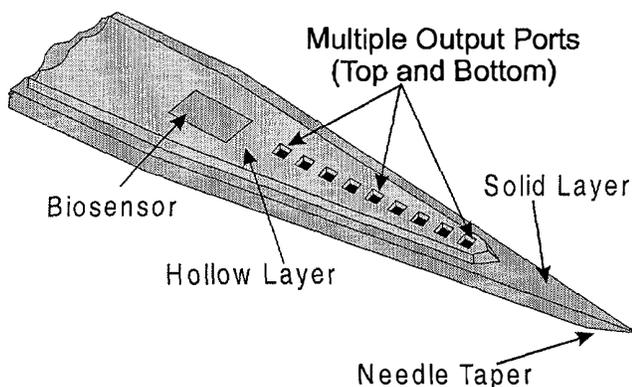

Figure 1. Schematic representation of a micromachined multiple output port needle with an integrated biosensor.
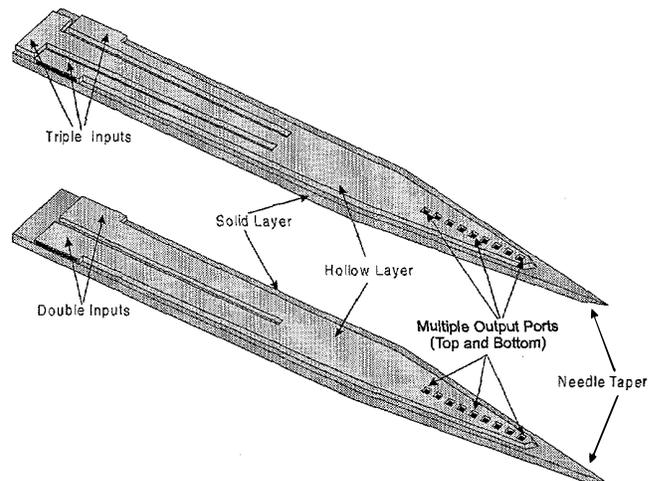

Figure 2. Schematic representation of multi-lumen design for the active microneedles.

using a sequential rinse in acetone, isopropanol, and de-ionized water, producing the hollow metallic microneedles. In the final step, the microneedles are released from the silicon substrate by first removing the chromium layer to expose the copper seed layer. Then the copper seed layer is etched away from underneath the needles by placing the wafer in a solution of ammonium hydroxide saturated with cupric sulfate. The solid silicon substrate is not consumed in this process and may be re-used. Using the given process offers the flexibility to integrate many additional functions directly onto the microneedle. Additionally, the basic process allows for great latitude in the routing and branching of the fluid

BIOLUMINESCENCE

Bioluminescence-based analysis is 100 to 1000 times more sensitive than conventional chromogenic (absorbance) measurements and is accurate over a five or more, orders of magnitude concentration range [17]. Firefly bioluminescence occurs by enzyme-catalyzed oxidation of luciferin utilizing adenosine triphosphate (ATP) [18]. Bacteria bioluminescence is closely coupled to nicotinamide adenine dinucleotide hydride (NADH). Since most of biochemistry depends on ATP and/or NADH, nearly all metabolic reactions can be monitored by bioluminescence via one or more enzymes catalyzed and linked reactions [17,18]. During the production or consumption of a metabolite of interest, enzymes linked reactions will cause the production or consumption of ATP (or NADH). Detectable light is then produced via the following reaction (for ATP):

$$ATP + O_2 + D\text{-Luciferin} + Mg^{++} \xrightarrow{\text{Luciferase}}$$
Oxyluciferin + AMP + CO_2 + Ppi + Light (560 nm)
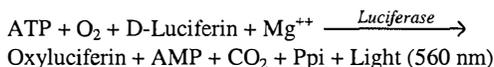
The change in light intensity will be stoichiometrically proportional to the time changing concentration of ATP (and thus proportional to the metabolite of interest). A photodiode or charge-coupled device (CCD) can be used for detection. Depending on the instrument used to detect the luminescence, nanomolar, picomolar, and even femtomolar analyses are possible [17]. With increased sensitivity, smaller amounts of sample fluid are needed for accurate analysis. 0.05-μL sample size that can be drawn by micromachined needles can painlessly access small amounts of body fluid that can be analyzed via bioluminescence.

The transmittance of light through a media across length λ is defined by Beer's law, $T = 10^{-Absorbance}$, where $Absorbance = \lceil_\lambda C_B(\lambda-x)$ and $\lceil_\lambda$ is the molar absorption coefficient in $dm^3/molEcm$ and C_B is the concentration in $mole/dm^3$ [19]. If the light is generated within a homogeneous media, the transmitting light is integrated along the length it travels and is then defined as:

$$T_{Total(\lambda)} = \int_0^\lambda 10^{-\varepsilon_\lambda C_B(\lambda-x)} dx = \frac{(1-10^{-\lambda})}{Ln10}.$$

where $\lceil_\lambda C_B$ is assumed to be 1 for simplicity. For a pyramid (which is the shape of the etched wells in the <100> Si wafers) with a peak of length λ, the transmittance is integrated through the viewing area and is equal to $T_{Total(\lambda)}Ea^2/3$. The assumption is that only the light being transmitted toward the viewing window is integrated and light traveling away from the viewing window is absorbed by the substrate. (Light a 560 nm is actually absorbed by crystallized Si). Thus, linear correlation of a volume verses integrated luminosity plot, for a given viewing window, should indicate that transmittance is proportional to signal intensity (averaged CCD counts over the viewing area). A flat bottom well with a reflective surface will have the transmittance of $T_{Total(2\lambda)}$, and will be about 2 time larger than $T_{Total(\lambda)}$ for $\lambda<<$1-cm (0.005 – 0.025-cm in this experiment). Thus, the slope of a linear fit of volume verses integrated luminosity plot, for a reflective surface of a given viewing window, should be about 2 times that of the slope for a non-reflective surface due to the increased transmittance.

RESULTS

Fabrication

The needle arrays fabricated to date include those electroformed from low-stress nickel sulphamate, gold cyanide, and palladium electroplating solutions [20,21]. Individual hollow metallic micromachined needles with multiple output ports fabricated to date have an inner cross-sectional area of 140×21 μm^2 (W×H) with outer dimensions of 200×60 μm^2. A micromachined needle fabricated on silicon and released from its substrate is shown in Figure 3. The needle shaft dimensions are 200 μm wide and 60 μm thick while the tip dimensions are less than 15 x 15 μm^2. The length of the tapered portion of the needle shaft is 1.0 mm and the distance from the tip to the first output port is approximately 300μm. The total length of the microneedle is 6.0 mm with input port inner dimensions of 140 μm wide and 21 μm high. The wall thickness of the needle is approximately 20 μm and the microneedle output ports are on the top and bottom with dimensions of 30 μm^2. The output ports are separated by 30μm and there are nine ports on top and 12 ports on the bottom. Bioluminescent detection of the biosensor output is performed at the output ports of the microneedle. Biosensing reagents are drawn into and dried down onto the inner walls of the hollow microneedle.
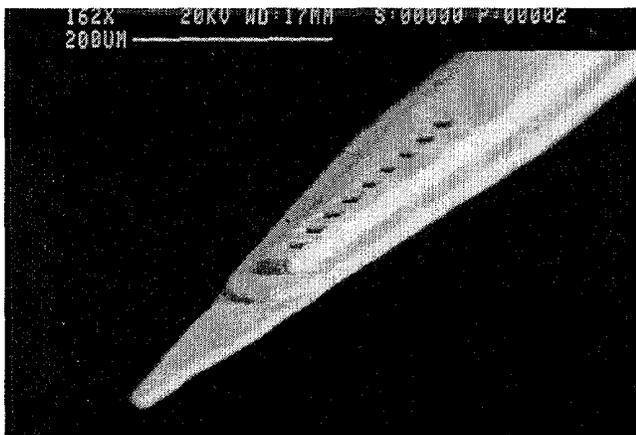

Figure 3. SEM of a multi-lumen, multiple output port micro needle. Shaft dimensions are 200 μm wide and 60 μm thick. Tip dimensions are less than 15 μm X 15 μm. Output ports are 30 μm^2.
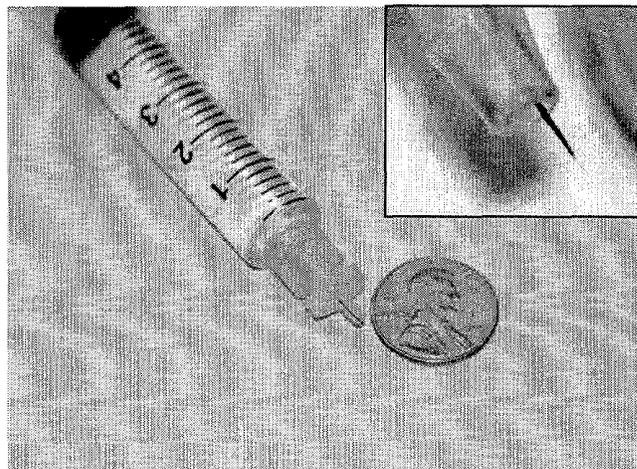

*Figure 4. Photograph of packaged microneedle connected to standard 5cc syringe as compared to a penny. **Insert:** Close-up of packaged microneedle.*

The completed micromachined needle structures were packaged into a standard luer-lock fitting using a polymeric medical-grade UV-curable adhesive (3321, Loctite Corp.). This interface between the microneedle and syringe consists of a simple female luer-to-1/16" barb adapter manufactured by Western Analytical (part# p-857x). Figure 4 is a photograph of a packaged microneedle connected to a standard 5cc syringe as compared to a penny. The insert of Figure 4 shows a close-up of the packaged microneedle that has been secured with UV epoxy on the barbed end of the interface. The UV curable epoxy was found to permanently affix the microneedle to the interface while providing a leak resistant seal. Fluid flow experiments were performed on five packaged microneedles at pressures ranging from 1psi to 70psi [9].

Bioluminescence Studies

In order to determine the feasibility and the physical limitations of using bioluminescence as a means for highly sensitive analyte measurement of small sample volumes, μ-reaction chambers (μRCs) were fabricated on silicon wafers using KOH anisotropic wet etching. An ATP firefly luciferase/luciferin solution was placed in the μRCs and observed through a close up lens with a CCD. The integrated CCD signal was recorded and compared with well size and depth. The attenuation of the CCD signal was also observed for wafers coated with titanium (500Å) followed by chromium (1500 Å).

A 5-mL firefly luciferase/luciferin solution consisted of 1.25-mg/mL bovine serum albumin (Sigma - reconstituted into the solution. Used for coating the glass vial to prevent denaturing of the luciferase), 1.25-mM ethylene diaminetetra acetic acid (Sigma), 12.5-mM Mg^{++} (Sigma – from $MgSO_4$), 1.84-μM firefly luciferase (Promega), and 1.25-mM luciferin (Biosynth) in a 1.25mM glycyl-glycine buffer. This mixture was able to maintain 90% activity for about 20-hrs when stored in the dark. A 5-mM ATP, glycyl-glycine buffer solution was also prepared. 20-μL of the firefly luciferase/luciferin solution was pipetted into 5-μL of the ATP solution that resulted in a 1.0-mM ATP mixture, saturated
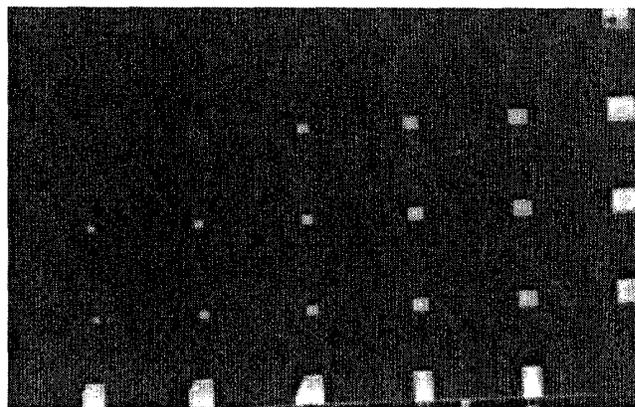
Figure 5. 20-sec integration of bioluminescence for the wafer that was etched 242.1-μm and coated with Ti/Cr. The μRCs seen here are the 750, 500, 400, 300, 250, and 200-μm wide squares.

μRC Width (μm)	P value from T-test between Intensity/Volume Slopes of Non-Reflective μRCs and Reflective μRCs	Ratio of Intensity/Volume Slopes for Reflective and Non-Reflective Substrates
750	0.00001	3.02
500	0.002	2.37
400	0.004	4.94
300	0.005	5.11
250	0.36	2.87
200	0.030	1.07
150	0.72	0.88
100	0.55	0.21

Table 1. Statistical comparison of the intensity/volume slope relationships for reflective vs. non-reflective surfaces of the same viewing window size.
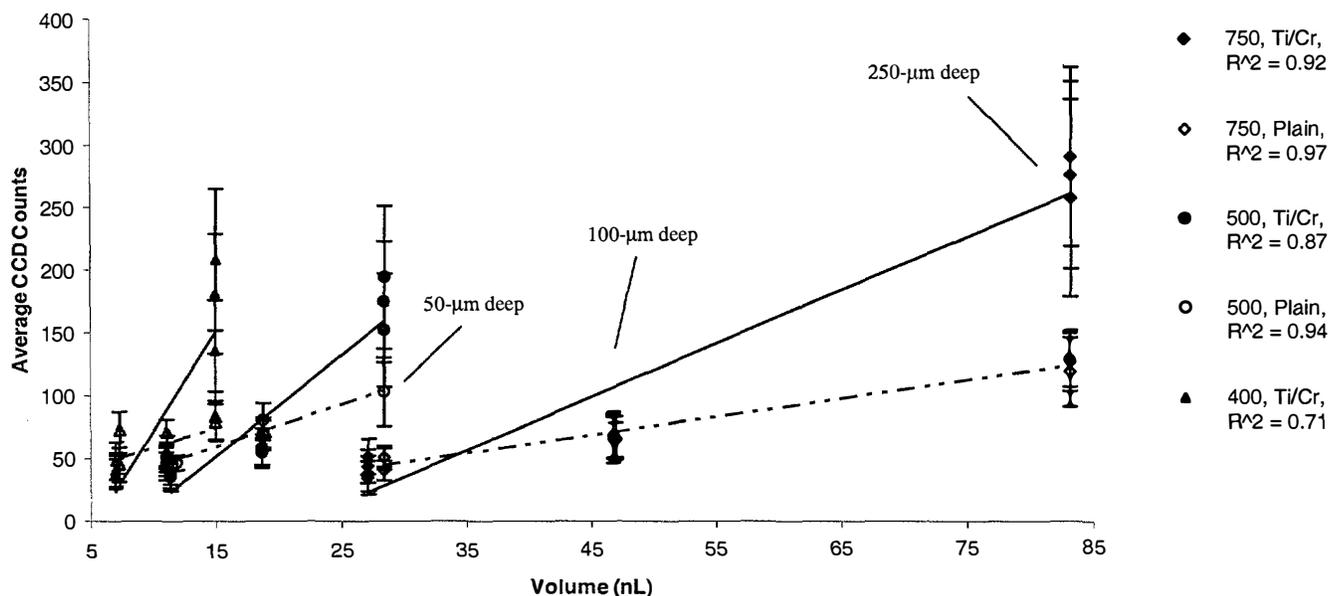
Figure 6. Average CCD counts from the 20-sec integrated CCD reading for different sample volumes. Data was plotted in sets for the same viewing area (or μRC square width). Data was also separated according for Ti/Cr reflective substrates and plain, non-reflective, substrates. For each set of data (wells with same width and coating), the increasing volume occurs from the increased etch depths 250-μm, 100-μm and 50-μm. Data from the different etch depths are pointed out for the 750-μm wide wells as an example. The higher intensity values occur for the deeper wells. The intensity/volume slope is greater for Ti/Cr reflective substrates than for plain substrates.

with luciferase and luciferin (which means the reaction rate was at it's peak). After the solution mixed, 20-µL of it was pipetted on an area about 20×15-mm wide. A thin glass cover slip was placed on the solution, starting at one end and tilting the cover slip as it was laid down, so that the excess bioluminescent fluid would disperse. For high concentrations of ATP (> 8-µM), the initial mixing of ATP with the luciferase/luciferin solution causes a peak luminescence within 3 seconds. The luminescence then tapers and levels off after 1 minute. Therefore, the light measurements for this experiment was integrated for 20 seconds, 2 minutes after the ATP and luciferase/luciferin solution were mixed, with the light intensity essentially constant. An ST6-A CCD camera, by Santa Barbara Instruments Group was fitted with an Olympus wide-angle lens and close-up ring. The experimental substrates were focused 55-mm below the lens with the aperture set at 2.8. The field of view was about 20×15-mm. The camera was operated at –20.00°C. 1.0-minute after the luciferase/luciferin solution was added to the ATP, a 20-sec dark field exposure was taken with the CCD. 2-min 5-sec after the ATP and luciferase/luciferin solution mixed, the CCD integrated a 20-sec exposure of the bioluminescent reaction while shrouded in a darkroom. The resulting image was saved as a TIF file (Range: 0 – 400). Scion Image (based on the NIH image software) was used to determine the average and standard deviation of the pixel values for each µRC.

Figure 5 shows an image of the 20-second integration of bioluminescence recorded by the CCD camera. One pixel of the CCD image was equal to 46.875-µm. Actual results failed to show light emitting from the 75, 50, 25, and 10-µm wide wells on all substrates. Figure 6 shows the average CCD counts from the 20-sec integrated CCD reading for different sample volumes. Table 1 statistically compares the intensity/volume slope relationships for reflective vs. non-reflective surfaces of the same viewing window size. The reflective wells with intensity/volume slopes that were statistically different from the intensity/volume slopes for the non-reflective surfaces ($P<0.05$ using the T-test) were over two times greater. This follows the hypothesis that transmittance is nearly doubled for chromium (reflectivity=0.67 @ 560 nm) coated versus uncoated silicon (reflectivity=0.33 @ 560 nm) µRCs.

The intensity/volume slope also increases as the µRC square width decreases. This is ideal for using long, narrow microneedles as the µRC. Table 2 statistically compares the intensity/volume slopes between reflective wells of different viewing widow sizes. It shows that there is a significance increase in the intensity/volumes slopes, as the µRC width decreases. The increase in intensity/volume is only significant down to wells that are 400-µm wide. For smaller µRC widths (150 – 300-µm), the non-reflective substrates show little correlation between average CCD counts and sample volumes ($R^2 < 0.25$). The Ti/Cr reflective substrates still show some correlation ($R^2>0.25$ for µRCs 250-µm wide and wider) between intensity and volume. However, as the µRC square width decreases (250-µm wide and smaller), there is little difference in signal intensity values for different etch depths.

µRC Square Widths Being Compared (µm)	P value from T-test between Intensity/Volume Slopes µRCs being compared	Ratio of Intensity/Volume Slopes for Smaller µRC Widths Verses Wider µRC Widths
500/750	0.004	1.89
400/500	0.04	1.98
300/400	0.28	1.59
250/300	0.89	1.11
200/250	0.81	1.28
150/200	0.91	0.73

Table 2. *Statistical comparison of the intensity/volume slopes between reflective wells of different viewing widow sizes. The intensity/volume slope ratio is the slope for the smaller window over the slope for the next larger window.*

The error in the signals for wells 200-µm wide and smaller is large enough to overlap with the background intensity. This suggests that 250-µm that smallest well window that would produce a discernable intensity signal. However, this does not mean that narrower, deeper µRCs like microneedles could not produce a good signal because their longer depths are able to hold a larger sample volume, which would produce a stronger signal.

CONCLUSION

Both silicon and metal electroformed micro needle technologies have been previously developed. This paper provides a basis for the next technological step in micro needle technologies through the development of active micro needles. Active micro needles include the integration of additional functionality into the basic micro needle structure. Additional functionality such as multiple lumens, multiple input and output ports, and the integration of biosensing capabilities has been shown.

REFERENCES

1. K. Najafi and K.D. Wise, "An implantable multielectrode array with on-chip signal processing," *IEEE J. Solid-State Circuits*, SC-21, pp. 1035-44, Dec. 1986.
2. S. Henry, D.V. McAllister, M.G. Allen, and M.R. Prausnitz, "Micromachined needles for the transdermal delivery of drugs", in *IEEE MEMS Conference*, Heidelberg, Germany, Jan. 25-29, 1998.
3. L. Lin, A.P. Pisano, R.S. Muller, "Silicon processed microneedles," in *Transducers 93', Int. Conf. Solid-State Sensors&Actuators*, pp.237-40, 1993.
4. L. Lin and A.P. Pisano, "Silicon-processed microneedles," in *IEEE J. Microelectromech. Sys.*, vol. 8, no. 1, pp. 78-84, 1999.
5. J. Chen and K.D. Wise, "A multichannel neural probe for selective chemical delivery at the cellular level," in *IEEE Solid-State Sensor & Actuator Workshop*, Hilton Head, SC., June 13-16, pp. 256-259, 1994.
6. N.H. Talbot and A.P. Pisano, "Polymolding: two wafer polysilicon micromolding of closed-flow passages for microneedles and microfluidic devices," in *Tech. Dig. Solid-State Sensor and Actuator Workshop*, Hilton Head, SC., June 8-11, pp. 265-268, 1998.
7. J. Brazzle, I. Papautsky, and A.B. Frazier, "Fluid-coupled metallic microfabricated needle arrays," in *Proc. SPIE Micro Fluidic Devices and Systems*, Santa Clara, CA, Sep. 21-24, pp. 116-124, 1998.
8. J. Brazzle, S. Mohanty, and A.B. Frazier, "Hollow Metallic Micromachined Needles with Multiple Output Ports", *SPIE Micro Fluidic Devices and Systems*, Santa Clara, CA, Sept. 20-21, pp. 257-266, 1999.
9. D.V. McAllister, F. Cros, S.P. Davis, L.M. Matta, M.R. Prausnitz, and M.G. Allen, "Three-dimensional hollow microneedle and microtube arrays," *Transducers '99*, Sendai, Japan, June 7-10, pp. 1098-1101, 1999.
10. I. Papautsky, J. Brazzle, H. Swerdlow, and A.B. Frazier, "A Low Temperature, IC Compatible Process for Fabricating Surface Micromachined Metallic Microchannels", *IEEE Journal of Microelectromechanical Systems*, (7) 267-73 (1998).
11. I. Papautsky, H. Swerdlow, and A.B. Frazier, "Surface Micromachined IC Compatible Technology for Fabricating Metallic MicroChannels", *10th IEEE International Workshop on Micro Electro Mechanical Systems*, Nagoya, Japan, January, 1997, pp. 104-109.
12. J. Brazzle, I. Papautsky, and A.B. Frazier, "Fluid-coupled metallic micromachined needle arrays," in Proc. *20th International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS '98)*, Hong Kong, Oct. 29 – Nov. 1, pp. 1837-1840, 1998.
13. I. Kadija, V. Chinchankar, V.T. Eckert, E.J. Kudrak, and J. Abys, "Electroplating of thick and ductile palladium: A new electroplating technology," in *Proc. of the 77th Annual Conference of the American Electroplaters and Surface Finishers Society*, July, 1990.
14. J. Black, *Biological performance of materials: Fundamentals of biocompatablity*, Marcel Dekker, New York, 1992.
15. S. D. Cramer and D. Schlain, "Electrodeposition of palladium and platinum from aqueous electrolytes," *Plating*, May, pp. 516-522, 1969.
16. E. J. Kudrak, J. A. Abys, H. K. Straschil, I. Kadija, and J. J. Maisano, "Palladium and palladium alloy plating for the 90's," presented at *Connectors '93*, East Midlands, England, May 19, 1993.
17. S. Brolin and G. Wettermark, *Bioluminescence Anal*, VCH Publ., 1992.
18. A. Campbell, *Chemiluminescence*, VCH Publ., 1989.
20. I.A. Levine, *Physical Chemistry 4th ed.*, McGraw-Hill, 1995, pg. 685.
21. W.H. Safranek, *The Properties of Electrodeposited Metals and Alloys*, New York: Elsevier, 1974.

FLOW CYTOMETRY ON MICROFLUIDIC DEVICES

Stephen C. Jacobson, Maxine A. McClain,
Christopher T. Culbertson, and J. Michael Ramsey

Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6142

ABSTRACT

Flow cytometry was demonstrated on a microfabricated fluidic device. The channels were coated to prevent cell adhesion, and the cells were transported electrophoretically by applying potentials to the fluid reservoirs. The sample stream was spatially confined in two dimensions at the cross intersection to conduct coincident light scattering and fluorescence detection. An *Escherichia coli* sample was labeled on-chip with a membrane permeable nucleic acid stain Syto 15 and counted.

INTRODUCTION

Interest in microfabricated instrumentation for chemical sensing and analysis has grown exponentially over the past decade primarily because these miniature instruments may provide information rapidly and reliably at low cost. Microfabricated fluidic devices (microchips) constructed on planar substrates are advantageous for manipulating small sample volumes, rapidly processing materials, and integrating sample pretreatment and separation strategies. The dexterity with which materials can be manipulated and the ability to machine structures with interconnecting channels with essentially zero dead volume contribute to the high performance of these devices. To carry out a complete assay, functional elements can be serially integrated on these devices and include filters, valves, pumps, mixers, reactors, separators, cytometers, and detectors. Coupling these elements together under computer control will enable the development of a wide range of microchip-based assays. One area of particular interest is the analysis of cells and cell populations, and a rapid screening technique for such assays is flow cytometry [1].

A few examples of cell manipulations on microfluidic devices have appeared where hydrodynamic [2,3] and electroosmotic [4,5] flows have been used to transport and sort cells on microchips. We are developing microfabricated fluidic devices for flow cytometry that incorporate electrokinetic focusing to spatially confine fluids [6] and particles [7]. The particles in this narrowed sample stream are then detected using light scattering and/or fluorescence detection. In this proceedings, we describe the operation and performance of a microfluidic device for cytometry that characterizes an *E. coli* sample labeled with a membrane permeable stain Syto 15.

EXPERIMENTAL

Fabrication. Two microchips designs similar to Figure 1 were transferred by UV exposure from the photomask (HTA Photomask) to glass substrates coated with chromium and positive photoresist (HOYA). After developing the photoresist, the chromium film was etched ($CeSO_4/HNO_3$), then the channels were etched into the substrate in a dilute, circulated HF/NH_4F bath. Channel access holes 2-mm in diameter were drilled in the cover plate. To form the closed network of channels, the cover plate was bonded to the substrate over the etched channels by hydrolyzing both surfaces, bringing the cover plate into contact with the substrate, and annealing at 500°C. Cylindrical glass reservoirs were then affixed on the cover plate using epoxy. The channels were 50 µm wide at the cross intersection. The channel widths were measured at half-depth using a stylus-based profiler. Before analyzing the sample, the microchip was sequentially rinsed for 15 minutes with 1.0 M NaOH, water, Run Buffer (CE-SDS Protein Kit, BioRad), water, and phosphate buffered saline (PBS). After each set of experiments, the microchips were rinsed in the same manner with 50 mM SDS and water.

Operation. Four high-voltage power supplies (UltraVolt) were connected to the sample, focus 1, focus 2, and waste reservoirs and independently controlled through a multifunction I/O card (PCI-MIO-16XE-50, National Instruments). Sample was continuously infused into the focusing chamber. The sample stream width was controlled by varying the potentials applied to the focus reservoirs relative to the sample reservoir.
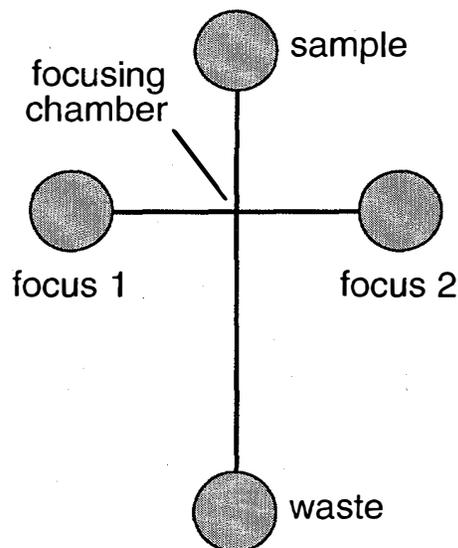
Figure 1. Schematic of microchip for flow cytometry.

Detection. For coincidence measurements, light scattering and laser-induced fluorescence signals were acquired simultaneously. The laser beam (Ar$^+$, 514 nm, 10 mW, Innova 300 FReD, Coherent) was brought to a focus 50 μm downstream of the focusing chamber. The incident angle of excitation was 45°. The scatter and fluorescence signals were collected simultaneously from the microchip using a 20x microscope objective (N.A. 0.4, Edmund Scientific). The scatter and fluorescence signals were spatially filtered (200 μm x 200 μm square), split using a dichroic filter (540 nm, Omega Optical), and measured by two photomultiplier tubes (PMT, 77348, Oriel). The fluorescence signal was again spectrally filtered using a bandpass filter centered at 560 nm with a 40 nm bandwidth (Omega Optical). The signals from the PMTs were then amplified (428-MAN, Keithley) and digitized at 1 kHz with the same multifunction I/O card and computer used for the voltage control. Data were collected in 60-second blocks. Statistical analysis from the Syto 15 stained cells was based on eight, one-minute runs.

Cells. All experiments were performed with the non-pathogenic *E. coli* strain Y1090. The cells are rod shaped, between 0.7 and 1.5 μm long, depending on their stage in cell division. A fresh culture was used for every experiment. A single colony, taken from an agar plate, was added to 2 ml of cell culture media with 50 μg/ml ampicillin (International Biotechnologies) and incubated for four to six hours. The culture was resuspended in water, and the cell concentration was measured at 600 nm with a spectrophotometer. A small volume of the suspension was added to diluted PBS (pH 7.5, J.T. Baker) in the sample reservoir of the microchip with the fluorescent dye prior to the experiment. The cell concentration in the sample reservoir was approximately 1.25 *10^8 cells/ml in a total volume of 40 μl. The Syto 15 dye (4 μM, Molecular Probes) was incubated with the cells in the sample reservoir for 10 minutes prior to the experiment. The focus 1,

focus 2, and waste reservoirs were each filled with 40 μl of PBS buffer.

RESULTS AND DISCUSSION

To prevent cell adhesion, the channel walls of the microchip in Figure 1 were coated with Run Buffer (BioRad). This product was effective in preventing adhesion, was hydrophilic, and had only to be flushed through the channel manifold. After the microchip was prepared, operation for several hours without noticeable degradation in cell flow was possible, but operation was not tested for periods over six hours. The Run Buffer, however, minimized electroosmotic flow, and as a result, the cells had to be electrophoretically confined. Focusing of cells is pictured in Figure 2 and was accomplished by applying 268 V to the sample reservoir, 43 V to the focus 1 reservoir, 0 V to the focus 2 reservoir, and 355 V to the waste reservoir. The cells were negatively charged in the PBS buffer and flowed from the sample to the waste reservoir in the presence of an electric field. At the cross intersection, the stream of cells was narrowed in two dimensions enabling single cell interrogation at the detection point.

Figure 3 shows the light scattering (bottom) and fluorescence (top) signals collected from Syto 15 labeled cells. A potential of 0 V was applied to the sample reservoir, 10 V to the focus 1 and focus 2 reservoirs, and 560 V to the waste reservoir. Each peak represents a single cell passing though the detection zone defined by the spatial filter. The correlation between the scatter and fluorescence channels was almost unity for these experiments. The fluorescence channel had a lower signal-to-noise ratio than the scatter channel, and consequently, some fluorescence peaks were not counted by the peak finding algorithm due to insufficient signal. The scatter peaks were nearly uniform in width and height, but a wide variation in fluorescence intensities was observed. The
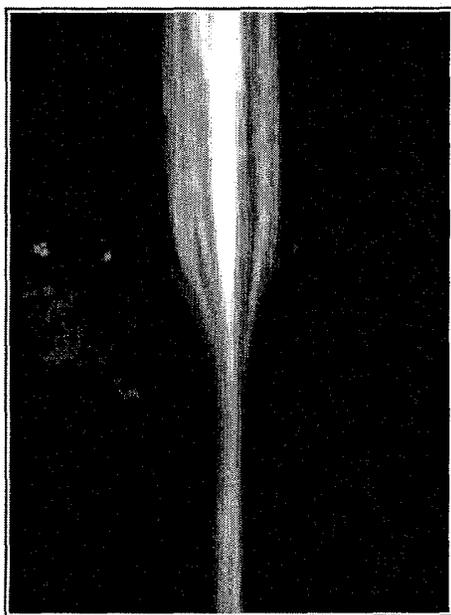
Figure 2. Fluorescence CCD image (5 s exposure) of electrophoretically focused cells labeled with Syto 15.
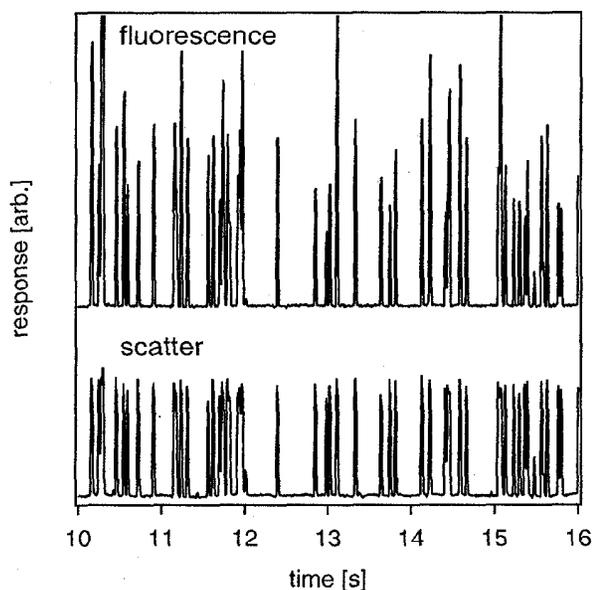
Figure 3. Light scattering (bottom) and fluorescence (top) signals from an *E. coli* sample stained with Syto 15.

variability in fluorescence from cells containing the Syto 15 stain was probably dependent on cell age and nucleic acid content. Also, with this configuration counting frequencies up to 14 Hz were recorded. Higher count rates were not possible here because at higher field strengths and cell concentrations, the cells started to aggregate hindering single cell detection.

In conclusion, these experiments demonstrate the potential of incorporating cytometry into a microfabricated fluidic device. Other assays to be presented include cell viability studies and immunoassays using a similar device and sample. Methods of increasing sample throughput, such as using higher operating flow rates and microchips with modified channel geometries, are being explored to achieve 100 to 1000 Hz counting rates per channel. Throughput can be further enhanced by multiplexing these cytometry elements in an array configuration [6].

ACKNOWLEDGMENTS

REFERENCES

1. H.M. Shapiro, "Practical flow cytometry," Wiley-Liss, New York (1995).

2. G. Blankenstein and U.D. Larsen, "Modular concept of a laboratory on a chip for chemical and biochemical analysis," *Biosensors and Bioelectronics, 13*, 427-438 (1998).

3. G. Blankenstein, L. Scampavia, J. Ruzicka, and G.D. Christian, "Coaxial flow mixer for real time monitoring of cellular responses in flow injection cytometry," *Cytometry, 13*, 200-204 (1996).

4. A.Y. Fu, C. Spence, A. Scherer , F.H. Arnold, and S.R. Quake, "A microfabricated fluorescence-activated cell sorter," *Nature Biotechnology, 17*, 1109-1111 (1999).

5. P.C.H. Li and D.J. Harrison, "Transport, manipulation and reaction of biological cells on-chip using electrokinetic effects," *Analytical Chemistry, 69*, 1564-1568 (1997).

6. S.C. Jacobson and J.M. Ramsey, "Electrokinetic focusing in microfabricated channel structures," *Analytical Chemistry, 69*, 3212-3217 (1997).

7. D.P. Schrum, C.T. Culbertson, S.C. Jacobson, and J.M. Ramsey, "Microchip flow cytometry using electrokinetic focusing," *Analytical Chemistry, 71*, 4173-4177 (1999).

A Multifunctional Silicon-based Microscale Surgical System

Il-Seok Son, Amit Lal*

Department of Electrical and Computer Engineering, University of Wisconsin
Madison, Wisconsin 53706

Bill Hubbard, Tim Olsen

Department of Ophthalmology, University of Minnesota
Minneapolis, Minnesota 55455

ABSTRACT

In this paper we report our progress on the development of silicon-based surgical tools. We report on the integration of piezoresistive sensors in noise insensitive Wheatstone configuration, silicon nitride/platinum composite overhanging blades, fluid channels, and electrical packaging features to make the device highly multifunctional. Multiple polysilicon piezoresistors were integrated onto bulk micromachined ultrasonic transducers, that were characterized for measuring the ultrasonic strain and the force applied to the cutting tip. The use of thin film silicon nitride/metal composites for cutting swine retinas at the cellular level is demonstrated.

INTRODUCTION

Microscale smart surgical tools can benefit many surgical operations where microscale or cellular level tissue manipulation is required [1]. For example, retinal surgery requires the manipulation of few layers of retina cells in an often-performed retinal-detachment surgical procedure. Other examples of cellular-scale removal of cells are epithelial cell transplants, plastic surgery, corneal incisions, and skin grafts. Although the glass pipette patch probes have existed for several decades to measure cellular potentials and manipulate tissue at the microscale, they requires high optical magnification to view the area being cut with expensive micromanipulation technology without integrated sensors. Sharp microscale blades made of diamond or other ceramics have been used but result the technology precludes the integration of solid-state sensors. The alternative of bonding sensors onto microscale cutting tools requires expensive packaging and leads to reliability problems.

Micromachined silicon probe technology allows one to fabricate cutting tools that have microscale cutting edges suitable for cells ranging in size from 10 to 100 microns. Micromachined neural probes have been designed to measure cellular potentials for at least a decade [5] and continued advances in that field have resulted in highly engineered systems. Force measurement probes with micromachined probes have also been demonstrated using piezoresistors [6].

We have previously demonstrated silicon-based high-intensity ultrasonic actuators for tissue cutting [2-4]. We have previously presented the concept of high-intensity ultrasonically driven silicon-based surgical tools with integrated sensors. Silicon's material properties of high yield strain and high speed of sound make it suitable for high-intensity ultrasound applications. The ability to micromachine and integrate sensors makes silicon ultrasonic actuators more versatile and allow integration of several functions on the transducer.

In this paper we present results on integration of piezoresistive strain gauges, microscale cutting edges, and packaging advances to result in a multifunctional tool that might allow tissue characterization while cutting the tissue. The fabricated device is a bulk-micromachined surgical tools that allow both mm^3 and μm^3 volume of tissue removal using piezoelectric actuators that can be driven at ultrasonic frequencies. Integrated piezoresistors are implemented on the silicon micromachined structure that can be used for sensing forces and closed-loop control of the ultrasonic amplitude. Spatial distribution of piezoresistors enables the measurement of spatial variations in strain enabling novel tissue characterization. With microfluidic channels for fluid sampling and delivery, this system has the possibility of integrating several surgical functions of cutting tissue at mm and μm scale, pumping cut tissue, sensing tissue stiffness and closed loop control of ultrasonic energy.

FABRICATION and ASSEMBLY

A schematic and photograph of the device is shown in Figure 1. The half needles are etched using the TMAH (Tetra-Methyl Ammonium Hydroxide) anisotropic etchant and bonded together to form the needle-like structure using silk-screened adhesive. Each half needle consists of integrated three strain-gauges using four polysilicon piezoresistors formed by the process shown in Figure 4. A 3" <100> wafer is coated with 1.0 μm of low-stress silicon nitride, which is later used as the mask layer for the TMAH etch. LPCVD polysilicon thin film (1.3 μm) is deposited on the silicon nitride and will form the piezoresistors. The polysilicon is doped heavily p-type by Boron implantation using 60 keV with a dose of 4.0×10^{15} /cm^2. After patterning the polysilicon layer in the shape of resistive elements, the layer is thermally oxidized to form oxide passivation. Contacts are etched into the polysilicon resistors followed by sputter deposition of 0.4 μm thick aluminum film. Aluminum interconnects to form the Wheatstone bridge and wires to the external electronics are defined. The exposed silicon nitride layer is now patterned in a front-back side alignment system to create the horn shape and the embedded

channels inside the needle structure. The exposed silicon is etched in 5.0 % TMAH solution with silicic acid as an aluminum passivation additive. The TMAH etch was chosen to ensure aluminum passivation and future compatibility with integration of active electronic devices on the cutting tool. The piezoresistors are arranged in a Wheatstone bride configuration such that the strain sensitivity in the axial direction is amplified over the other possible direction as shown in Figure 5.

The Wheatstone bridge also eliminates the RF coupling from the PZT electrode driven at large voltages. The piezoresistors were also placed at three locations along the length of the needle structure as shown in Figure 3. The distribution of the strain gauges allows the measurement of the ultrasonic strain and strain due to forces during tissue cutting.
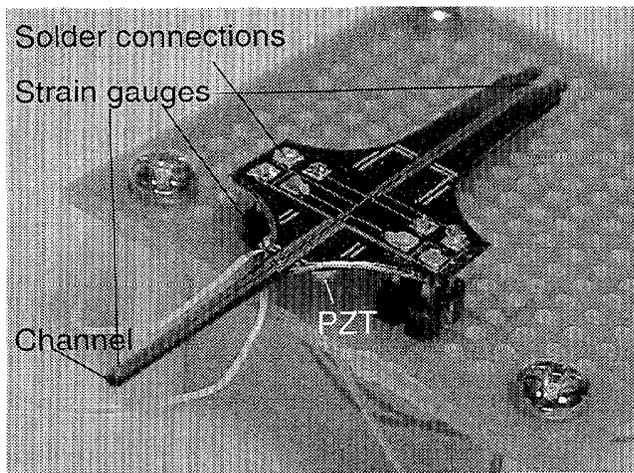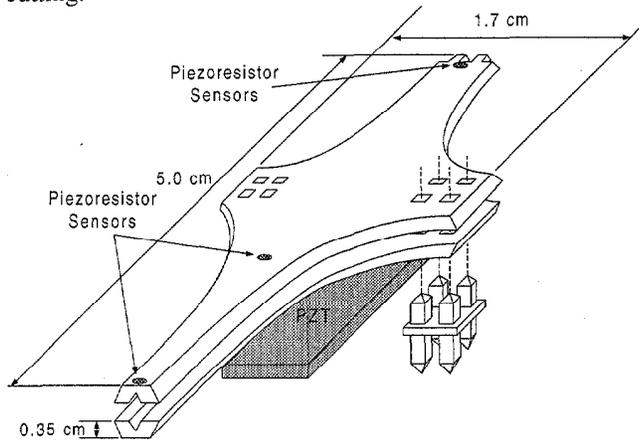
Figure 1. The schematic and photograph of the silicon needle structure mounted on a PC board

A piezoelectric PZT (lead-zirconate-titanate) ceramic plate is adhesively attached at the center of the needle structure. The placement at the node and the length of the PZT are chosen to optimize the electromechanical coupling to the silicon structure [8].

The structure is designed to operate in its half-wavelength resonance. Hence every point along the structure vibrates expect at the node. The piezoresistors are connected

to the external electronics via metal pins placed at the node. Physical wires placed anywhere else on the vibrating surface would effect the modal shape by mass and spring loading. Aluminum pads are shaped around a square anisotropically etched orifice as shown in Figure 1. The pins of a DIP package are fed through the orifice and silver-epoxied or wire-bonded to the aluminum interconnect. In addition to providing the electrical connections the pins also provide a rigid anchor for the vibrating structure. The device mounted on the DIP package can easily be placed on a PC board with electronics.

A silicon nitride blade is formed at the edge of the silicon structure by using a double-side exposure technique (Figure 2) The nitride blades were shadow coated with sputtered titanium/platinum thin film sandwich of 0.3 μm thickness. The metal coating substantially decreases the brittleness of the cantilever and makes possible the ability to penetrate tissue without breaking.
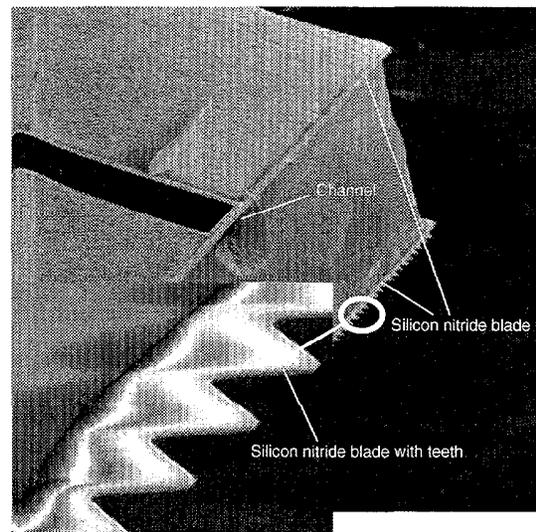
Figure 2. The cutting tip of the ultrasonic device can have silicon nitride blades for microscale cutting.

DEVICE CHARACTERIZATION

The ultrasonic response of the device can be characterized by measuring the electrical impedance versus frequency of the device. For example, the electromechanical coupling to the different bimorph and longitudinal modes can be obtained from the measured electrical impedance versus drive frequency at the PZT terminals. The ½-wavelength longitudinal mode was found at 112 kHz as predicted from an analytical model. The parameters of the equivalent circuits can be extracted from the impedance data. The quality factor of the resonance was measured to be 306 and K^2_{eff} was .025. The product of the two quantities is the figure of merit of an ultrasonic transudcer and is 7.7 for the transducer. The displacement versus applied drive voltage was measured optically, and found to be linearly dependent on drive voltage as 50 nm/V_{pp} drive at the resonance frequency of 112 kHz.

The piezoresistor value obtained after the fabrication and annealing of the contacts was 20 kΩ. A variation in resistance value of approximately 4 % was measured across the wafer resulting from polysilicon thickness, doping, and oxidation variations. These small differences lead to DC offsets for the resulting in large DC voltages from the differential amplifiers. In future, the resistors can be laser-trimmed to eliminate the DC offsets.

The strain gauges were characterized by measuring the differential signal across the Wheatstone bridge amplified by an Analog Devices AD620 instrumentation amplifier (set at a gain of 30) as a function of applied forces. A force transducer (Sensym FS-03) was mounted against the needle structure, as the needle was pressed against the transducer, the calibrated readings from the Sensym device and the needle piezoresistors were measured. Careful attention was given to eliminate bending as that also produces a strain signal. In particular the tip piezoresistor has negligible bending even if bending moments are applied. Figure 6 shows the measured voltage from the instrumentation amplifier as a function of applied force. Using this data and assuming that the compression was uniaxial from the needle tip to the anchor pins, one can derive an expression for the gauge factor of the device to be

$$G = \frac{\Delta V}{2V} \frac{1}{F} * E_{Si} * A,$$

where V is the DC voltage across the Wheatstone bridge, F is the applied force, E_{Si} is the silicon Young's modulus, and A is the silicon needle cross section at the piezoresistor. Using this equation and the measured data, we obtain a gauge factor of 25, which is in very good agreement with value of 27 for similarly doped polysilicon piezoresistor published earlier [7].

ULTRASONIC STRAIN MEASUREMENT

The ultrasonic strain as a function of applied drive voltage was also measured. The differential strain value as a function of applied voltage to the PZT is shown in Figure 7. The strain value was calculated using the gauge factor of 25 and the measured ultrasonic differential signal. The strain values obtained using the piezoresistors agree with the value obtained from knowing the tip displacement and the mode shape of the ½-wavelength resonance. This data clearly demonstrates that the piezoresistors can be used to measure ultrasonic strain for feedback control of the device.

THIN FILM SILICON NITRIDE/METAL BLADES

In order to investigate the ability of the silicon nitride/metal composite blades to withstand cutting forces, the devices, mounted on a micrometer, were pushed into a swine retina. The silicon nitride cutting edge was able to penetrate the vitreous and the retina layer several times without any blade fracture (Figure 3). Furthermore, the cantilever was able to release retinal cells, much more so than when cut without the silicon nitride blade. Cutting results on other tissues such as onion resulted in precision cellular cutting (Figure 3). In both cases, applied ultrasound energy resulted in acoustic streaming-related fluid motion

and release of materials in the tissue. Notably blades made only of silicon nitride sometimes broke during cutting indicating the importance of metal films to increase blade toughness.
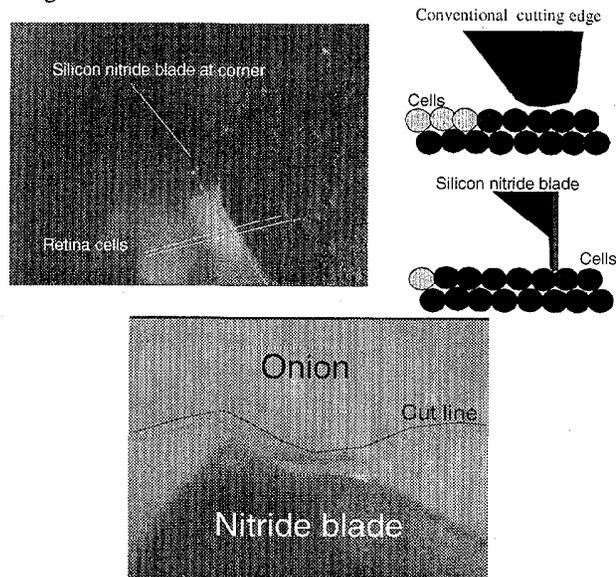
Figure 3. Top shows the penetration of the silicon nitride cutting edge into swine retina and releasing highly reflective cells. Bottom shows the silicon nitride blade cutting an onion cell along a controllable cut line.

CONCLUSIONS

We have presented a multifunctional ultrasonic transducer suitable for surgery that includes integrated strain gauges, silicon nitride blades, microfluidic channels, and through holes for easy electrical interconnects to external circuitry. The integrated sensors have been characterized for use in feedback control of the ultrasonic amplitude and measure forces during cutting tissue. One of the exciting possibilities of the integrated sensors is to measure the applied force to the silicon micromachine the surgical tool operator. Such a capability might allow the doctor to a highly sensitive version of the traditional percussive testing of tissue properties. The silicon nitride blades coated with Ti/Pt coatings were strong enough to penetrate soft tissues repeatedly. The device allows cutting tissue ultrasonically or statically, while sensing the tissue forces being applied. In the near future we plan to study the efficacy of the device in performing ophthalmic surgery in animal eyes.

REFERENCES

[1] Charles, S., *Vitreous Microsurgery*, 2nd Edition, Wilkins and Williams, Baltimore, Madison, 1987.

[2] Lal, A., White, R. M., "Silicon micro-fabricated horns for power ultrasonics," Transducers 95, June 1995, Stockholm, Sweden.

[3] Lal, A., White, R. M., "Micromachined silicon ultrasonic needle," IEEE Ultrasonics Symposium, Nov. 7-10, 1995, Seattle, USA.

[4] Son, Il-Seok, Lal, A., "Silicon Ultrasonic Horns with Integrated Strain Gauges," presented at the Ultrasonics International 1999, Copenhagen.

[5] Najafi, K., Hetke, J.F., "Strength Characterization of Silicon Miroprobes in Neurophysiological Tissues," IEEE-Transactions on Biomedical Engineering, vol. 37, no. 5, May 1990, pp. 474-481.

[6] Kewley, D. T., et al., "Plasma Etched Neural probes," Hilton Head, 1996, pp. 266-271

[7] Obermeier, E., Kopystynski, P., Niebl, R., "Characteristics of Polysilicon Layers and Their Application in Sensors," Records of the IEEE Solid State Sensors Workshop, 1986.

[8] Lal, A., White, R. M., "Optimization of the silicon/PZT longitudinal mode resonant transducer," Microelectromechanical Systems, Proceedings of the ASME World Congress 1997, Dallas, pp.61-67.

Figure 5. SEM of the polysilicon piezoresistors arranged in a Wheatstone configuration to measure strain in the axial direction of the transducer

Figure 4. The fabrication process for the silicon cutter. (a) Low- stress silicon nitride deposition, (b) polysilicon deposition and ion implantation, (c) oxidation of polysilicon after patterning, (d) contact opening to polysilicon, (e) metal deposition, (f) after patterning aluminum, patterning silicon nitride film, and bulk silicon etching in TMAH
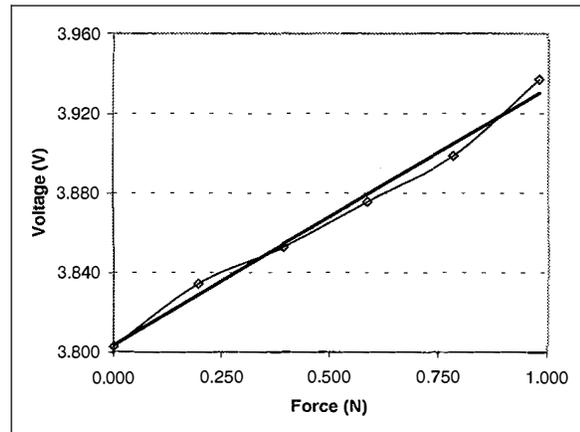
Figure 6. Amplified differential output from the tip strain gauge as a function of applied compressive force on needle yields linear response and a gauge factor of 25 for the polysilicon piezoresistors
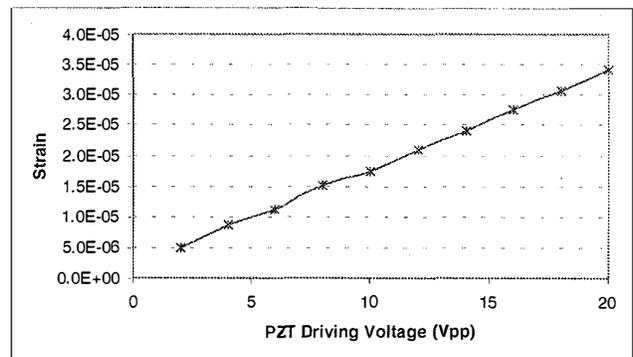
Figure 7. Measured strain at the middle strain gauge as a function of drive sinusoidal voltage across the piezoelectric PZT ceramic

A NEW CLASS OF HIGH FORCE, LOW-VOLTAGE, COMPLIANT ACTUATION SYSTEMS

**M. Steven Rodgers[1], Sridhar Kota[2], Joel Hetrick[3], Zhe Li[2], Brian D. Jensen[2],
Thomas W. Krygowski[1], Samuel L. Miller[1], Stephen M. Barnes[1], and Michael S. Burg[1]**

1) Intelligent Micromachine Department, Sandia National Laboratories
Albuquerque, New Mexico 87185-1080
http://www.mems.sandia.gov

2) Department of Mechanical Engineering
and Applied Mechanics
The University of Michigan-Ann Arbor
Ann Arbor, Michigan 48109-2125

3) Department of Mechanical Engineering
The University of Wisconsin-Madison
Madison, Wisconsin 53706

ABSTRACT

Although many actuators employing electrostatic comb drives [1] have been demonstrated in a laboratory environment, widespread acceptance in mass produced microelectromechanical systems (MEMS) may be limited due to issues associated with low drive force, large real estate demands, high operating voltages, and reliability concerns due to stiction. On the other hand, comb drives require very low drive currents, offer predictable response, and are highly compatible with the fabrication technology. To expand the application space and facilitate the widespread deployment of self-actuated MEMS, a new class of advanced actuation systems has been developed that maintains the highly desirable aspects of existing components, while significantly diminishing the issues that could impede large scale acceptance. In this paper we will present low-voltage electrostatic actuators that offer a dramatic increase in force over conventional comb drive designs. In addition, these actuators consume only a small fraction of the chip area previously used, yielding significant gains in power density. To increase the stroke length of these novel electrostatic actuators, we have developed highly efficient compliant stroke amplifiers [2]. The coupling of compact, high-force actuators with fully compliant displacement multipliers sets a new paradigm for highly integrated microelectromechanical systems.

INTRODUCTION

Electrostatic comb drives [1] are ideal candidates for microelectromechanical applications requiring high frequency, very low current, and well controlled force generation over several microns of displacement. However, conventional comb drives typically only produce a few µN's of drive force, while often requiring non-resonant mode operating voltages on the order of 100 volts [3]. In addition, conventional comb drives consume significant real estate and can occupy more than a square millimeter of chip area, placing a fundamental limit on the size and complexity of advanced microelectromechanical systems. In response to this, a new class of advanced actuation systems has been developed. Actuators that provide more than 200 times the force per unit area of previous designs and multipliers with gains from 5 to 50 and efficiencies greater than 80% (force out x displacement out > 0.80 x force in x displacement in) have been designed, fabricated, and successfully demonstrated. As the method for synthesizing the novel displacement multipliers has previously been reported [2], this paper focuses on the design of the electrostatic actuation component.

COMB DRIVE STABILITY ISSUES

The equation governing the drive force generated by an electrostatic comb drive [3] is given in equation 1:

$$F = \frac{n \varepsilon t}{g} V^2 \quad (1)$$

Where n = number of fingers, ε = electrical permitivity, t = finger thickness, g = gap between fingers, and V = voltage applied. Note that n, t, and g are the parameters related to the mechanical design, with the goal being to maximize F while consuming the least amount of die area. Therefore, n and t should be as large as possible and g should be as small as possible. All of the devices presented in this paper were fabricated in the SUMMiT-V surface micromaching process [4,5] at Sandia's Microelectronics Development Laboratory, so t was maximized by always using every mechanical layer available in this technology to define the comb fingers. This leaves only n and g to work with. The highest finger density can be achieved by using the minimum width and space geometry permitted by the design rules, which are both 1.0 µm for this process. Although it is easy to design such a structure, to actually fabricate it and have it work is a different matter. This is because lateral instabilities become increasingly difficult to overcome as the comb geometry is downsized.

In addition to the force along the desired line of motion stated by equation 1, there is a substantial parallel plate attractive force between each of the fingers as shown in Figure 1. Equation 2 shows this force is directly proportional to the area A between the fingers and inversely proportional to the gap squared [3]. As long as the

$$F_{pp} = \frac{\varepsilon A}{2g^2} V^2 \quad (2)$$
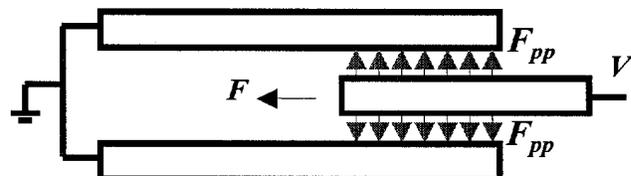
Figure 1. *Forces acting on a comb drive finger when a voltage is applied. Vertical arrows represent parallel plate attraction.*
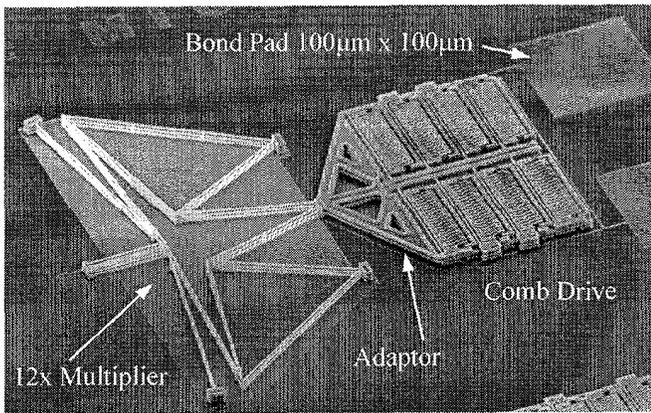
**Figure 2.** Scanning electron microscope (SEM) image of new comb drive with compliant displacement multiplier.

Figure 4. This structure firmly holds the banks of comb fingers in proper alignment under high electrostatic forces.

center comb finger remains precisely between the outer two, the parallel plate forces cancel. Any asymmetry or off axis loading, however, will create an imbalance that could easily overcome the restoring force of the structure resulting in lateral clamping of the comb fingers [6] and actuator failure. The issue with reducing the gaps between fingers to generate more force is this instability increases as the square of the gap dimension while the desired force only increases in direct proportionality. The problem is further compounded when very thin comb fingers are used to create a high density array, since their lateral stiffness varies as the cube of their width.

DESIGNING FOR STABILITY

The authors know from experience that developing a successful high force per unit area comb drive requires taking extreme care to ensure symmetry and mechanical stability. One of the actuation systems resulting from this effort is shown above in figure 2. The first issue to be addressed was the prevention of lateral bending of thin highly-compliant comb fingers due to asymmetrical electrostatic fields. This is normally only a problem for the end fingers, since these typically see a comb finger on one side and very different geometry on the other. Therefore, the comb configuration in Figure 3 was employed to maintain balanced lateral forces on both sides of the fingers regardless of actuator displacement.

The second task was to create a rigid comb mounting structure that would not distort under high electrostatic forces. Common practice is to suspend banks of comb fingers by attaching one end to a single central beam and letting the other end float [1,3]. The new design incorporates outer beams as shown in Figure 4 to add additional support, significantly stiffening the assembly. The beams and the fingers mounted directly to them form the moving
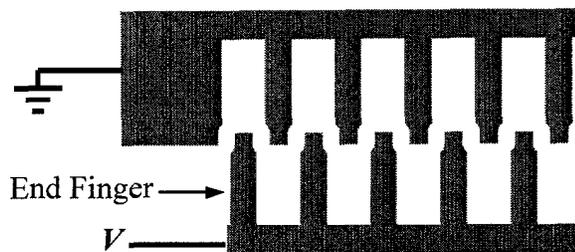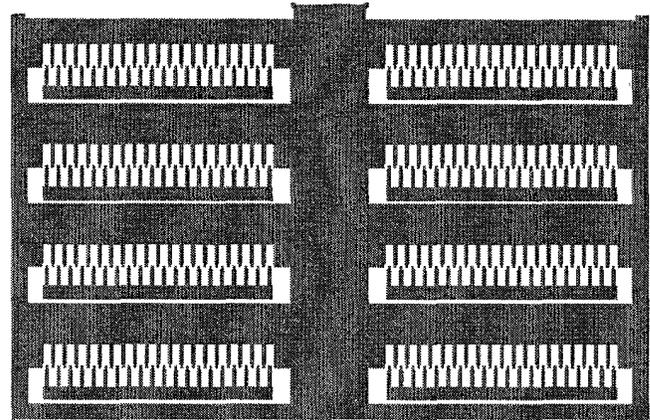
shuttle assembly, which is held at ground potential. Actuation occurs when voltage is applied to the stationary mechanically isolated banks of fingers. Desired motion is displacement of the shuttle towards the bottom of the figure. However, this movement is opposed by a parallel plate attraction force that develops between the main support beams for the stationary fingers and the nearby horizontal shuttle beams. This force is effectively eliminated by shielding with grounded electrostatic fences as shown in figure 5.

Comb drive shuttle assemblies are supported by springs that must be compliant enough to allow movement along a line parallel to the fingers, yet stiff enough to prevent transverse movement which could result in lateral clamping of the comb fingers. These springs are normally attached to both sides of the shuttle assembly. Since this approach consumes significant amounts of valuable die area, a new high-stability suspension system was developed that allows the springs to be defined almost completely within the boundary established by the perimeter of the shuttle assembly. Support springs were fabricated in both the upper and lower mechanical layers and anchored together through openings cut in the central shuttle beam, which significantly increases the vertical or out of plane stiffness. The close-up in Figure 6 shows details of the suspension system and other previously described features.

The SUMMiT-V process has 4 mechanical levels of polysilicon and a poly0 electrical interconnect layer that also serves as a ground plane. The first 2 mechanical layers, poly1 and poly2, were laminated together to form a rigid 2.5-μm thick lower layer that was patterned by a single mask, while the 2.25-μm thick poly3 and poly4 layers were used to form the middle and upper levels of the actuator. Therefore, only 3 mechanical levels are apparent.

Figure 3. Balancing electrostatic forces on both sides of the thin comb fingers prevents shorting due to bending.
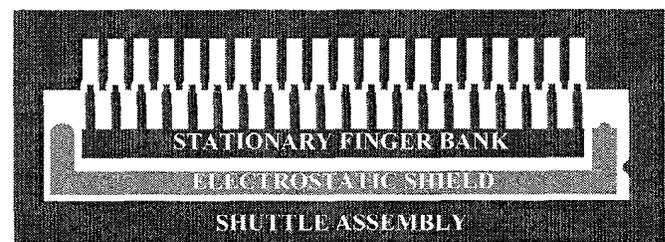
Figure 5. Grounded electrostatic shields prevent parallel plate attraction between the shuttle and stationary fingers.
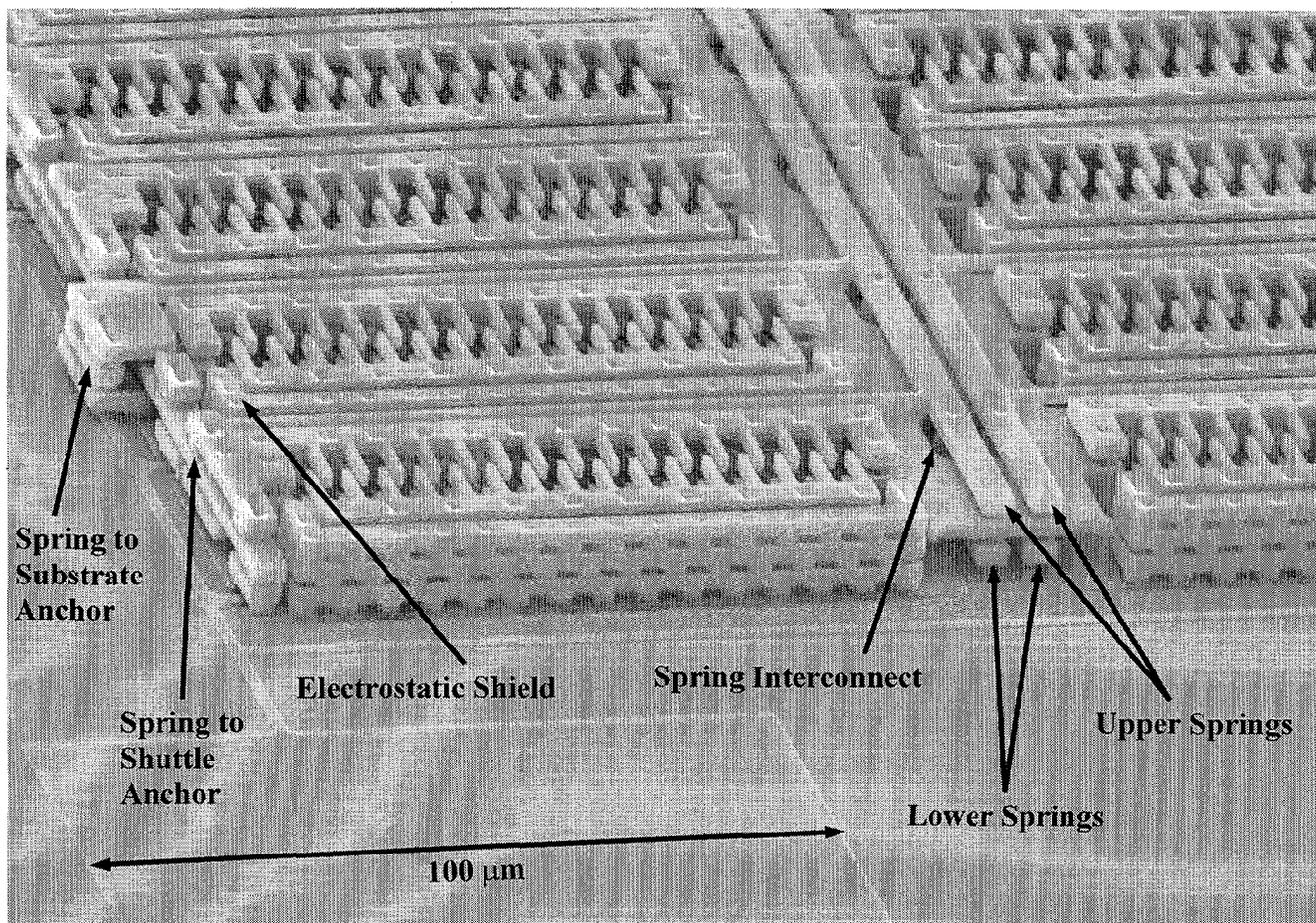
Figure 6. This SEM close-up of a portion of the new actuator assembly shows details of features discussed in text.

The actuator just described provides a 6-μm stroke with full displacement being obtainable with just 22-35 volts, depending on the specifics of the suspension system. The actuator was also designed as a cluster of modular subassemblies that can be easily arrayed to provide the force necessary for the intended application. The largest fabricated to date is 1450-μm long by 475-μm wide, and has a calculated drive force of more than 2.5mN at 100 volts.

SECOND ACTUATOR DESIGN

A second actuator has been developed that only provides a 2-μm stroke. However, it does this in a very small area. The entire electrode and shuttle assembly is just 116 x 71 μm, or less than the area of a typical 100 x 100-μm bond pad. Yet it produces about the same force (~15 μN at 100V) as the large 1200 x 1000-μm comb drives that Sandia has been using to drive the microengines [7]. Illustrated in Figure 7, this device has alternating rows of stationary electrodes that are anchored to the substrate, and electrodes that are connected to support beams. The beams are attached at both ends to a rigid rectangular frame which is held at ground potential. When a voltage is applied to the stationary fingers, the shuttle moves to align the electrodes horizontally as show in upper right of figure. This actuator features self-limiting displacement, so no mechanical stops are required. Once enough voltage is applied to almost fully displace the shuttle assembly, additional voltage provides very little additional displacement.
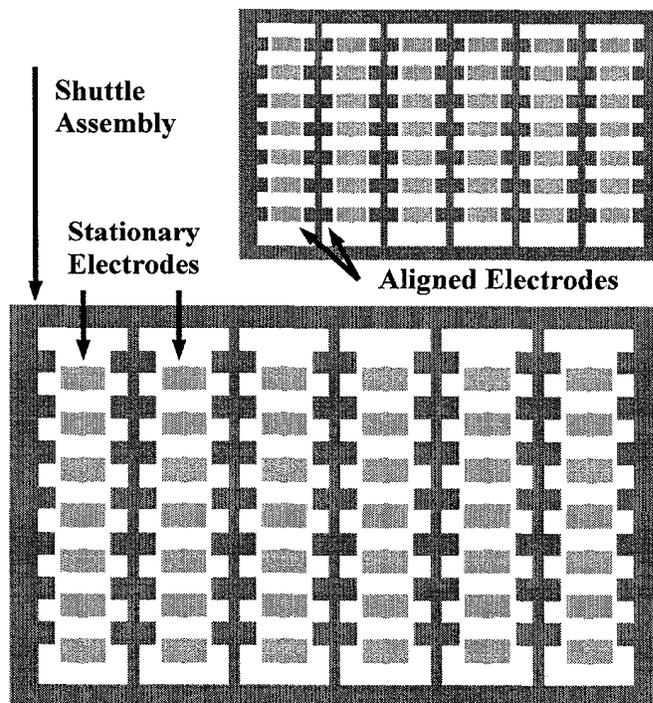
Figure 7. Second actuator design. Fabricated position shown on bottom; actuated position upper right.

212
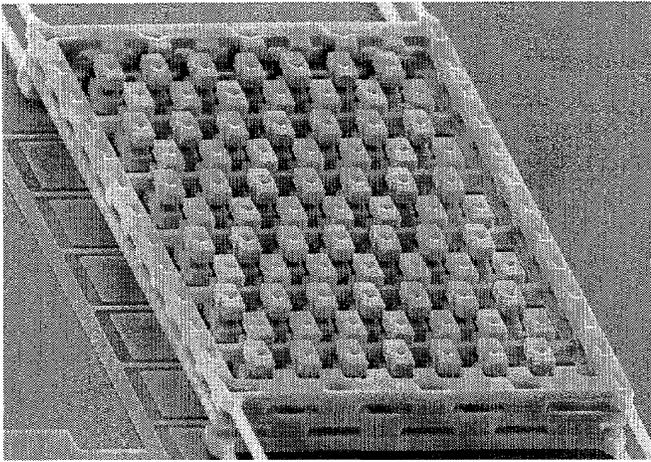
Figure 8. *SEM image of second actuator design.*

A close-up view of this actuator is shown above in Figure 8. One benefit from designing this device to only move 2 or 3 μm is that an extremely rigid and stable suspension system can be used. Simple 100-μm to 200-μm long and 2-μm wide cantilever beams have successfully been used to support the structure shown in the figure. This device has demonstrated the ability to handle 190V, our test system limit, without becoming unstable. So very high forces can be generated in very small spaces in applications where supplying high voltage is not a problem.

Most applications, however, require lower voltage actuation. Like the actuator shown in figure 6, these components were also designed to be modular. In figure 9 a 3 x 3 array of the actuator shown in figure 8 is being used in conjunction with a compliant displacement multiplier to create an actuator that provides enough displacement (~ 10 μm) at 30 V to drive a ratcheting component.

CONCLUSIONS

In spite of their generally high voltage and large die area requirements, surface micromachined electrostatic actuation systems have typically only provided marginal drive forces. This paper
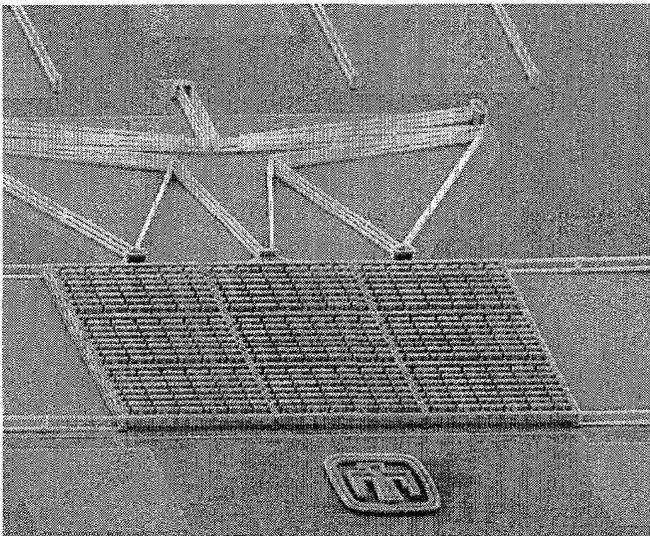
Figure 9. *This compliant displacement multiplier is being driven with a 3 x 3 array of the actuator in figure 8 to reduce the voltage required for operation.*

shows that such attributes are not fundamental limitations. Novel short-stroke electrostatic actuators that provide dramatic increases in power density have been demonstrated, and further gains are still possible. These devices can either be used directly or coupled to very efficient displacement multipliers to meet application requirements. Since drive force will be reduced by the multiplication factor, best performance will be obtained when the multiplier is optimized for a given actuator and load.

Yield and reliability are also improved compared to previous designs. The short, rigid suspension systems used in both the drives and multipliers keeps real estate requirements low while simultaneously making the structure very resistant to surface tension effects during the final release process, thus reducing stiction concerns.

The actuation systems described here are already being used in many of the surface micromachining programs at Sandia National Laboratories, and design enhancement is ongoing.

ACKNOWLEDGEMENTS

REFERENCES

1) W. C. Tang, Ph.D. Thesis, University of California, Berkeley, CA. (1990).
2) Sridhar Kota, Joel Hetrick, Zhe Li, Steve Rodgers, and Thomas Krygowski, "Synthesing High-Performance Compliant Stroke Amplification Systems for MEMS", *Proceedings of the IEEE Thirteenth International Micro Electro Mechanical Systems Conference*, Miyazaki, Japan, 1/23-27/2000, pp. 164-169.
3) S. L. Miller, J. J. Sniegowski, G. LaVigne, and P. J. McWhorter, "Friction in surface micromachined microengines", *Proc. SPIE Smart Electronics and MEMS*, 2722, 2/28-29/96, San Diego, CA, (1996) pp. 197-204.
4) M. S. Rodgers and J. J. Sniegowski, "5-Level Polysilicon Surface Micromachine Technology: Application To Complex Mechanical Systems", *Technical Digest of the 1998 Solid State Sensor and Actuator Workshop*, Hilton Head Island, SC, 6/8-11/1998, pp. 144-149.
5) More technical information can be found at the web site http://www.mems.sandia.gov.
6) S. L. Miller, M. S. Rodgers, G. La Vigne, J. J. sniegowski, P. Clews, D. M. Tanner, and K. A. Peterson, "Failure Modes in Surface Micromachined Microelectromechanical Actuation Systems", *Microelectronics Reliability, 39*, 1229 (1999).
7) J. J. Sniegowski, S. L. Miller, G. LaVigne, M. S. Rodgers, and P. J. McWhorter, "Monolithic Geared Mechanisms Driven by a Polysilicon Surface Machined On-chip Electrostatic Engine", *Technical Digest of the 1996 Solid State Sensor and Actuator Workshop*, Hilton Head Island, SC, 6/3-6/96, Transducers Research Foundation, Cleveland (1996) pp.178-182.

Optimized Scratch Drive Actuator for Tethered Nanometer Positioning of Chip-Sized Components

Ryan J. Linderman, Victor M. Bright
NSF Center for Advanced Manufacturing and Packaging of Microwave,
Optical, and Digital Electronics (http://mems.colorado.edu)
University of Colorado at Boulder, CO 80309-0427

ABSTRACT

A novel surface micromachined robot, capable of pushing a 2x2x0.5 mm chip several millimeters while being powered by a gold wire tether is presented. The micro robot uses a unique assembly and packaging process and is created from large arrays of scratch drive actuators (SDA) tethered by gold wires bonded prior to release in an HF acid bath. The wire tether allows for robot transfer to other substrates and driving surfaces, where assembly line positioning tasks can be performed. A device model has been derived to define optimum SDA plate design for any fabrication process. The driving signal has also been optimized, consuming much less electrical power than previously published signals. Because the step size of the individual SDA is approximately 30 nanometers, the tethered SDA robots are expected to achieve similar placement precision for chip sized components. The SDA robot (Scratchuator) demonstrates the potential for low cost nanometer positioning of chip sized components for alignment and flip chip assembly.

INTRODUCTION

The scratch drive actuator is the only micro system to provide high precision (~30nm) long range motion with a large output force. Previously published papers have highlighted the use of the SDA for optical autocorrelators[1], 3D micro stages[2] as well as novel mechanisms for ultra-thin active catheters[3]. The stepping motion of the SDA occurs when an electrostatic load deflects the plate until the bushing is scratched forward upon which the voltage is reduced and the plate is pulled forward (see figure 3). Little modeling of optimum plate design and voltage response has been published on SDA technology. Previously published SDA papers have reported plate designs for a specific fabrication process, however when an alternative fabrication method is used, the SDA plate must be redesigned to account for the new bushing geometry, plate thickness, insulator dielectric properties, Young's modulus, etc. (see figure 1). In this paper modeling functions to define optimum plate length for any fabrication process and material selection are presented to assist in the development of new applications and alternative microsystem material choices.
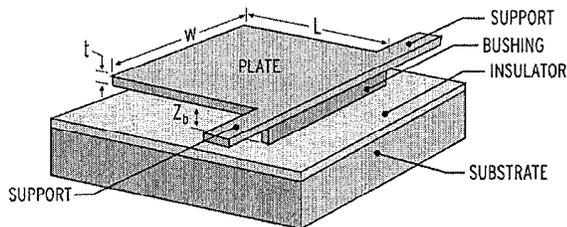
Figure 1. Scratch drive actuator nomenclature.

Once the individual SDA has been modeled for a given fabrication process, an optimum plate design can be chosen and experimentally verified. By designing a series of test stands with incremental variations in design parameters, the experimental optimum can be evaluated with intended applications in mind.

Because the SDA offers high force, long range and nanometer precision placement capabilities, it is well suited to meet the needs of future optics and flip-chip assembly operations. The limits of precision machined lead screws and large scale positioning systems have already been realized at significant costs. Thus the development of a microsystem based alignment and positioning actuator, such as shown in figure 2, stands to have a dramatic impact on chip and wafer based assembly operations.
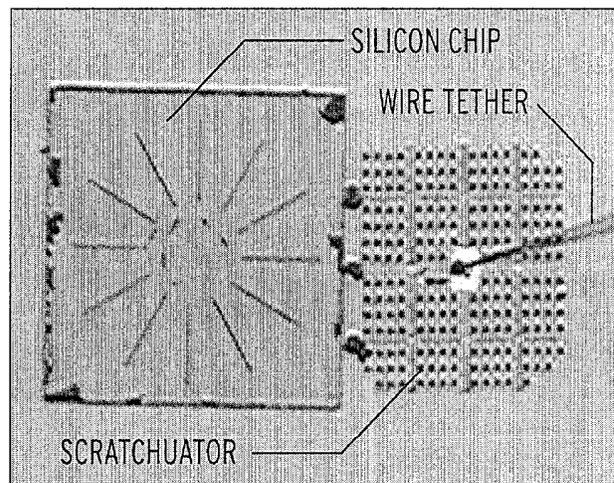
Figure 2. 188 SDA robot (Scratchuator) positioning a 2x2x0.5 mm chip.

DEVICE MODEL

The problem encountered when designing SDA for a given fabrication process is defining an appropriate plate length for a given stiffness material and bushing height. Plates that are too long flatten out against the insulating nitride layer after the snap through voltage is reached — limiting the output force of the SDA. Plates that are too short require exceptionally high voltage levels to step forward which have adverse effects on wear and life span of the mechanism. Because nearly all the work from the SDA is produced from longitudinal deflection, the width of the plate does not play a dominant role in optimization however, the width does have a dramatic effect on the squeeze film damping and frequency response of the SDA.

Two characteristic deflections must be modeled in order to determine the optimum plate length for a given fabrication process. The snap through voltage occurs when the electrostatic loading critically overpowers the elastic stiffness of the SDA resulting in the plate tip snapping down to touch the nitride dielectric layer. The priming deflection is the next function of plate length to be modeled and occurs when the SDA plate tip is deflected enough to flatten to a zero slope at the free end as shown in figure 3. For an optimized SDA plate, the snap-through voltage is equal to the priming voltage — resulting in the best balance of drag friction and voltage response.
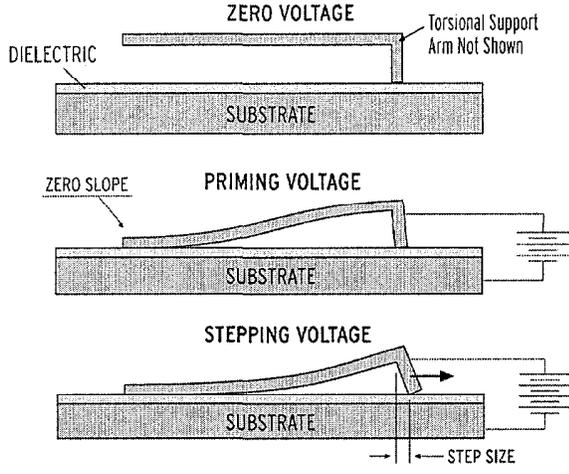
Figure 3. Concept of SDA stepping motion. In order to produce forward motion the SDA oscillates between the stepping and priming deflection.
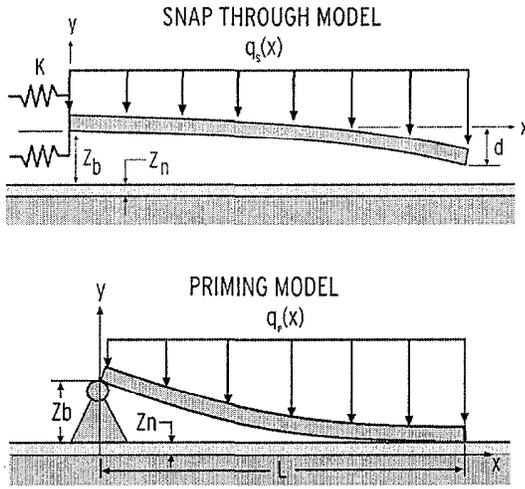
Figure 4. Free body diagrams for snap through and priming model.

The relationship for tip deflection of a torsionally supported cantilever with a distributed electrostatic load, $q_s(x)$, is given by[4]:

$$d = \int_0^L q_s\left(x\right) \frac{wx}{6EI}\left(3xL-x^2+6KL\right)dx \qquad (1)$$

Where K is the torsional flexibility defined by the rectangular support arm dimensions, ε is the permittivity of free space, E is Young's modulus, and I is the moment of inertia. Assuming square law curvature for the electrostatic loading[5] and a maximum separation distance of $Zm=Zb+Zn$:

$$q_s\left(x\right) = \frac{\varepsilon \kappa_1}{2}\left[\frac{V}{z_m-\left(\frac{x}{L}\right)^2 d}\right]^2 \qquad (2)$$

Where V is the applied voltage, κ_1 is the effective dielectric constant of the air ($\kappa_A=1$) and nitride ($\kappa_N=5.7$) layers below the SDA plate for the snap through model. Because the plate is curved towards the nitride layer, κ_1 is estimated based on the average height of the plate. For the MUMPs process and a bushing of 1.5

μm, a value for κ_1 of 2.0 works well for the model presented. Substituting equation (2) into (1), integrating and solving for the voltage as a function of deflection yields equation (3).

$$V = \frac{EI}{w\varepsilon\kappa_1 L^2} \frac{\sqrt{\frac{6w\varepsilon\kappa_1}{dEI}\left[2d\frac{Z_m+3\frac{Kd}{L}}{Z_m(Z_m-d)}-\frac{3\sqrt{d}}{\sqrt{Z_m}}\tanh^{-1}\left(\frac{d}{\sqrt{Z_m d}}\right)-\ln\left(\frac{Z_m-d}{Z_m}\right)\right]}}{\left(\frac{1}{d}+\frac{3K}{Z_m L}\right)\frac{1}{Z_m-d}-\frac{3}{2d\sqrt{Z_m d}}\tanh^{-1}\left(\frac{d}{\sqrt{Z_m d}}\right)-\frac{1}{2d^2}\ln\left(\frac{Z_m-d}{Z_m}\right)}$$

Equation (3) can then be differentiated with respect to d to determine the deflection at snap through given by:

$$d = \frac{45}{8}Z_m\frac{L+4K}{13L+45K} \qquad (4)$$

Substituting (4) back into (3) defines the snap through voltage as a function of SDA plate length. The next step is to derive an expression for the priming voltage as a function of SDA plate length.

Based on the geometry of the priming model the slope at the bushing, Zb/L, can be substituted into a standard formula for simply supported beams[4].

$$\frac{Z_b}{L} = \frac{1}{6LEI}\int_0^L q_p(x)(L-x)\left[L^2-(L-x)^2\right]dx \qquad (5)$$

The electrostatic loading function for the priming model differs from the snap through loading due to the fact that the plate has collapsed against the insulating layer.

$$q_p\left(x\right) = \frac{\varepsilon\kappa_2}{2}\left[\frac{V}{Z_n+\left(\frac{x}{L}\right)^2 Z_b}\right]^2 \qquad (6)$$

Where κ_2 is the effective dielectric for the nitride and air layers as defined by the priming model. The value for κ_2 is larger than κ_1 because the effective dielectric for the priming model is predominantly the value for silicon nitride. A value of 2.7 for κ_2 is a reasonable estimate for a MUMPS SDA with primed deflection. Substituting $q_p(x)$, into (5), integrating and solving for the voltage yields.

$$V_p = \sqrt{\frac{24Z_b^2 EI}{\varepsilon\kappa_2 L^4 w}\left(\frac{3}{\sqrt{Z_n Z_b}}\tan^{-1}\left(\frac{Z_b}{\sqrt{Z_b Z_n}}\right)-\frac{\ln(Z_n+Z_b)}{Z_b}-\frac{2}{Z_n}+\ln Z_n\right)^{-1}} \qquad (7)$$

With functions for both the snap through and priming voltage derived, an optimum plate is defined by plotting both functions verses plate length. Figure 5 plots equation (7) and the result of substituting (4) into (3) for the CRONOS® MUMPs fabrication data listed in table 1.

CRONOS® MUMPs Fabricated SDA Data	
Poly-2 thickness (plate, supports)	1.5 μm
Bushing (dimple + via)	1.5 μm
Support length	29 μm
Support width	3 μm
Nitride layer thickness	0.6 μm
Nitride dielectric constant	5.7
Young's modulus (polysilicon)	169 GPa
SDA width	65 μm

Table 1. SDA model data for CRONOS® MUMPs fabrication.

VOLTAGE RESPONSE VS. SDA PLATE LENGTH
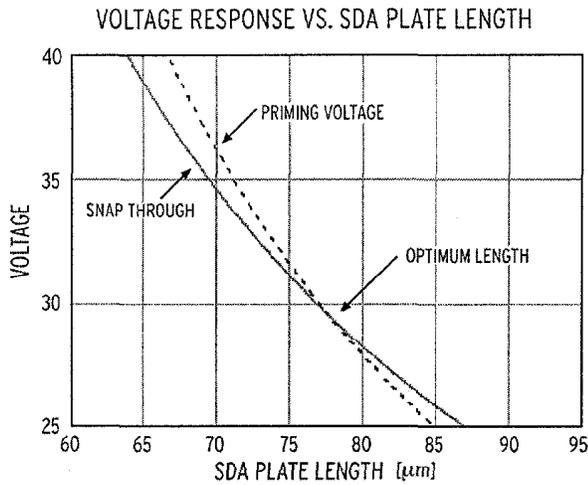
Figure 5. Optimum SDA plate length modeling results.

It should be noted from figure 5 that small errors in either voltage function can have a large effect on the predicted optimum length. Thus, the model should be used to provide an estimate for an initial starting point for experimental verification.

Because the priming deflection is held between stepping cycles, the optimal drive signal for SDAs is an AC signal offset by a DC priming voltage. This drive signal consumes significantly less electrical power than previously published drive signals[1],[3] and results in extended fatigue life and reduced wear to SDA mechanisms. Techniques for experimental evaluation are considered next.
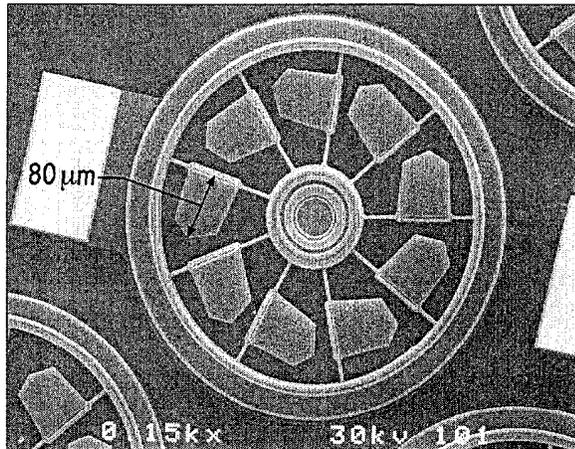
Figure 6. Rotary SDA test stand.

EXPERIMENTAL VERIFICATION

The model was verified using a series of rotary SDA test stands (see figure 6) with incremental changes in plate length to characterize minimum driving voltage with respect to plate length. The rotary motors were fabricated using the CRONOS® MUMPs process corresponding to Table 1, with plate lengths varying from 65 to 80 μm in 5 μm increments. Based on the modeling results, the driving voltage will be minimized at the experimental optimum due to the balancing of drag and friction forces. The experimental results are plotted in figure 7 with the error bars representing the average standard deviation for eight device data points at each plate length. The error is largely due to variations in the release process and storage conditions of individual devices.
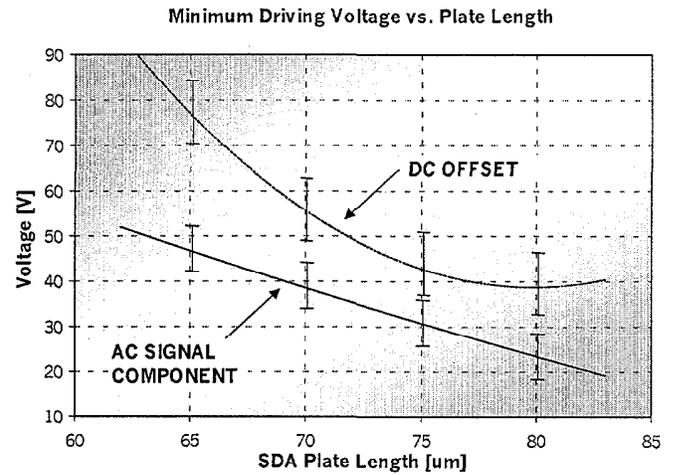
Figure 7. Experimental minimum driving voltage for a 1kHz signal.

The AC component used to drive the SDAs shown in figure 7 was recorded as a zero to peak value with all tests using a 1 kHz signal. It should be noted that the drive signal voltages reach a minimum for a plate length of approximately 78 μm, which agrees well with the modeling prediction for optimum length (see figure 5). Other tests have been performed using linear arrays to determine the output force from SDA arrays as a function of plate length. These tests have revealed a nearly linear decrease in output force from 20 to 9 μN for plate lengths ranging from 55 to 85 μm. Thus for short range and high force SDA applications a shorter plate length than the predicted optimum may be required. However, as can be observed from figure 7, the increase in output force from shorter plate lengths comes at a cost of higher drive voltages, and in turn reduced fatigue life due to the adverse wear affects of the higher voltages.

TETHERED SDA ROBOTS

With the design of the SDA optimized, the actuator can be arrayed in large numbers to provide forces capable of moving chip sized and larger components. The SDA robot (Scratchuator) shown in figure 8 was made from an array of 188 SDA and is tethered by a 25 μm gold wire. The gold wire is bonded to the Scratchuator before release in an HF bath which separates the device from the substrate it was fabricated on. The gold bumps on the front end of the Scratchuator are used as a bumper for pushing objects or as a means to bond the device to a larger ceramic stage
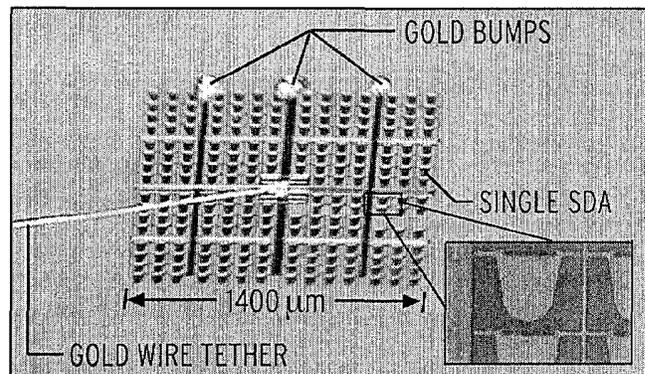
Figure 8. Wire tethered Scratchuator with array of 188 SDA.
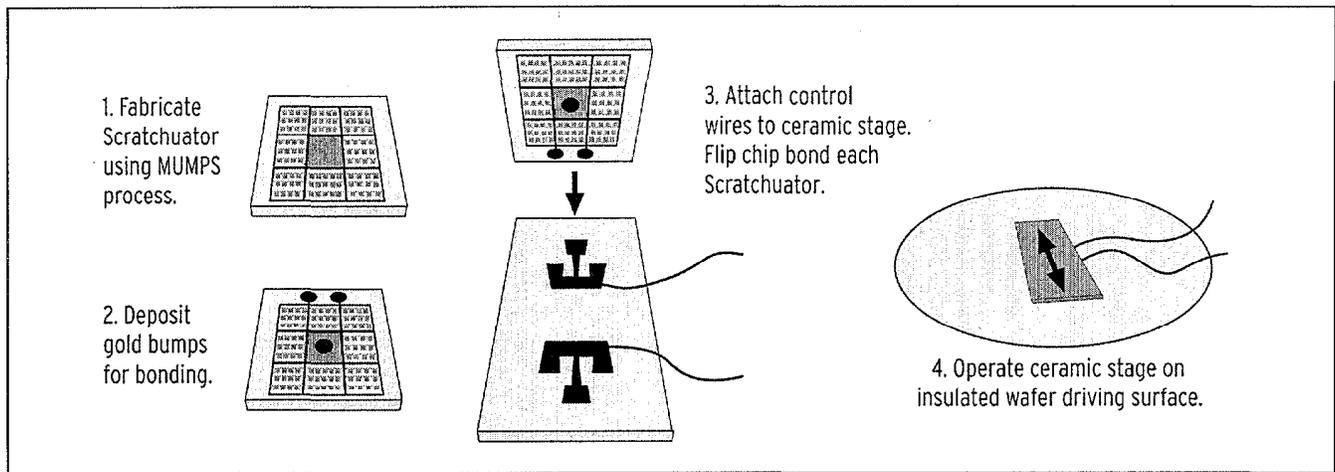
Figure 9. Assembly process for flip chip bonding of SDA robots to a ceramic stage.

as shown in figure 10. An SDA array is required for each desired direction of motion when positioning objects. Assembly techniques for the ceramic stage are shown in figure 9.

Because Scratchuators can be attached uniformly to the bottom of a ceramic substrate, large forces and multiple directions of motion are expected from a level stage designed to withstand the normal forces required for use as a flip chip bonding aligner. An silicon wafer can be used as a large driving surface for the Scratchuators as it has a smooth uniform surface with a ground plane and insulating layer.

Preliminary testing has shown a single Scratchuator (figure 8) is capable of pushing a 2x2x0.5mm silicon chip over 8 millimeters while being tethered by a 25 µm gold wire even after nearly half the SDAs from the array have been damaged. Failure mechanisms are related to the stiction of individual SDA to the driving surface which causes fracturing when the array is repositioned. Proper designs require a strong central frame that minimizes stress concentrations at the supports of the SDA.
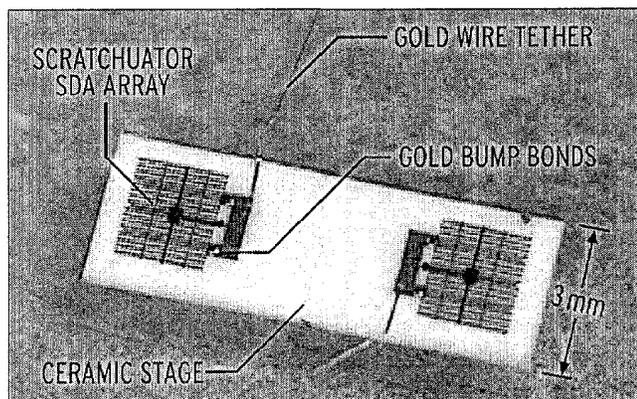
Figure 10. Bottom view of ceramic stage with two Scratchuators gold bump bonded. The gold wire tethers provide electrical power to each Scratchuator allowing controlled motion in two directions when driven across an insulated wafer surface.

CONCLUSIONS

A novel surface micromachined robot, capable of pushing a 2x2x0.5 mm silicon chip over 8 millimeters while being powered by a 25 µm gold wire tether has been developed. The micro robot uses wire bonding and flip chip techniques for assembly and is

created from large arrays of scratch drive actuators (SDA). A model has been derived to define optimum SDA plate design for any fabrication process. The robot is transferred to the large driving surface of an insulated wafer, where assembly line positioning tasks can be performed. The driving signal consists of an AC stepping component with a DC offset used to hold the SDA in a primed deflection between steps. Because the step size of the SDA is approximately 30 nanometers, similar placement precision is expected for positioning of components. The SDA robot (Scratchuator) demonstrates the potential for low cost nanometer alignment of components for flip chip assembly.

REFERENCES

1. G. D. Su, L. Y. Lin and M. C. Wu, "Single-chip femtosecond autocorrelator realized by surface-micromachined integrated optics," in *Proc. Conf. Lasers and Electro-Optics (CLEO)*, Baltimore, MD, 1997, paper CWL2.

2. L.Fan, M.C. Wu, K. Choquette, and M.H. Crawford, "Self-assembled microactuated XYZ stages for optical scanning and alignment," in *Proc. 1997 Int. Conf. Solid-State Sensors and Actuators (TRANSDUCERS 97)*, 1997, paper 2A2.01.

3. G. Bourbon, P. Minotti, P. Langlet, T. Masuzawa and H. Fujita, "Three dimensional active microcatheter combining shape memory alloy actuators and direct-drive tubular electrostatic micromotors," *SPIE Conference on Micromachined Devices and Components IV*, Santa Clara, California, September 1998, SPIE Vol. 3514 pp. 147-158.

4. R. B. Hopkins, "Design analysis of shafts and beams," McGraw-Hill, 1970.

5. K. E. Petersen, "Dynamic micromechanics on silicon: techniques and devices," *IEEE Transactions on Electron Devices*, Vol ED-25, No. 10 pp. 1241-1250, October 1978.

ELECTRO-THERMAL ACTUATORS USING OPTIMIZED COMPLIANT MICROTRANSMISSIONS AS RECTILINEAR MOTION AMPLIFIERS

Larry L. Chu*, Joel A. Hetrick**, and Yogesh B. Gianchandani*[1]
*Department of Electrical and Computer Engineering
**Department of Mechanical Engineering
University of Wisconsin, Madison

ABSTRACT

This paper reports on the use of microtransmissions to tailor the force-displacement relationship of rectilinear electro-thermal actuators. A two-stage design approach is described for the microtransmissions, which effectively automates the synthesis procedure. It produces detailed designs of high performance compliant microtransmissions given magnification the requirements while considering buckling and minimum width (fabrication) limitations. Measurements of mechanisms fabricated from electroplated Ni on Si substrates demonstrate displacement amplification factors of 3-14, and match theoretical predictions within 10%. Higher amplification factors can be achieved by changing certain dimensional parameters.

I. INTRODUCTION

Electrothermal actuators are attractive for applications that benefit from high output forces, low actuation voltages, and electrically conductive structural materials. Simple and cascaded bent-beam electrothermal actuators have been used for rectilinear motion parallel to the substrate plane [1,2], producing maximum displacements and maximum blocking forces in the range of 8 μm and 2.5 mN, respectively, for a silicon device of 1000 μm length, 6 μm width, 4.5 μm thickness, and 0.2 rad. bending angle, operating at 400°C. However, in many applications a higher displacement is necessary, a smaller force is adequate, and the operating speed of incremental mechanisms like inchworms is insufficient. In such cases, the force-displacement compromises can be tailored with compliant mechanisms that serve as backlash-free microtransmissions.

Compliant mechanisms (often called flexures) are structures that exploit elastic deformation to effectively transmit forces and motions. The advantages of compliant mechanisms include the elimination of friction, wear and backlash common in conventional mechanical joints. Due to their monolithic construction, compliant mechanisms are also much easier to fabricate at the micro-scale making them very attractive for MEMS applications.

As shown in Fig. 1, zero-backlash compliant mechanisms are useful as micro-transmissions, effectively altering the force-displacement characteristics of unconventional actuation sources such as smart materials and electro-thermal actuation sources [3]. In this paper, compliant mechanisms are used to provide efficient rectilinear motion while amplifying the strains from electro-thermal bent-beam actuators. The challenge in designing a high performance compliant mechanism is to optimize the structural topology, size and geometry to efficiently transform the forces and displacements of an input actuation source to drive an output load.

While synthesis of compliant mechanisms [4-6] and combined electro-thermal-compliant actuators [7,8] has been studied by past researchers, the approach developed here can precisely control the force-displacement trade-off of the actuator at levels of extreme magnification (up to 20X displacement amplification depending on fabrication limitations).

II. SYNTHESIS OF COMPLIANT TRANSMISSIONS

The design procedure is broken into topology synthesis and dimensional synthesis. Each stage is posed as a structural optimization problem, where the objective is to maximize the blocking force and unloaded output displacement given all relevant design constraints. Structural analysis is performed using structural (truss and beam) elements and linear, static finite element theory. For both topology synthesis and dimensional synthesis, the finite element equilibrium analysis and the constrained optimization problem are solved within MATLAB™. Optimization is performed using MATLAB's Sequential Quadratic Programming (SQP) algorithm. Design derivatives are calculated using the adjoint variable method and by directly differentiating the stiffness matrix with respect to size and node position design variables.

The magnification factor, M, is expressed as a function of the slope change between the input and output force-displacement characteristics:

$$M = \sqrt{\frac{d_{\max2} \cdot F_{block1}}{d_{\max1} \cdot F_{block2}}} \qquad (1)$$

where d_{max} denotes displacement in the absence of external loading forces.

While slope change is one method for determining the magnification factor, other methods such as specifying the ratio of unloaded output displacement or blocking force can also be readily accommodated. The generic structural optimization problem (for both topology and dimensional synthesis) is posed as follows:

$$Maximize \quad F_{block2} \cdot d_{\max2} - P\left(\frac{d_{\max2}}{F_{block2}} - M^2 \frac{d_{\max1}}{F_{block1}}\right)^2$$

such that

$$\sigma_{axial} - \sigma_{buckling} \leq 0; \quad V - V_{allow} \leq 0$$

$$h_{\min} \leq h_i \leq h_{\max}; \quad X_{\min} \leq X_j \leq X_{\max}$$

$$Y_{\min} \leq Y_k \leq Y_{\max}$$

$$(2)$$

[1]*Corresponding author:* 1415 Engineering Drive, Madison, WI 53706-1691; Tel: (608) 262-2233, Fax: 262-1267, E-mail: yogesh@engr.wisc.edu

The required magnification factor, M, is enforced by penalizing the objective function by the penalty factor, P. Buckling is enforced by limiting the maximum axial stress to $\sigma_{buckling}$ based on Euler's buckling criteria, modified for end conditions (during dimensional synthesis). The total volume constraint is enforced by V_{allow}. Design variables include element width and X and Y node position, denoted by h_i, X_j, and Y_k respectively. These are given local upper and lower bounds which serve to limit the search space and also to enforce fabrication limitations (during dimensional synthesis).

Topology Synthesis

To establish the optimal topology of a mechanism, the design domain is discretized using a network of truss elements, defined as a ground structure. The topology optimization process seeks to resize each element thickness until the performance of the structure is maximized. Upon completion, elements which converge to the lower bound are removed yielding the final mechanism topology. The topology synthesis approach described here utilizes a modular ground structure where nodes are allowed to change location within 'wandering limits'. Research has shown that this technique produces good topology designs that accurately satisfy magnification requirements [9].

Figure 2 illustrates the example compliant mechanism design problem. Given a positive actuator input force, the output is designed to move in the negative vertical direction. Taking advantage of symmetry allows only half of the topology to be designed, which lowers computational requirements. For the inverter example, the modulus was set to 1000, the width was set to 1, the total volume constraint was set to 2, and the lower element bound was set at 1E-4. To achieve the magnification factor of 20, the penalty multiplier was set to 1. Figure 3 shows the floating-node ground structure, discretized using 25 nodes and 72 elements; 23 nodes were activated and given wandering ranges resulting in 107 design variables (wandering ranges indicated by gray areas). Figure 4 indicates the optimal topology of the compliant mechanism amplifier with elements converging to the lower bound removed.

Dimensional Synthesis

To proceed with dimensional synthesis, each segment of the topology is re-meshed using arrays of beam elements; the number of elements along an array is directly controlled by the designer. Exact specifications for the material modulus, out-of-plane thickness, input actuator characteristics, and desired magnification are input into the optimization routine. The designer may choose to activate end nodes of topology segments, allowing for geometric variation. Intermediate nodes along the length of an activated segment are adjusted by maintaining a collinear equidistant relationship.

Figure 5 shows the inverter example from the topology synthesis remeshed for the secondary dimensional synthesis stage. The mechanism is to be fabricated using electroplated nickel as the structural material. Constraints within the fabrication sequence limit the minimum thickness and width of the structure to about 15 μm. The mechanism is designed to be driven using a nickel bent-beam actuator ($H = W = 15$ μm, $L = 1000$ μm) operating at a temperature difference of 200 °C with respect to the substrate ($F_{block} = 25000$ μN, $d_{max} = 6.8$ μm). The thermal expansion coefficient of the metal is approximately 13.5 ppm/K [10]. Figure 6 shows the optimized structure ($M = 20$, $P = 1$, $V_{allow} = 1 \times 10^6$ μm^3). This particular design offers a blocking force of 470 μN (assuming a Young's modulus of 100 GPa) and a free displacement of 51 μm. Because of the minimum width limitation, substantial energy is spent deforming the compliant transmission (transmission efficiency is 14.1 percent). Larger blocking force and free displacement can be achieved by fabricating more flexible structures with thinner beam widths (within buckling limitations).

III. EXPERIMENTAL RESULTS

Thermal actuators with microtransmissions were fabricated on silicon wafers insulated with thermal oxide and LPCVD nitride. A 2.2 μm thick sputtered Ti sacrificial layer was patterned and covered with a Cr/Ni seed layer. The devices were electroplated into a photoresist mold from a nickel sulphamate solution. At 55°C temperature, the plating current was adjusted to provide a structural layer thickness of 16 μm in 120 min. Devices with other thicknesses were also electroplated at the same rate. The photoresist mold was subsequently stripped and the sacrificial material etched away. Some samples were then covered with a thin layer of gold by electroless plating to provide an inert coating [10]. Following this, a coating of self-assembled monolayers was applied. Optical micrographs of a fabricated structures are shown in Fig. 7.

The dimensional parameters of designs that were fabricated are defined in Fig. 8 and listed in Table I. All fabricated devices were tested electrically. A sample set of measurements is plotted in Fig. 9, showing the displacement at the input and output of the microtransmission. The measurements were taken by a calibrated visual method, with an uncertainty of <0.5 μm. The measured displacement amplification, A, for this device was 3.1, which is very close to theoretical prediction of 3.2. Electrothermal tests of a device that was 16 μm thick but similar to design 2a in all other respects showed an output displacement of 9.7 μm at a drive voltage of 0.6 V. Due to the high A offered by devices 2a and 3, the displacement at the input to the microtransmission was not measurable at the peak operating temperatures. In order to verify theoretically predicted A of these devices, the microtransmissions were mechanically actuated with a probe tip. The results matched expectations within 10%. Sample measurements for device 3 are shown in Fig. 10.

Note that the output blocking forces are linearly related to the Young's modulus of the structural material. When electroplated metals are used, there is significant variability in this parameter [10]. Values ranging from 35 Gpa to 150 GPa hav been reported for electroplated nickel, whereas the Young's modulus of bulk nickel is 208 GPa.

IV. CONCLUSIONS

Microtransmissions have been used to tailor the force-displacement relationships of electro-thermal actuators. The microtransmissions are flexural elements that provide backlash-free operation and rectilinear input and output displacements that are suitable for use with bent-beam actuators and convenient for numerous applications. Topology synthesis and dimensional synthesis are demonstrated to effectively automate the design of compliant mechanisms obtaining extreme magnification

values. The two-stage process produces detailed designs given actuator and material input parameters while considering magnification, buckling, and minimum width constraints.

The resulting electro-thermal actuators with integrated microtransmissions were fabricated by electroplating nickel on silicon substrates. Electrical and mechanical tests confirmed the model predicitons, with the measured multiplication factor matching the theoretical value to within 10%. Displacement amplification factors of 3-14 have been demonstrated. However, these are constrained by the modest fabrication limits, and higher performance can be achieved.

REFERENCES

[1] L. Que, J.-S. Park, Y.B. Gianchandani, "Bent-Beam Electro-Thermal Actuators for High Force Applications," *IEEE Intl. Conf. on Micro Electro Mechanical Systems (MEMS99)*, Orlando, Florida, Jan. 1999

[2] J.-S. Park, L.L Chu, E. Siwapornsathain, A.D. Oliver, Y.B. Gianchandani, "Long Throw and Rotary Output Electro-Thermal Actuators Based on Bent-Beam Suspensions," *IEEE Intl. Conf. on Micro Electro Mechanical Systems (MEMS00)*, Miyazaki, Japan, Jan. 2000

[3] S. Kota, J. Hetrick, Z. Li, L. Saggere, "Tailoring Unconventional actuators using compliant transmissions: Design Methods and Applications," *IEEE/ASME Transactions on Mechatronics*, 4(4), 1999, pp. 396-408

[4] S. Nishiwaki, M. Frecker, M. Seungjae, N. Kikuchi, "Topology Optimization of Compliant Mechanisms Using the Homogenization Method,", *Intl. Journal for Numerical Methods in Engineering*, 42(3), 1998, pp. 535-559

[5] O. Sigmund, "Tailoring Materials with Prescribed Elastic Properties," *Mechanics of Materials*, v. 20, 1995, pp. 351-368

[6] G. Ananthasuresh, S. Kota, N. Kikuchi, "Strategies for Systematic Synthesis of Compliant MEMS," *Dynamic Systems and Control*, DSC-Vol. 55-2, 1994 ASME Winter Annual Meeting, Chicago, IL, Nov. 1994, pp. 677-686

[7] T. Moulton, G. Ananthasuresh, "Micromechanical Devices with Embedded Electro-Thermal-Compliant Actuation," *MEMS-Vol. 1*, ASME International Mechanical Engineering Conference and Exposition, MEMS, Nov. 1999, Nashville, Tenessee, pp. 553-560

[8] J. Jonsmann, O. Sigmund, S. Bouwstra, "Compliant Thermal Microactuators," *Sensors and Actuators (A)*, Vol. 76, 1999, pp. 463-469

[9] J. Hetrick, *A Unified Approach for Topological and Dimensional Synthesis of Compliant Mechanisms*, Ph.D. Thesis, University of Michigan, Ann Arbor, Michigan, 1999

[10] L. Chu, L. Que, Y. Gianchandani, "Temperature Coefficients of Material Properties Using Differential Capacitive Strain Sensors," *Proc., Solid-State Sensors and Actuators Workshop*, Hilton Head, S.C., June 2000

Table I: Summary of simulated and measured results. Dimensions are defined in Fig. 8. H is device thickness; d_{max} and F_{block} are calculated at the output, assuming a 200°C temp. difference with respect to the substrate, and the Young's modulus is 100 GPa.

Dev #	L0 μm	L1 μm	W μm	H μm	d_{max} μm	F_{block} μN	A_{calc}	A_{meas}
1	720	1500	70	35			3.2	3.1
2a	1000	750	10.8	9.7	31.5	244	11.1	10.6
3	1000	1000	10.8	16.0	134	394	14.1	12.9
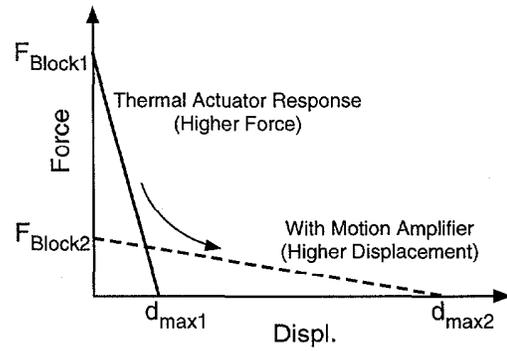
Fig. 1: A compliant micro-transmission provides rectilinear motion and allows the force-displacement trade-off to be tailored for an application.
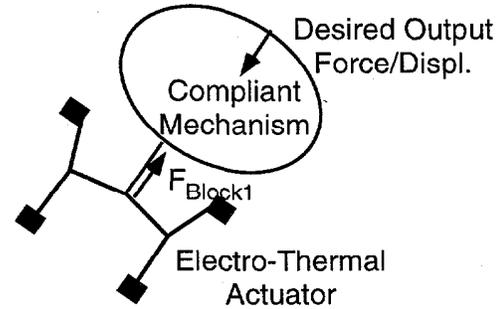
Fig. 2: The topology design problem for a compliant mechanism rectilinear amplifier.
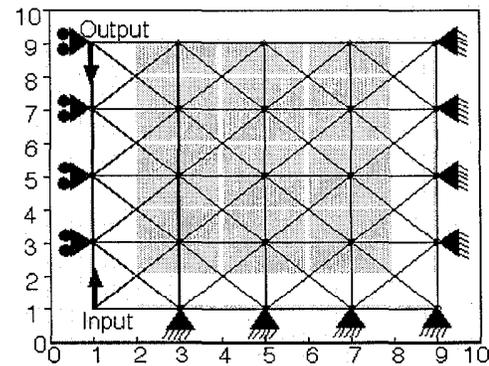
Fig. 3: Half of the design domain is meshed using truss elements (gray areas denote wandering regions).
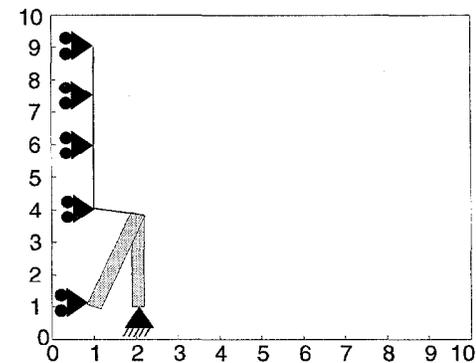
Fig. 4: The optimized mechanism topology.
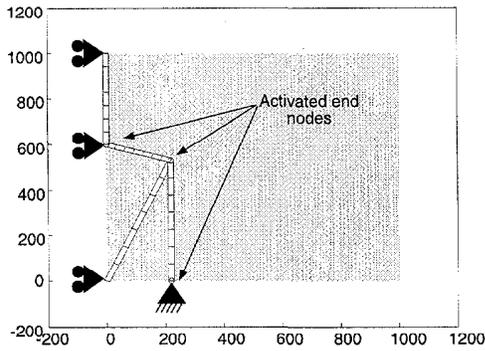
Fig. 5: The dimensional synthesis problem with the structure modeled using beam elements (gray areas denote end node wandering ranges).
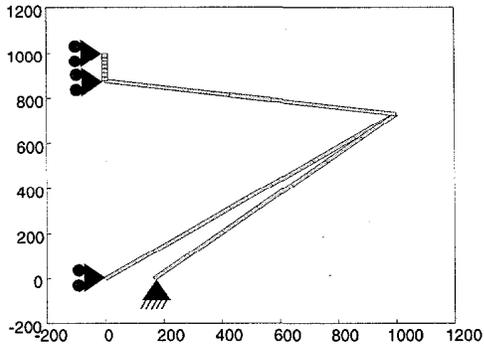
Fig. 6: Half of the finalized compliant mechanism.
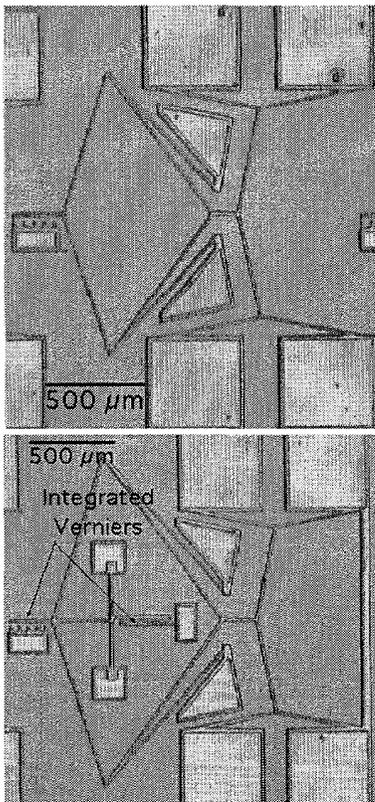
Fig. 7: Optical micrographs of cascaded bent-beam actuators with attached microtransmissions. The devices were fabricated from Ni electroplated on Si substrates.
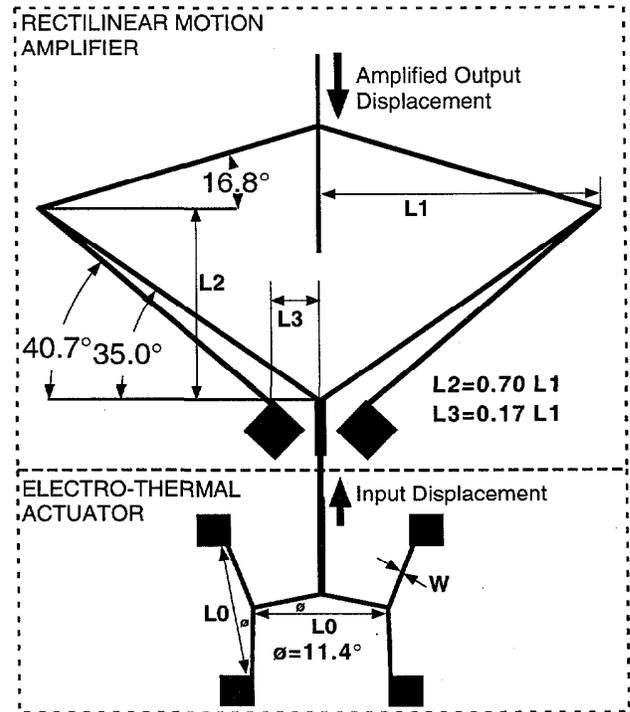
Fig. 8: Schematic of the rectilinear microtransmission and cascaded actuator showing the dimensional variables.
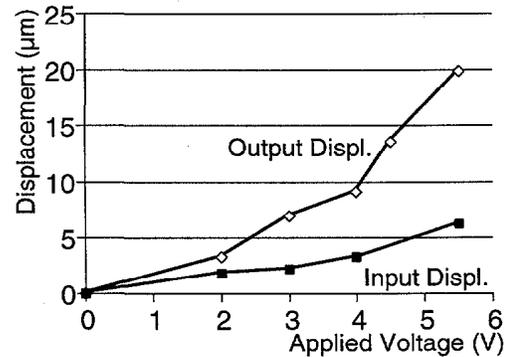
Fig. 9: Electro-thermally generated displacements measured at the input and output of the transmission of device 1, with dimensional parameters as in Table I.
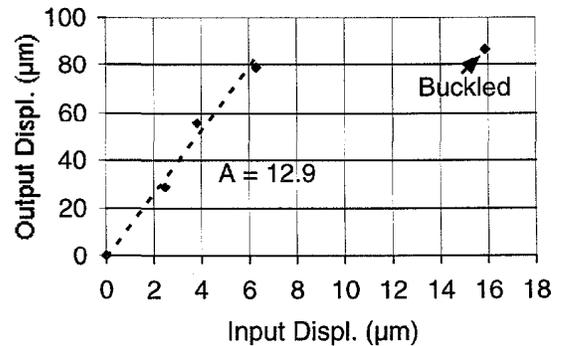
Fig. 10: Measured ratio of input to output displacement for the microtransmission of device 3 (Table I). This measurement was taken mechanically because of operating temperature limitations.

HYDROGEL MICROVALVES FABRICATED USING *IN-SITU* POLYMERIZATION

Robin H. Liu*, Qing Yu*, Joseph M. Bauer*, Jeffrey S. Moore*, David J. Beebe*[†]

*The Beckman Institute for Advanced Science and Technology
University of Illinois at Urbana-Champaign
Urbana, Illinios 61801, USA
[†]Department of Biomedical Engineering, University of Wisconsin-Madison,
Madison, Wisconsin 53706, USA (dbeebe@engr.wisc.edu)

ABSTRACT

Hydrogel microvalves, which consist of a single smart hydrogel material that undergoes a volume change in response to changes in local pH, have been developed. The hydrogel components were fabricated inside microchannels using *in-situ* photopolymerization. Two 2D shut-off valves and one 3D hybrid valve were fabricated and tested. Flow and pressure performances were studied. Analytical modeling was performed on the 3D valve. The behavior of hydrogel valves in electrokinetic flow was also investigated. Hydrogel microvalves have a number of advantages over conventional microvalves, including relatively simple fabrication, no external power requirement, no integrated electronics and large displacement.

Keywords: *in-situ* polymerization, microvalves, hydrogel

INTRODUCTION

Most conventional active microvalves couple a flexible diaphragm to a thermopneumatic, piezoelectric, electrostatic, electromagnetic, or bimetallic actuator [1]. The scaling of these actuation forces with respect to the size of the devices is often unfavorable to micro-actuators [2]. Some actuator components are made heavy and large (mm size) in order to produce enough force to regulate a microscale fluid. The displacement of conventional diaphragms (e.g., silicon, silicon nitride) is often limited to tens of microns or less. Many actuators use resistive heating (thermopneumatic), high voltage (electrostatic), and bulky magnets (electromagnetic), which often require high power consumption. Moreover, the integration of many microvalves composed of electronic components into complex microfluidic systems has proven to be non-trivial using traditional approaches, such as modular or lithographic methods [3]. In general, most conventional microvalves are complex and costly, and integrating them into micro-systems is challenging

We have recently developed an *in-situ* liquid phase polymerization process to integrate functional materials into microfluidic channels to construct microvalves [4, 5]. The microvalves consist of a stimuli-responsive hydrogel material that undergoes a volume change in response to changes in local pH. Stimuli-responsive hydrogels have a significant advantage over conventional microfluidic actuators due to their ability to undergo abrupt volume changes in response to the surrounding environment without the requirement of an external power source. Stimuli-responsive hydrogels are crosslinked polymer networks that can undergo discontinuous volume changes in response to chemical signals in their surrounding environment [6]. Among stimuli-

responsive hydrogels, pH sensitive hydrogels are one of the most commonly studied. As the environmental pH varies, the hydrogel networks change from neutral to charged because of the protonation of amine groups or the deprotonation of acid groups. The hydrogels thus undergo a volume transition from the collapsed state to the expanded state due to osmotic pressure exerted by mobile counter-ions neutralizing the network charges. In this paper, we report an extension of our previous work and detailed characterizations of the microvalves with optimized design.

ANALYTICAL MODELING

Among the hydrogel microvalves constructed, the 3D hybrid valve is the only one that couples an axisymmetric flexible poly(dimethylsiloxane) (PDMS) membrane to a hydrogel actuator. An analytical model was used to optimize the design of the 3D valve. The objective is to determine the relationship between the membrane deflection (z), the membrane thickness (t), and the diameter of the membrane ($2R$), for a given pressure applied by a hydrogel actuator (Fig.1). Consider a thin circular membrane of radius R that is clamped along the perimeter and subjected to a net uniform pressure p exerted by the hydrogel actuator. Since the deflection of the membrane z (hundreds of microns) is an order of magnitude larger than its thickness t (tens of microns), and the lateral deflection w is more than a few tenths of thickness t, the "membrane" theory was employed instead of the "thin plate" theory [7, 8]. Bending moments are assumed to be negligibly small. The membrane force that is acting tangent to the deflected middle surface due to tension is relatively large, and it is assumed to be uniform across the thickness of the membrane. The model assumes that the membrane material is isotropic and linearly elastic at small strain.

Let T denote the tension per unit length of the membrane. The net force on the membrane perpendicular to a section of radius due to p is $\pi r^2 p$ (see r in Fig.1). Note that h in Fig.1 is assumed to be large enough so that the membrane does not touch the orifice. The restoring force due to tension in the membrane is $2\pi r T \sin\theta$. Force equilibrium requires: $2\pi r T \sin\theta + \pi r^2 p = 0$, where $\sin\theta = dz/dr$. The solution is given by

$$z = -\frac{p}{4T}r^2$$

(1)

The strain along the radial direction r is

$$\varepsilon_r = (\int_0^R \sqrt{1+z'^2}\,dr - R)/R \approx \frac{p^2 R^2}{24 T^2}$$

Together with Hooke's law $\varepsilon_r = \sigma_r/E^*$, where $E^* = E/(1-v^2)$, we have

Solid-State Sensors, Actuators, and Microsystems Workshop
Hilton Head Island, South Carolina, June 4-8, 2000

$$T = \sigma_r t = \frac{\varepsilon_r E}{1 - v^2} t = (\frac{1}{24} \frac{E}{1 - v^2} t p^2 R^2)^{1/3}$$

which together with Eq. (1), gives

$$z_{max}\big|_{r=R} = -\frac{1}{4}[\frac{24(1 - v^2)p}{E}]^{1/3} \frac{R^{4/3}}{t^{1/3}}$$ (2)

or

$$z_{max} \propto (\frac{R^4}{t})^{1/3}$$
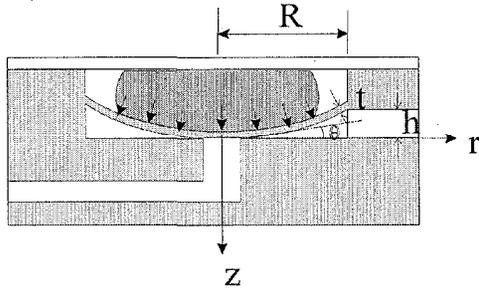
Fig. 1 Schematic of the 3D hybrid valve.

Eq. (2) indicates that the maximum deflection of the membrane increases only slightly when the thickness of the membrane is decreased, while a large change results from increased membrane radius. In order to obtain large z_{max}, we can either make a thinner membrane or a membrane with larger diameter.

The pressure p generated by the hydrogel actuator is unknown. But if z_{max} can be measured experimentally, then the pressure can be determined by Eq. (2):

$$p = -\frac{8z_{max}^3 Et}{3(1 - v^2)R^4}$$ (3)

EXPERIMENTS

Two 2D shut-off valves and one 3D hybrid valve were fabricated. The fabrication process consists of two steps: fabrication of microchannels and in-situ polymerization. For the 2D valves, microchannels are made of EPON (Nano XP SU-8 50, Microchem Corp., Newton, MA). A transparent adhesive tape (regular packaging tape) was used to seal the 200 μm deep EPON microchannels. The microchannels of the 3D valve were constructed using multi-layer PDMS [9]. A glass cover slip served as the top of the channel. After the fabrication of microchannels, an in-situ polymerization that introduces active hydrogel components into microchannels via direct photopatterning of a liquid phase was used to complete the microvalve construction. The liquid phase mixture that consists of monomers (acrylic acid and 2-hydroxyethyl methacrylate in a 1:4 mol. ratio), a crosslinker (ethylene glycol dimethacrylate, 1 wt %) and a photoinitiator (2,2-dimethoxy phenylacetophenone, 3 wt %) was first flowed into the microchannel. After the liquid reached a quiescent state in the microchannel, it was then exposed to UV light through a photomask placed on top of the channels. When the polymerization was finished, the channel was flushed with deionized (DI) water to remove unpolymerized liquid. The resulting hydrogel expands in basic solutions and collapses in acidic solutions, with the volume

transition at pH 4.5 ~ 7.0. The details of microvalve fabrication process can be found elsewhere [4, 5].

Pressure drop measurements using a pressure transducer (Validyne Engineering Co., Northridge, CA) with a pressure range of 0 to 125 psi were taken over the regulated channel length to characterize the hydrogel microvalves. In order to verify fluid flow in the various device designs, a variety of flow visualization techniques such as particle tracking, dyes, and air bubbles were used. To view the membrane deflection of the 3D hybrid microvalve, the cross-section images of the valve were obtained by cutting the PDMS using a razor blade without damaging the valve.

RESULTS AND DISCUSSION

A shut-off valve, consisting of a single hydrogel component that can sense the chemical environment (pH) in one channel and regulate the flow in an adjacent channel, was constructed as shown in Fig.2. The hydrogel structure was polymerized in the gap (300 μm wide) between two parallel microchannels. The sensing channel (600 μm wide) was constructed wider than the regulated channel (300 μm wide) in order to keep the sensing channel open when the regulated channel is completely shut off. Solutions of different pH values were pumped through the sensing channel while DI water flowed through the regulated channel. As the pH in the sensing channel was changed from 3.8 to 5.7, the shut-off response time of the hydrogel is slow (~ 160 min) due to the slow migration of ions across the hydrogel. Cross talk between two channels is another problem that may limit some practical applications.
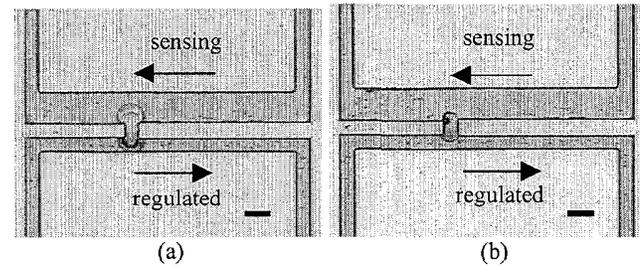
Fig. 2 Micrographs of a 2D hydrogel shut-off valve. (a) and (b) illustrate flow regulation in the regulated channel as the pH is changed from 3.8 to 5.7 in the sensing channel.

Since diffusion is the rate-limiting factor governing the swelling process of hydrogels [10], the time response of the volume change approximately follows the square of the hydrogel dimension. In order to improve response time, smaller hydrogels with a larger surface area exposed to surrounding chemical environment are preferred. However, it was found that small hydrogel objects tend to buckle or migrate during a volume change. In order to fabricate stable hydrogel actuators with fast response times, a hydrogel jacket (50 μm thick) was polymerized around a prefabricated circular EPON post (50 μm diameter) in the microchannel (Fig.3). When a pH 11 solution was flowed into a side branch, the hydrogel jacket expanded and closed the regulated channel (pH 7). If a pH 2 solution was flowed into the other side branch instead, the contracted hydrogel allowed the fluid in the regulated channel to flow down the waste channel. The post provides a robust support and also improves time response due to the short diffusion path of the hydrogel jackets surrounding the posts. The response time of the post valve design is 12 seconds

223

(Fig. 4). This integration of hydrogels into microfluidic systems provides the scaling necessary to overcome the primary drawback (slow time response) of hydrogels. The maximum differential pressure the valve can sustain is 56.5 psi when the valve is closed. At least 200 operating cycles were achieved over a pH range of 2 to 11 without failure.

(a)   (b)
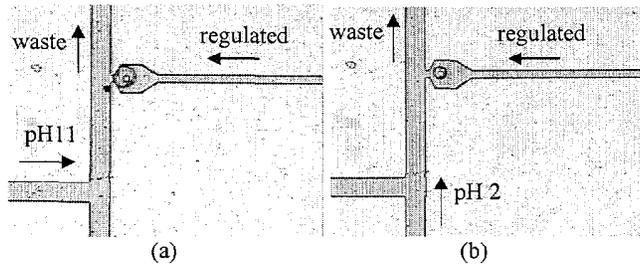
Fig. 3 (a) Micrograph (2x) of the hydrogel jackets blocking the regulated channel (pH 7) in their expanded state in a pH 11 solution. (b) Micrograph showing the contracted hydrogels allowing the fluid to flow down the side branch.
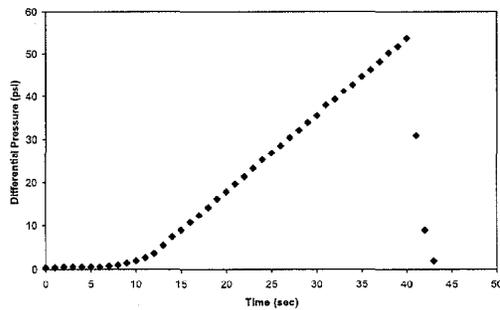
Fig.4 Pressure drop measurement of the 2D post valve illustrates the time response of the valve during swelling process. The transition to a linear response (at approx. 12 sec) indicates valve closure. At 41 sec, the pressure dropped substantially due to the failure of an external tubing connection.

The post hydrogel valve was also tested in electrokinetic flow (EOF) experiments. As shown in Fig. 5, the regulated channel was first filled with a buffer solution (pH 7) using a syringe pump. The hydrogel jacket expanded and closed the channel. A voltage of 500 V was then applied to the ends of the channel. The electrical field resulted in a rapid additional expansion (~2 sec) of the hydrogel jacket. The resulting force was large enough that it delaminated the EPON post from the channel bottom. Although this phenomenon is still being investigated, the preliminary result demonstrates a potential use of hydrogel components as actuators in EOF systems.
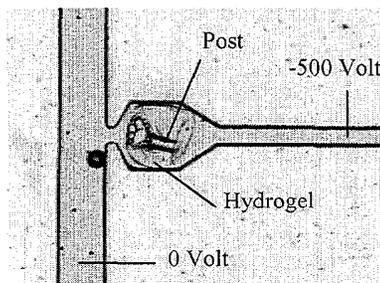
Fig.5 Micrograph (4x) illustrates a post hydrogel microvalve in EOF.

Since hydrogels are generally porous, ions can diffuse through a hydrogel object, resulting in a leaking or cross talk problem in the above 2D hydrogel valves. To address this problem, the 3D hybrid microvalve couples a flexible PDMS membrane to a hydrogel actuator (Fig.6). The membrane acts to physically separate the sensing and regulated flows. The hydrogel actuator can sense the pH in the upper channel and expands or contracts as the surrounding pH is changed. The force associated with these volumetric changes is sufficient to deform the membrane and consequently control the flow in the lower channel. The membrane can deform to completely block the orifice with a displacement of up to 150 μm. No leakage was observed. We previously reported a maximum differential pressure P_{max} of 7.5 psi if the lower channel height h is 150 μm [5]. If h decreases while other parameters are kept unchanged (i.e., the diameter of the hydrogel actuator is 250 μm, the PDMS membrane has a thickness of 50 μm and a 700 μm diameter), P_{max} goes up as shown in Table 1. In other words, h should be kept as small as possible in order to achieve high P_{max}. However, it was found that if h is less than 70 μm, the PDMS membrane is prone to stick to the bottom of the lower channel due to a surface force during the PDMS bonding process, rendering the orifice sealed permanently. Therefore, it is recommended that the ratio of h to the diameter of the membrane be at least 0.1 if high P_{max} is required. The maximum differential pressure P_{max} was measured to be 26.7 psi with a height h of 75 μm (Fig. 6).

(a)

(b)   (c)

(d)   (e)
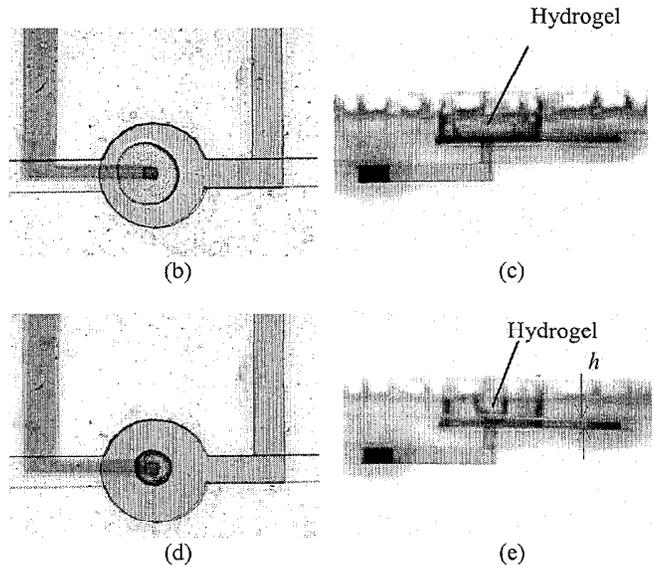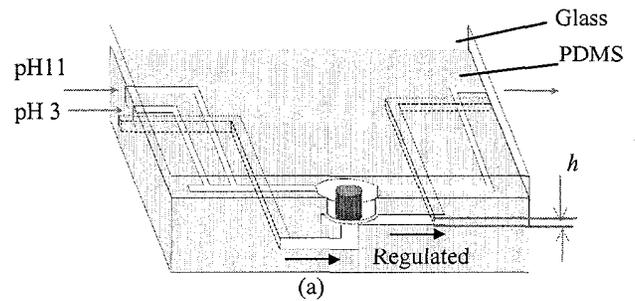
Fig. 6 (a) Schematic of a 3D hybrid microvalve. (b), (c), The top view and the cross section, respectively, of the valve when a pH 11 solution is flowing in the upper channel. The hydrogel actuator expands and deforms the PDMS membrane to seal the orifice in the regulated channel. (d), (e) show that the hydrogel contracts at a pH 3 solution and reopens the regulated channel.

Table 1 The effect of the lower channel height h on the maximum differential pressure P_{max} of the 3D hybrid microvalve.

h (μm)	P_{max} (psi)
75	26.7
100	18.4
125	11.3
150	7.5

Moreover, small h gives rise to rapid closure of the orifice. The shut-off time response of the 3D hybrid valve with h of 75 μm is 19 sec, compared to 50 sec if h is 150 μm. At least 200 operating cycles were achieved over a pH range of 3 to 11 without failure.

From the experiment, it was found that z_{max} is 185 μm for a hydrogel actuator (250 μm diameter and 200 μm high) at pH 11 deforming a PDMS membrane of 50 μm thickness and 700 μm diameter (note: the PDMS membrane does not touch the orifice because h is 200 μm here). Therefore, Eq.(3) is used to determine the pressure generated by the hydrogel actuator. Given that the Young's modulus E of PDMS (1:10 mixing ratio) is 750 kPa and the Poisson ratio is 0.5 [11], the pressure p is approximately 56.26 kPa, which corresponds to a force of 22 mN.

Due to the unique ability to convert chemical energy into mechanical work, the pH-sensitive hydrogels exhibit both sensing and actuating functions simultaneously. This approach avoids the need for integrated electrical connections for microvalve operation. Moreover, the fabrication process of hydrogel microvalves is less complicated than that of most conventional microvalves which involve various bulk processes (e.g., bulk etching of Si wafers) and surface processes (e.g., thin film processes). This method avoids the complicated integration process required to construct conventional microvalves inside complex microfluidic systems. Although the time response of the hydrogel microvalves shown above is relatively slow compared to most conventional microvalve (~ms), it is believed that changes in hydrogel chemistry and valve design can further improve the time response and make hydrogel actuation widely applicable to microfluidic systems.

CONCLUSION

Two 2D shut-off hydrogel microvalves and one 3D hybrid microvalve have been fabricated using an *in-situ* polymerization. The time response (~ sec) of hydrogel actuators has been improved by scaling down the hydrogel objects. The maximum differential pressure was 56.5 psi for the 2D valve and 26.7 psi for the 3D valve. Simple fabrication and no power consumption or integrated electronics make hydrogel microvalves attractive in many microfluidic applications.

ACKNOWLEDGEMENT

REFERENCES

1. G. T. A. Kovacs, *Micromachined Transducers Sourcebook*, New York: McGraw-Hill, 1998.

2. W. S. N. Trimmer, "Microrobots and micromechanical systems," *Sensors and Actuators*, vol. 19, pp. 267-287, 1988.

3. M. Madou, *Fundamentals of Microfabrication*. Boca Raton, FL: CRC Press LLC, 1997.

4. D. J. Beebe, J. S. Moore, J. M. Bauer, Q. Yu, R. H. Liu, C. Devadoss, B.-H. Jo, "Functional Structures For Autonomous Flow Control Inside Microfluidic Channels", *Nature*, in press, 2000.

5. R. H. Liu, Q. Yu, J. M. Bauer, B.-H. Jo, J. S. Moore, D. J. Beebe, "In-channel Processing to Create Autonomous Hydrogel Microvalves", *accepted in μTAS 2000*, Netherlands, 2000.

6. Y. Osada and S.B.Ross-Murphy, "Intelligent Gels," *Scientific American*, pp.82-87, 1993.

7. R. D. Cook and W. C. Young, *Advanced Mechanics of Materials*, New York: Macmillan, 1994.

8. M. T. A. Saif, B. E. Alaca, and H. Sehitoglu, "Analytical Modeling of Electrostatic Membrane Actuator for Micro Pumps", *IEEE J. of Microelectromechanical Systems*, Vol. 8, No. 3, Sept., 1999, pp. 335-345.

9. B. H. Jo, L. M. Van Lerberghe, K. M. Motsegood and D. J. Beebe, "Three-dimensional Micro-channel Fabrication in Polydimethylsiloxane(PDMS) Elastomer", *Journal of Micro-electro-mechanical Systems*, Vol. 9, No. 1, March, 2000, pp. 76-81.

10. E. S. Matsuo and T. Tanaka, "Kinetics of Discontinuous Volume-Phase Transition of Gels," *J. Chem. Phys.*, vol. 89, pp. 1695-1703, 1988.

11. D. Armani and C. Liu, "Re-configurable Fluid Circuits By PDMS Elastomer Micromachining", *12th International Conference on MEMS, MEMS 99*, Orland, FL, 1998, pp.222-227.

SWITCHABLE ELECTROSTATIC MICRO-VALVES WITH HIGH HOLD-OFF PRESSURE

Mike L. Philpott, David J. Beebe*, Anthony Fischer, Bruce Flachsbart,
Mike Marshall, Norman R. Miller, John C. Selby, Mark A. Shannon, and Yan Wu

Department of Mechanical Engineering
University of Illinois at Urbana-Champaign
Urbana, IL 61801

*Department of Biomedical Engineering
University of Wisconsin at Madison
Madison, WI 53706

ABSTRACT

A family of electrostatically-actuated micro-valves is presented. The valves are comprised of a silicon-based substrate and a polyimide-based membrane flapper or slit valve. The valves employ touch-mode capacitance actuation, allowing very large forward pressures to be held, extremely small reverse flows, plus rapid and in-phase timing. However, the peak pressures attainable, as well as the operating characteristics, are strongly dependent on material selection and process parameters.

INTRODUCTION

Electrostatically actuated micro-valves were developed to control high-pressure flows through a micro-compressor for an integrated mesoscopic cooler circuit (IMCC) [1]. The goal was to create valves that could be timed with the compressor stroke, would have large forward and zero reverse flow rates, and consume little power. Many micro-valves have been previously described [2-11], however, none could meet the necessary specifications for this application. The new electrostatically actuated micro-valves presented in this paper are able to achieve the rigorous specifications of the IMCC's refrigeration system. These valves have held over 18 atm (> 1.8 MPa) of forward pressure across ports greater than 100 μm wide, allowing large flow rates in excess of 3 ml/s. Moreover, these valves have demonstrated immeasurably small reverse flows, and rapid switching (order milliseconds). After charging during a cycle, the valves can continue to hold off large pressures, without further application of voltage or current flow. Power consumption only occurs during switching. A number of applications that require low-power on-off flow control are now possible using these micro-valves. With changes to the fabrication recipe, bi-stable micro-valves that are either normally open or closed can be created.

The mode of operation for these valves is via touch-capacitance [12], which theoretically can hold-off high pressures. However, during valve development, we found that charge migration through the dielectric layers and charge trapping at the interface dominate the operation at high-applied fields. These charging phenomena are strongly material and process dependent. This paper describes the design, fabrication, and operation of these metalized polyimide (PI) on silicon micro-valves. Results are given that demonstrate the performance of the valves with respect to different material stacks and processing variables.

FABRICATION

The valves are fabricated in two parts: a Si-based substrate with an etched fluid-flow orifice, and a polyimide-based flexible membrane. The valve substrate shown in Fig. 1 uses double-sided polished <100> n-type silicon wafer. Square valve orifices ranging from 75 to 200 μm (150 μm typical) on a side at the top of the wafer surface, are etched through the wafer using a 30% KOH solution at 55° C, patterned using a 1500 Å dry oxide and a 400 Å Cr mask on top. After the Cr and oxide mask is removed in standard Cr-etchant and buffered oxide etch (BOE), respectively, the wafers are heavily phosphorous doped at 1150° C to concentration levels approximately 10^{21} atoms/cm^2 to create a n$^+$ conductive plane a few microns deep along the surface of the silicon. A thermally grown 1500 Å dry oxide layer is then grown on top of the doped layer to provide a high-strength dielectric layer. Standard 1000 Å (99%Al-1%Si) contact layers are deposited as pads, followed by a 450° C anneal for 20 minutes to lessen the Schottkey barrier. For the purposes of testing, a pipette is epoxied to the backside to apply pressure.
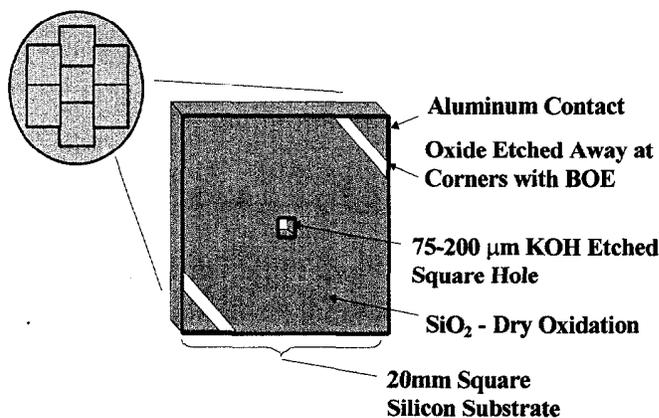
Figure 1. Valve substrates (Si) have a KOH etched port for a fluid orifice, and are heavily doped to form a conductive plane. A high dielectric strength thermal oxide is grown on top the n+ layer. Electric contact is made at both corners to insure continuity.

The flexible valve membrane is made with a layered stack of polyimide (PI) and metals, as shown in Fig. 2 for a typical case. PI 2611 from DuPont (BPDA-PPD polyimide) was used for the structural and dielectric portions of the membrane due to its chemically inertness, excellent temperature and fatigue resistances, and dielectric properties. First, a PI dielectric layer is spun and cured on a polished silicon wafer. A metal stack is then deposited

to create a conductive layer that includes 100 Å of adhesion metal, 1000 Å of conductive plane metal, followed by another 100 Å of adhesion metal. Two metal stack systems were tested: Cr/Al/Cr was sputtered using a RF powered dual-target Cooke sputtering system, and Cr/Au/Cr was evaporated using a NRC evaporator with two thermal sources and an e-beam gun. The metal layers were then patterned to remove metal from the region where openings in the polyimide were to be etched, as shown in Fig. 3. After metal patterning, one or two PI structural layers were spun on top of the metal stack and cured. The curing will be presented below. After curing of the structural layers, the slits shown in Fig. 3 are then patterned in a RIE oxygen plasma, using sputtered and patterned Al as the RIE mask. In the plasma, areas of the PI structural layer are etched away to allow electrical contact to the metal layer. The Al is removed with photoresist developer, to avoid Al ion migration into the PI. After final patterning, each sample is peeled from the substrate with a straight razor edge and a pair of flat-headed tweezers. The membrane is then sliced with the edge of a straight razor into the final valve strip, approximately 7 mm wide, which is then transferred and aligned to the substrate for testing.

Figure 3. Photograph of a flapper-style electrostatically actuated micro-valve mounted over a fluid port and channel, schematically depicted in the wire frame drawing to the left. The clear section is plasma etched through, the darker gray is two layers of PI, and the remaining opaque regions are metalized with Cr/Au/Cr stack.

Figure 2. A typical fabrication sequence of film stacking for the flexible valve membrane. PI dielectric layer is spun onto a polished Si wafer and cured, a metal stack is deposited to provide a conductive plane, then one or more PI structural layers are spun and cured to provide additional stiffness.

The curing of the PI proved to be a very important parameter for the dielectric layer. For the control group, we cured on a hotplate in air by ramping the temperature from 110° C to 350° C at 60° C/hr over the span of four hours, holding it at 350° C for one hour, and then ramping the hotplate back to ambient conditions under natural cooling. On other samples cured on the hotplate, we increased the curing time up to a maximum of 16 hrs. The remaining samples were cured in a 1 Torr nitrogen atmosphere at temperatures up to 400°C. The structural PI layers were all cured under the hotplate control conditions. Each cure schedule changed the results, as will be discussed below.

The testing was conducted using the pressure apparatus shown in Fig. 4, and a similar device with flow and pressure transducers to measure forward and reverse flows. The valve assembly was encased in a plexiglass container for flow measurements and to visualize the operation, as shown in the inset of Fig. 4. The hold-off pressure for a particular actuation voltage was determined by increasing the pressure until the valve opened.
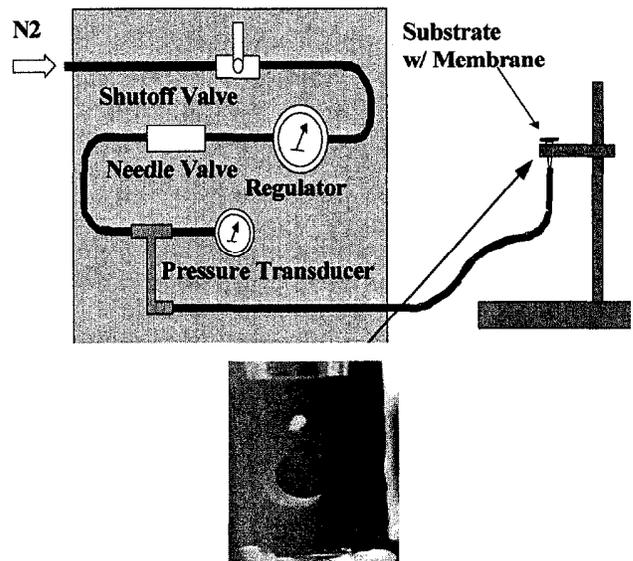
Figure 4. The hold-off pressure test rig with a blow-up of the valve mounted within a plexiglass holder, which is used to measure forward and reverse flows and pressures.

RESULTS AND DISCUSSION

To optimize the polyimide based micro-valve in order to hold a statically applied forward pressure of 200 psi over a 150 µm hole, we varied and tested five variables: (i) the PI dielectric thickness, (ii) the number of PI dielectric layers, (iii) the type of metals, (iv) the number of PI structural layers, and (v) the curing of the PI.

The results of extensive testing over several hundreds of valves can be summarized as follows with regards to (i) to (v). (i) *The thickness of the PI dielectric layer does not affect the*

maximum hold-off pressure achievable, for a given applied field strength. This result is due to the essentially 1-D electrostatic attraction between the conductive layers when the valve is closed. Therefore, the voltage required to attain the same hold off pressure increases linearly with the dielectric thickness, keeping the applied electric field constant. (ii) *The number of PI dielectric layers has a strong effect on hold-off pressure.* In fact, a multi-layered PI dielectric film exhibits a dramatically reduced maximum hold-off pressure when compared to a single layer PI film of identical thickness. This result is due to charge migration and trapping at the interface between the two layers, thereby shielding the two conductive planes. Shielding decreases the effective field between the conductive planes, and thus the attractive force that withstands the pressure acting on the valve. (iii) *The type of metals used in the stack has a strong effect on valve performance.* If Al is used in the stack, over extended time and straining, electromigration of Al into the PI dielectric appears to occur, even if a thin Cr adhesion layer is present. The electromigration causes shielding to occur between the conductive planes, again reducing the overall pressures attainable. When replaced by Au, the shielding problem was not observed to occur. (iv) *The number and thickness of the structural layers exhibits strong affects on maximum pressure attainable.* The stiffer the membrane, up to a point, the easier the valve can sustain the applied pressure, without pulling up from the substrate. The tradeoff is that stiff membranes, particularly those in the slit valve configuration, demonstrate reduced flow rates when in the open position. (v) *The curing schedule strongly affects the maximum pressures attainable, and the power required to actuate the valve.* In general, as the membrane is cured longer, the hold-off pressure is increased, even if full imidization of the polyimide has occurred. This result appears to be due to the effect of NMP solvent concentration within the PI layer, and its effect on charge migration. Reduced levels of NMP decreased the steady-state current through the dielectric layers, and increased the attainable hold-off pressures.

Figure 5 shows some of the hold-off pressure data for a variety of membrane dielectric thicknesses and fields. The apparent saturation of pressure in Fig. 5 is due to the maximum pressure attainable by the apparatus of 13 atm. The testing apparatus was recently extended to handle a maximum hold-off pressure of 18 atm. Subsequently, many valves have reached this hold-off pressure. The forward pressure is not a strong function of the dielectric thickness, rather it strongly depends on the thickness of the structural layer and the cure time. Longer cures, and/or vacuum cures can increase the hold-off pressure by an order of magnitude or more, compared to the standard cure schedule.

A typical valve operation in our work involves forward pressurization and subsequent valve opening. Fig. 6 shows the forward pressure as a function of time for a normally open micro-valve (Fig. 3), as it is switched off. The pressure rise and fall response is primarily a function of the pressure tester. The valves open and close within milliseconds. When closed, a reverse pressure up to the limits of the system have been applied without any recordable reverse flow or pressure rise. When operated passively, measurable flow rates and pressure rises occur, showing the effectiveness of electrostatic clamping for micro-valves.

Slit-type valves were also fabricated, where the side channels shown in Fig. 3 were eliminated. The slit valves followed the same performance as the flapper-style valves. The main reason for using a slit-type valve was due to curling of the flapper valves from residual stresses that arose during fabrication

and membrane transfer. Slit valves are simply highly constrained, or tethered, flapper valves. The performance of the valves was very similar to the flapper-style valve, save that the yield was more robust. It is conjectured that slit-type valves with thick structural layers would produce a lower flow rate in the forward direction. However, due to the size of our slit (1.5 mm) vs. the size of the orifice (150 μm^2), no changes were observed.
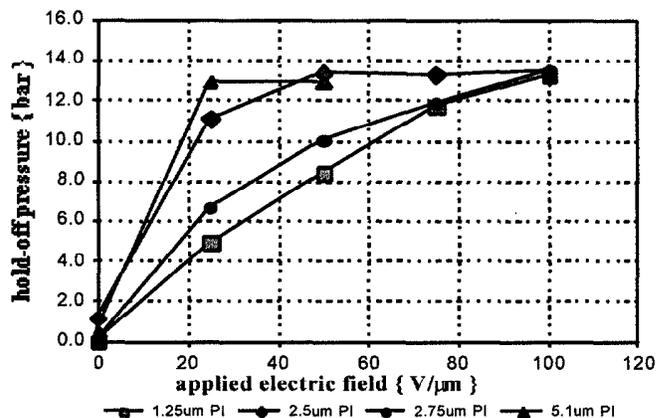
Figure 5. Some hold-off pressure data for membranes with different PI dielectric layer thicknesses and cure times vs. the applied electric field. The test rig's maximum pressure is 13 atm for this data. The differences in hold-off are determined most strongly by the PI dielectric layer cure time, not thickness.
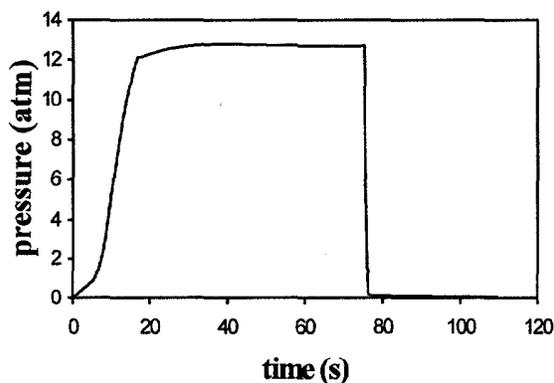
Figure 6. The operation of a flapper style valve under forward pressure, as the valve is pressurized and then shut-off (opened) electrostatically.

ACKNOWLEDGEMENT

REFERENCES

[1] M.A. Shannon, M.L. Philpott, N.R. Miller, C.W. Bullard, D.J. Beebe, A.M. Jacobi, P.S. Hrnjak, T. Saif, N. Aluru, H. Sehitoglu, A. Rockett, and J. Economy, "Integrated Mesoscopic Cooler Circuits (IMCCs)," *1999 ASME International Mechanical Engineering Congress and Exhibition*, Nashville, TN, Nov. 15-20, 1999, pp. 75-82.

[2] A. Lee, J. Hamilton, and J. Trevino, "A Low Power, Tight Seal, Polyimide Electrostatic Microvalve," MEMS ASME 1996, Vol. 59, pp. 345-349.

[3] M. Hirano, K. Yanagisawa, H. Kuwano, and S. Nakano, "Microvalve with Ultra-low Leakage," Proceedings of the 10th Annual International Workshop on Micro Electro Mechanical Systems, (1997) pp. 323-326.

[4] W.K. Schomburg and C. Goll, "Design Optimization of Bistable Microdiaphragm Valves," Sensors and Actuators A, Vol. 64, (1998), pp. 259-264.

[5] D.S. Popescu, D.C. Dascalu, M. Elwenspoek, and T. Lammerink, "Silicon Active Microvalves Using Buckled Membranes for Actuation," The 8th International Conference on Solid-State Sensors and Actuators, and Eurosensors IX, Transducers '95 (June 25-29, 1995), pp. 305-308.

[6] X. Yang, C. Grosjean, Y-C. Tai, and C-M. Ho, "A MEMS Thermopneumatic Silicone Membrane Valve," Proceedings of the 10th Annual International Workshop on Micro Electro Mechanical Systems, (1997) pp. 114-118.

[7] M. Shikida, K. Sato, and S. Tanaka, "Characteristic of an Electrostatically Driven Microvalve," Micromechanical Systems ASME 1993. Vol. 46, pp. 43-47.

[8] N. Vandelli, D. Wroblewski, M. Velonis, and T. Bifano, "Development of a MEMS Microvalve Array for Fluid Flow Control," Journal of Microelectromechanical Systems, Vol. 7, No. 4, (December 1998) pp. 395-403.

[9] C. Vieider, O. Ohman, and H. Elderstig, "A Pneumatically actuated Micro Valve with a Silicone Rubber Membrane for Integration with Fluid-Handling Systems," The 8th International Conference on Solid-State Sensors and Actuators, and Eurosensors IX, (June 25-29, 1995) pp. 284-286.

[10] X. Wang, Z. Zhou, and W. Zhang, "A Micro Valve made of PSPI," MEMS ASME 1998, Vol. 66, pp. 31-36.

[11] X-Q. Wang, Q. Lin, and Y-C. Tai, "A Parylene Micro Check Valve," Proceedings of the 12th Annual International Workshop on Micro Electro Mechanical Systems, (1999),   pp. 177-182.

[12] , J.R. Gilbert and S.D. Senturia, "Two-phase actuators: Stable zipping devices without fabrication of curved surfaces," Solid-State Sensor and Actuator Workshop, Hilton Head SC, 3-6 June 1996, pp. 98-100.

FULLY MICROFABRICATED, SILICON SPRING BIASED, SHAPE MEMORY ACTUATED MICROVALVE

Geon Hahm, Hal Kahn*, Stephen M. Phillips, Arthur H. Heuer*
Department of Electrical Engineering and Computer Science, *Materials Science and Engineering
Case Western Reserve University
Cleveland, Ohio 44106-7221

ABSTRACT

A microfabricated normally-closed microvalve consisting of a flat silicon spring, a co-sputtered and patterned Titanium-Nickel (TiNi) shape memory alloy (SMA) actuator, and an orifice die, is presented. All three components are batch microfabricated using silicon substrates. This microvalve is designed for an actuation mechanism for a pneumatically controlled, biologically inspired six-legged microrobot. We have used a TiNi SMA thin film as an actuator since it is capable of both high strains (3%) and high forces (work density of 5×10^7 J/m^3). Flow rates of up to 0.17 lpm are achieved when the input air pressure is 34.5 kPa (5.0 psi). A CNC-machined plastic package accommodates the electrical and fluidic interconnects.

INTRODUCTION

There has been extensive research on various types of valves, especially active valves, using microfabrication technology since the late 1980s. Thermopneumatic [1], bimetallic [2, 3], shape memory alloy [4, 5], electrostatic [6], piezoelectric [7], magnetic [8], and giant magnetostrictive [9] based mechanisms have all been used for actuation. Each actuation method has its own benefits and would be used for specific applications. While thermal inkjet printer nozzles currently have the largest market of all MEMS fluid control devices, the potential is also growing for medical applications. The requirements in pressure tolerance and low leak rates must be fulfilled to compete with conventional valves for many applications [10].

Microvalve design specifications include flow rate, operating pressure, power consumption, actuation speed, leakage, packaging and connectivity to its operating domain. A TiNi SMA actuator has been chosen because its high strains and actuation forces enable high fluid flow rates and high working pressures, respectively.

TiNi SMA exhibits reversible transformation between two characteristic phases: relatively stiff austenite at high temperatures, and relatively ductile martensite at low temperatures. Both forward and reverse transformations occur above room temperature. Since the transformations occur over a range of temperatures, proportional flow control can be achieved by precise temperature control. The temperature of the SMA actuator can be controlled by joule heating, using an applied current, and forces are generated by the phase transformation.

The successful use of TiNi thin films in MEMS applications has been hindered in previous efforts. One of the most serious problems is the reliable batch fabrication of TiNi thin films with reproducible transformation temperatures and transformation strains. The transformation temperatures and strains are very sensitive to compositional variation. Since the Ti and Ni constituents in alloy sputtering targets have different sputtering yields during deposition, a co-sputtering procedure has been developed which uses an alloy TiNi target and an elemental Ti target to reliably achieve stoichiometric SMA films [11].

Packaging and assembly of MEMS devices are important considerations which should be improved for product completeness and are application specific [12]. For our microvalve, a package has been developed, which enables easy, reliable assembly and repeatable testing of the fluidic device components. A package for arrays of microvalves is also being designed.

OPERATION OF MICROVALVE

The microvalve is composed of three components - a flat silicon spring, a patterned TiNi SMA actuator, and an orifice die. They are assembled on top of each other (Figure 1). The valve is normally closed; the fluid flow is proportional to the current applied. The patterned TiNi actuator and silicon spring allow "flow-through" operation. When the fluid direction is reversed, this leads to pressure balancing in the valve which enables much higher input pressures to be modulated.

The spring provides an initial closing force against the orifice making the valve normally closed when the SMA actuator is in its low temperature martensite phase (Figure 1a). When the actuator is heated, it transforms to austenite, forcing the spring upward, opening the valve (Figure 1b). The orifice die

a)

b)
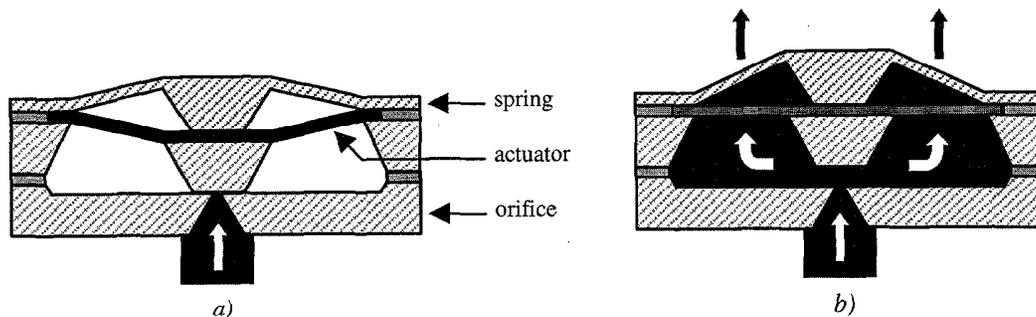
Figure 1. Cross-section of a microvalve , a) closed position, b) open position
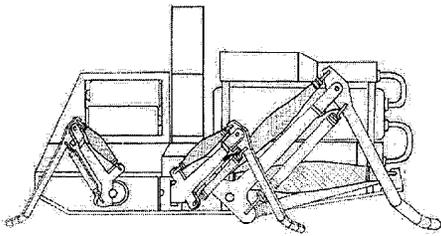
Figure 2. *Conceptual design of the microrobot*

simply contains a square inlet.

Assembled microvalves will be integrated as an array with pneumatic systems to control actuators in legs of a microrobot. Figure 2 shows the conceptual design of the microrobot.

DESIGN

A microvalve using an SMA diaphragm actuator has previously been developed in our lab [13]. Its shortcomings included non-uniform temperature across the diaphragm, slow transformation times, and large power consumption. To address these in the current valve design, the TiNi film is patterned into a thin strip, as shown in Figure 3. Two types of finite element analysis have been performed to optimize the design – coupled thermal-electrical analysis and stress-displacement analysis [14]. The reduction of thermal mass greatly reduces the response time and lowers the power consumption.

Component dimensions are intimately related to the mechanical characteristics of the TiNi film and the fluidic device specifications. Specifically, the orifice size is determined by the desired flow rate and the stroke of the boss. The stroke is determined by the dimensions of the patterned TiNi SMA actuator whose design also depends on the force required to modulate the fluid, which depends on the pressure. The spring design is determined by the recovery force of the actuator. Since the spacings between each component are also important factors determining the performance of the valve, accurate control of the thickness of the spacing or bonding material during assembly must be accomplished.

FABRICATION

One of the most important considerations in fabricating MEMS devices containing TiNi is process compatibility,
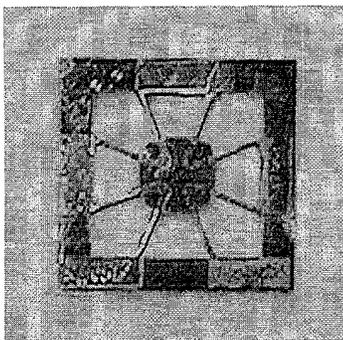
Figure 3. *Patterned TiNi SMA actuator, 7.5 mm × 7.5 mm*
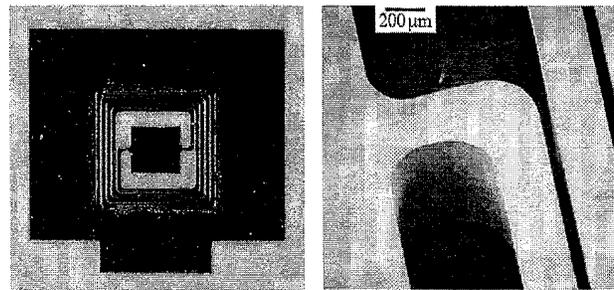
Figure 4. *a) Plan view of the bias silicon spring, 7.5 mm × 7.5 mm; the bottom corners are etched for electrical connection to the actuator, b) SEM of thick silicon spring fabricated using a DRIE etcher*

especially with regards to temperature. Since TiNi oxidizes easily when it is heated above ~300 °C in air, process flows are constrained.

The actuators are fabricated by the following sequence. First, a thin (25 μm) silicon diaphragm with a central boss is created by tetramethylammonium hydroxide (TMAH) etching. Convex corner compensation is applied to generate the square boss shape in a limited area. Next, a 2-μm TiNi thin film is co-sputtered on the silicon diaphragm. Then, the TiNi is patterned into a thin line with standard photolithography and wet etching. The silicon diaphragm is removed in an SF$_6$ plasma to release the patterned TiNi film which is attached to the central boss (Figure 3).

The TiNi thin film SMA's are co-sputtered using a TiNi alloy target and an elemental Ti target with two independently-controlled DC magnetron guns. During deposition, the chamber is evacuated and back-filled with argon gas to 3.5 mTorr. The wafer is rotated and heated to 230 °C, and a deposition rate of ~0.25nm/sec is used. After deposition, the film is annealed *in situ* at 400 °C for 15 minutes after evacuating the Ar gas from the deposition chamber. Stoichiometric SMA films are routinely achieved.

400-μm thick silicon springs with 200-μm wide arms are bulk micromachined using deep-reactive-ion-etching (DRIE) with a 10 μm-thick photoresist layer as a mask (Figure 4). The 480 μm × 480 μm orifices are bulk micromachined using TMAH. The etch rate is ~48 μm/hour at 90 °C.

PACKAGING AND ASSEMBLY

Two types of packages have been developed. One is for a Au/Si eutectic bonded stand-alone device (Figure 5) and the other is for testing device components that have been assembled but not permanently bonded (Figure 6). Computer aided machining is used to cut the packages directly from CAD drawings using acrylic as the structural material. Both designs use two flexible interconnect contacts (gold plated 2.5 mm long bellows contact springs) to supply the current to heat the SMA actuator, without any direct bonding to the TiNi film surface. The absence of rigid bonds allows for slight misalignment in assembly, and easy re-use of the package.

For the component testing package (Figure 6), the three components of the valve are stacked in the trench machined into the lower half of the package, and the top cover is screwed on tightly. An o-ring (outer diameter of 2.4 mm) provides sealing between the orifice and package at the inlet port, and a push-in
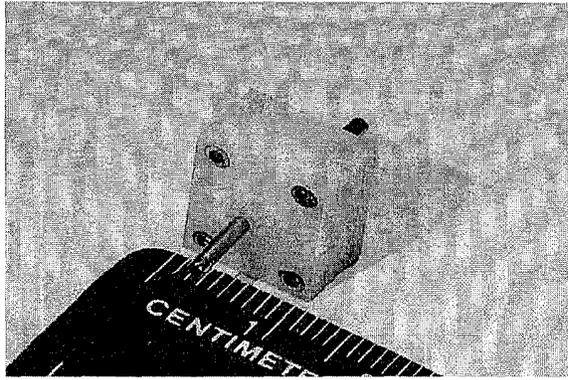
Figure 5. Valve with stand-alone package
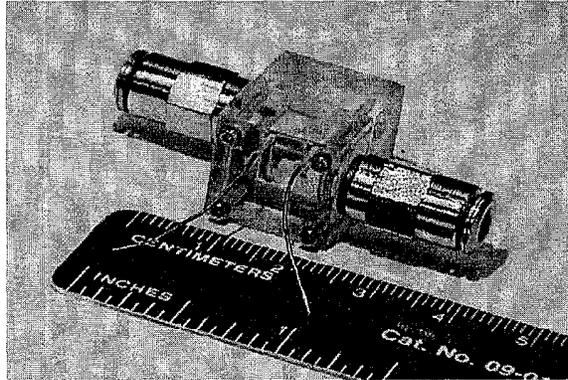
Figure 6. Valve with test package

fitting system (tube diameter 5/32") is adapted in the test fixture package for easy pneumatic connections. This packaging scheme allows reliable and repeatable testing of the device components. The stand-alone package (Figure 5) uses thin stainless steel tubing (1/32" diameter) for fluidic connections.

The dimensions of an assembled valve without packaging are 7.5 mm × 7.5 mm × 2 mm. Those with packaging are 11 mm × 13 mm × 7.5 mm for stand-alone, and 16 mm × 16 mm × 20 mm for component testing, without fittings.

TESTING AND RESULTS

The silicon spring force-deflection behavior was measured by a precision balance and a transfer stage with 1 micron
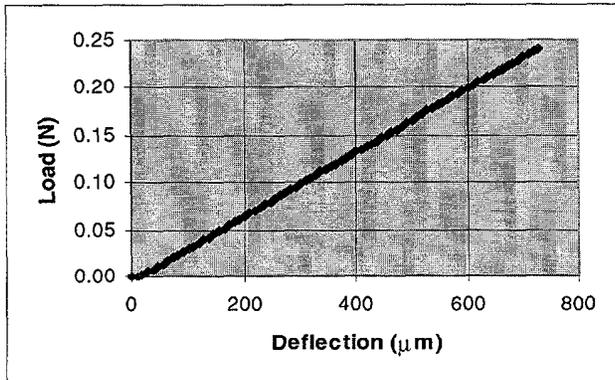
Figure 7. Spring load-deflection characteristic
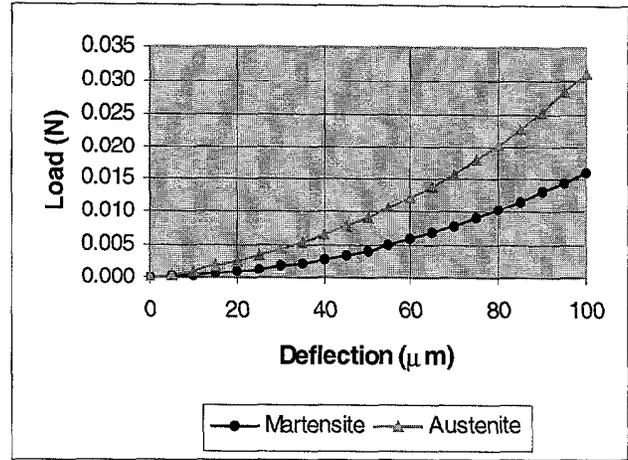
Figure 8. Actuator load-deflection characteristic
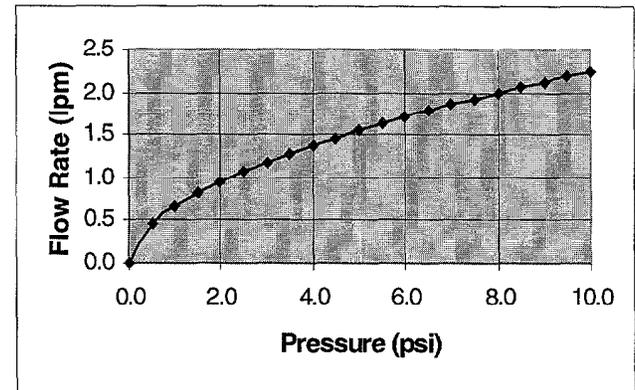
Figure 9. Flow characteristics of an orifice (480 μm × 480 μm)

resolution. It was linear to greater than 730 μm with a spring constant of 330 N/m (Figure 7).

The load-deflection characteristics of a patterned actuator were also acquired (Figure 8). From this data, both recovery force (at a given deflection) and recovery deflection (at a given load) can be determined.

A non-contact, fiber optic, displacement sensor (880-nm light source) was used to measure displacement and time response during the heating cycle of an SMA actuator attached to a silicon spring. The full deflection range of the combination of a spring and an actuator with a 400-μm thick spacer is ~75 μm. Time response was measured with a fully assembled device without fluid by applying a 1-Hz square wave signal. The opening time was 50 msec and the closing time was 18 msec. The orifice was also tested, and Figure 9 shows air flow rates through the orifice versus inlet pressure.

Based on dimensions and measurements from all three elements, proper spacer thickness was chosen. The components were assembled and microvalves were investigated for gas (air) flow modulation capabilities with the test setup in Figure 10. Flow rates of up to 0.17 lpm were measured when the input air pressure was 34.5 kPa (Figure 11) and the leakage was 0.005 lpm. The leakage was most likely due to the surface roughness of the underside of the boss and the top of the orifice, and possible misalignment between the boss and the orifice. It could be reduced by microfabricating a valve seat structure, depositing a flexible thin film at the valve seat, or making series of circular grooves around the inlet port [7].
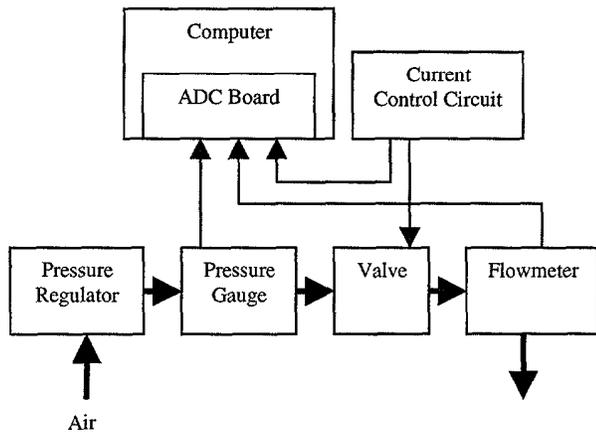
Figure 10. Test setup

CONCLUSION

A normally-closed TiNi SMA actuated microvalve for gas flow control has been fabricated and successfully tested. It consists of a silicon spring, a patterned TiNi SMA actuator, and an orifice layer. All components are microfabricated from silicon wafers and are compatible with batch processing. Two types of packages have been developed using computer aided machining. This enables easy assembly and repeatable testing. The valve is tested for gas (air) flow modulation. Flow rates of up to 0.17 lpm are demonstrated when the input air pressure is 34.5 kPa.

FUTURE WORK

To optimize device performance, two enhancements are necessary - a reduction in the leakage flow and more reliable packaging. A new package will be developed for a matrix of valves, since one microrobot requires 32 valves. Au/Si eutectic wafer bonding will be further refined for batch assembly.

Polyimide thermal insulation layers are being pursued to reduce the power consumption and response time, and for fluid isolation when using liquid media. In addition, driving the actuator with pulses of current are being investigated to further reduce the power consumption. Efforts to reduce the leakage are underway.
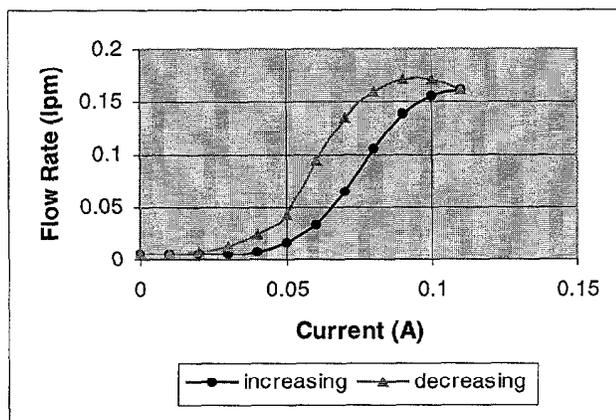
Figure 11. Flow rate vs. current (air, 5 psi)

ACKNOWLEDGMENTS

REFERENCES

1. M. J. Zdeblick, R. Anderson, J. Jankowski, B. Kline-Schoder, L. Christel, R. Miles, W. Weber, "Thermopneumatically actuated microvalves and integrated electro-fluidic circuits," *Solid-State Sensor and Actuator Workshop*, 1994, pp. 251-255.

2. P. W. Barth, C. C. Beatty, L. A. Field, J. W. Baker, G. B. Gordon, "A robust normally-closed silicon microvalve," *Solid-State Sensor and Actuator Workshop*, 1994, pp. 248-250.

3. H. Jerman, "Electrically-activated, micromachined diaphragm valves," *J. Micromech. Microeng.* 4, 1994, pp. 210-216.

4. C. A. Ray, C. L. Sloan, A. D. Johnson, J. D. Busch, B. R. Petty, "A silicon-based shape memory alloy microvalve," *Mat. Res. Soc. Symp. Proc.* Vol. 276, 1992, pp. 161-166.

5. W. L. Benard, H. Kahn, A. H. Heuer, M. A. Huff, "A Titanium-Nickel shape-memory alloy actuated micropump," *Int. Conf. Solid-State Sensors and Actuators*, 1997, pp. 361-364.

6. M. A. Huff, M. S. Mettner, T. A. Lober, M. A. Schmidt, "Fabrication, packaging, and testing of a wafer-bonded microvalve," *Solid-State Sensor and Actuator Workshop*, 1992, pp. 194-197.

7. I. Chakraborty, W. C. Tang, D. P. Bame, T. K. Tang, "MEMS micro-valve for space applications," *Transducers '99*, 1999, pp. 1820-1823.

8. R. L. Smith, R. W. Bower, S. D. Collins, "The design and fabrication of a magnetically actuated micromachined flow valve," *Sensors and Actuators A*, 24, 1990, pp. 47-53.

9. E. Quandt and K. Seemann, Fabrication of giant magnetostrictive thin film actuators, *Proc. IEEE MEMS 1995*, 1995, pp. 273-277.

10. *MEMS 1999 Emerging Applications and Markets*, 1999. System Planning Corporation,

11. C.-L. Shih, Deposition and characterization of TiNi shape memory alloy thin film for MEMS applications, MS thesis, Case Western Reserve University, Department of Materials Science and Engineering, 1999.

12. Microelectromechanical Systems: Advanced Materials and Fabrication Methods, National Materials Advisory Board, National Academy Press, 1997.

13. H. Kahn, W. L. Benard, M. A. Huff, A. H. Heuer, "Titanium-nickel shape memory thin film actuators for micromachined valves," *Mat. Res. Soc. Symp. Proc.*, vol. 444, 1997, pp. 227-232.

14. A. G. Parr Jr., Design and characterization of a patterned actuator and the design and development of a smart controller for a microfluidic valve, MS thesis, Case Western Reserve University, Department of Electrical Engineering and Applied Physics, 1998.

A THERMOPNEUMATICALLY-ACTUATED MICROVALVE WITH IMPROVED THERMAL EFFICIENCY AND INTEGRATED STATE SENSING

Collin A. Rich[*] and Kensall D. Wise
Department of EECS, University of Michigan
Ann Arbor, MI 48109-2122

ABSTRACT

A thermopneumatically-actuated microvalve that is suitable for integration into batch-fabricated microsystems is reported. The valve achieves a 420sccm open flow under 1500torr differential and <0.01sccm leak rate, yielding a dynamic range >4×10^4. The saturated vapor pressure of a resistively heated working fluid deflects a corrugated silicon valve plate to close the valve. The heater grids are elevated above the substrate and the cavity is only partially filled to increase thermal efficiency. An integral capacitive pressure sensor allows direct monitoring of actuator pressure. Pentane-filled actuators sustain a 2070torr pressure rise with 500mW. A device tested *in-situ* closes with 350mW at 1000torr inlet (venting to vacuum) and maintains closure with 30mW input. Thermodynamic modeling matches experimental power, pressure, and transient response data to within a few percent. This model is used to suggest an optimized structure capable of a 2000torr pressure rise with 50mW input and a 1sec response time.

INTRODUCTION

Microfluidic systems constitute a steadily growing sector of the MEMS field. Potential applications include microreaction chambers, implantable drug delivery systems, "intelligent" petri dishes for microanalysis, a compact gas microchromatograph, and miniaturized mass flow controllers for precision applications. A suitable microvalve for such systems will have a wide dynamic range, low leak rate, fast response, and low power consumption. Additionally, compatibility with batch fabrication is important for developing integrated microfluidic systems.

This paper reports a thermopneumatically-actuated microvalve with a simple, glass-and-silicon composition that is suitable for integration into batch-fabricated microsystems. Thermopneumatic actuation was chosen for its robustness and ability to generate large forces over considerable distances, facilitating a valve with substantial dynamic range and low leak rate. Furthermore, response time and power requirements can be minimized with thermal optimization of the valve structure.

DEVICE STRUCTURE

The basic microvalve (see Fig. 1) comprises a silicon microstructure capped above and below by anodically-bonded glass plates. The valve diaphragm in the 1st Si bulk consists of a central boss supported by a corrugated suspension. In the open state, gas flow is determined by the inlet and outlet pressures, along with the dimensions of the flow channel. To close the valve, the diaphragm is deflected to occlude the gas inlet. The 1st Si fabrication process, reported elsewhere [1], uses an RIE etch to define the gas flow channel and valve plate corrugation profile.

Diffused boron etch-stops define the rim, boss, and corrugation thickness; an EDP etch step then releases the structure.
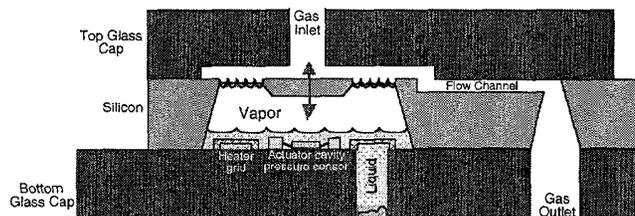
***Figure 1.** Thermopneumatically-actuated microvalve structure.*

Deflection force is generated by the vapor pressure of the heated working fluid, which can be a volatile solvent such as pentane or methanol. Because a saturated liquid-vapor phase exists in the cavity, temperature determines the vapor pressure regardless of fill level. Earlier work by Bergstrom [2] demonstrated that the elimination of excess working fluid decreases system thermal mass and improves transient response.

The heater grids and integral capacitive pressure sensor (for real-time monitoring of actuator cavity pressure) are fabricated in a 2nd Si layer bonded to the glass substrate. Elevating the heater grids above the glass substrate decreases steady-state power consumption by reducing direct heat loss to the substrate. Each of four heater grids consists of a lattice of 100μm-long beam segments forming diamond-shaped pores, and each beam has a 4μm-wide cross-section, shown in Figure 2. A 0.7μm-thick polysilicon layer is used for the actual heating element (rather than the heavily boron-doped bulk Si grid frame) in order to allow independent control of the heater resistance by varying the polysilicon doping level. (The grid frame doping is fixed at >10^{20}cm^{-3} to ensure an adequate etch stop.) The dielectrics electrically isolate the poly-Si and protect it from attack during the final EDP wafer-dissolution step.
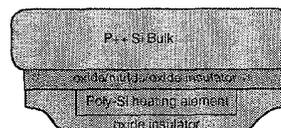
***Figure 2.** Cross-section of a beam in the heater grid.*

A modified version of the dissolved-wafer process developed by Chavan for vacuum-sealed pressure sensors [3] is used to simultaneously fabricate the heater grids and pressure sensors on the glass substrate. Chavan's work gives an in-depth analysis of capacitive pressure sensor dynamics. Table 1 presents the design parameters used for the actuator pressure sensor in this work.

[*]Collin Rich is now with Integrated Sensing Systems (ISSYS), Inc., Ypsilanti, MI 48197.

Table 1: Design parameters for actuator pressure sensor.

| Webbing thickness | 2.4μm | Boss/electrode radius | 124μm |
| | | Outer radius | 250μm |

	Pressure [torr]	Deflection [μm]	Gap [μm]	Capacitance [pF]	Sensitivity [fF/psi]
Min:	760	2.2	6.8	0.0625	1.4
Max:	3000	8.7	0.3	1.4162	701.0

A KOH recess step forms the heater grid posts and pressure sensor rim in the 2nd Si, followed by boron etch stops to define the heater grid and pressure sensor bulk profiles. CVD polysilicon sandwiched between CVD dielectrics forms the active heating element in the grids, along with the sensor and grid lead transfers. The 2nd Si is then wafer-level anodically bonded to a metallized lower glass substrate. Finally, an EDP etch-back step dissolves the wafer bulk, leaving the pressure sensors and heater grids bonded to the glass substrate. The entire device is assembled with two additional anodic bonds, after which the cavities are insulated with 3μm of vapor-phase-deposited parylene, partially filled with pentane, and sealed. Parylene, which has very low thermal conductivity (k = 0.0837W/mK), reduces heat loss from the working fluid to the surrounding cavity.

THERMAL MODELING

A thermal model was developed for the valve structure. The fabricated device has been designed to facilitate an 8-bit microflow controller architecture, in conjunction with a separate microcontroller. It includes eight individually-valved, binarily weighted flow-metering channels in parallel, along with inlet and outlet pressure sensors. A steady-state model was derived from device geometry and material parameters to determine the thermal resistance between an energized actuator and the surrounding ambient. Conduction was assumed in the interior and through the lower substrate surface (i.e., mounting surface), whereas convection was assumed from the top and sides of the device (see Fig 3.)
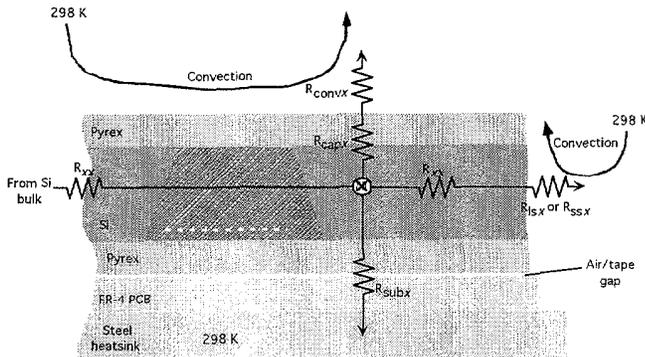
Figure 3. Simplified bulk thermal circuit for an analytical segment (including a non-energized actuator).

The steady-state thermal resistive network for the entire device is shown in Fig. 4. Simplifying the network and reducing the thermal mass of the system to a lumped-parameter model yields the final energized cavity model of Figure 5. Similar analysis shows that heat loss through the grid posts is negligible compared to dissipation into the working fluid.
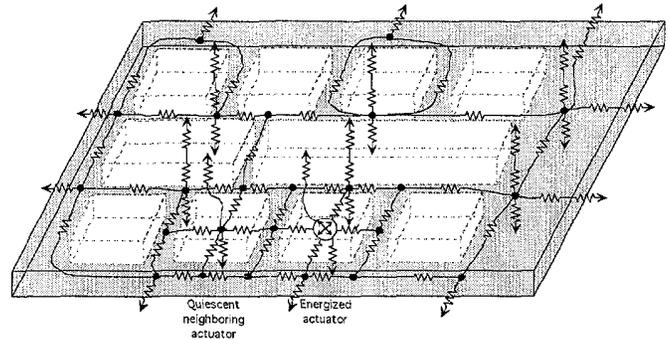
Figure 4. Thermal network for entire device. The eight small cavities are actuators. The two larger cavities contain the inlet and outlet pressure sensors.
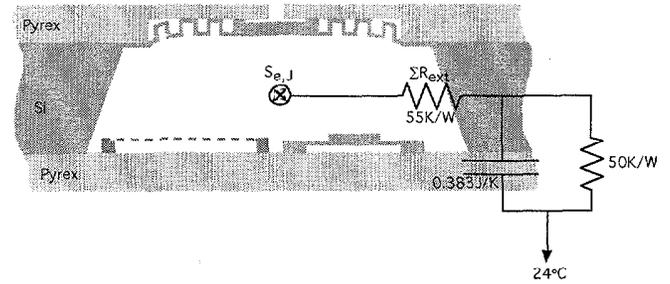
Figure 5. Lumped thermal model for energized actuator. Model relates input power to cavity temperature over time.

The Clausius-Clapeyron relation may be used to predict the pressure p_g of an actuator cavity at temperature T_{lg} (all temperatures in K), assuming the cavity contains working fluid in both liquid and gaseous states (i.e. the fluid is saturated). This relation may be cast in integrated form [4] as:

$$p_g = p_{g,o} \exp^{-\frac{M\Delta h_{rg}}{R_g}\left(\frac{1}{T_{lg}} - \frac{1}{T_{lg,o}}\right)}.$$

An additional partial pressure term must also be added to account for residual air (an ideal gas) in the cavity. After substituting appropriate parameters for pentane, simulations yield an expected response as shown in Fig. 6.
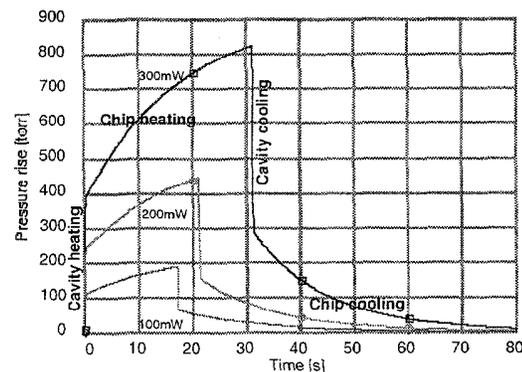
Figure 6 Simulated transient pressure vs. time for device mounted on semi-insulating substrate, at varied power levels.

EXPERIMENTAL RESULTS

An assembled device along with an EDP-released lower substrate is shown in Figure 7. Figure 8 shows a SEM cross-section of a corrugated diaphragm, and Fig. 9 is a close-up of an actuator pressure sensor and heater grid cluster.
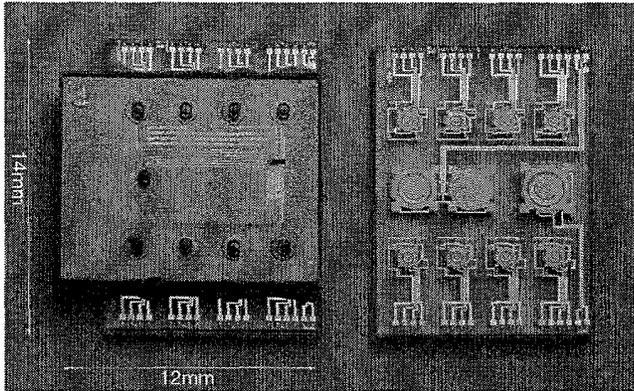
Figure 7. Left: assembled die with microvalves, flow channels, and pressure sensors for μFC configuration; Right: lower substrate with heater grids and pressure sensors.
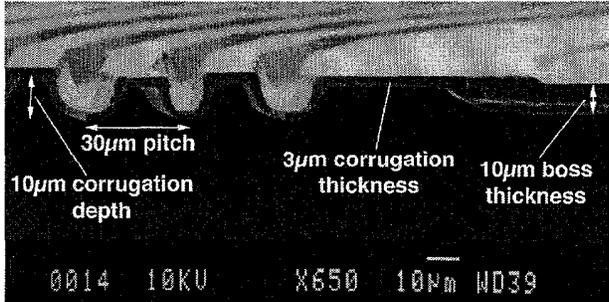
Figure 8. Cross-section of corrugations in valve diaphragm.
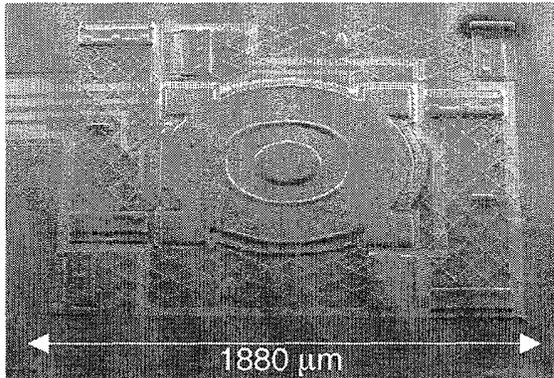
Figure 9. SEM of sensor/actuator cluster.

Flow rates for the structure varied linearly (flow was sonically limited) from 60-420sccm over a 700-1500torr range. Typical leak rates were 0.005-0.01sccm, giving a dynamic range of 4×10^4 at 1500torr. A set of seven diaphragms was also pulsed with 15psi for one million cycles to test long-term reliability. No failures occurred.

Figures 10-13 plot actuator performance for several devices mounted on an insulating FR-4 (PCB) base, in room air ambient.
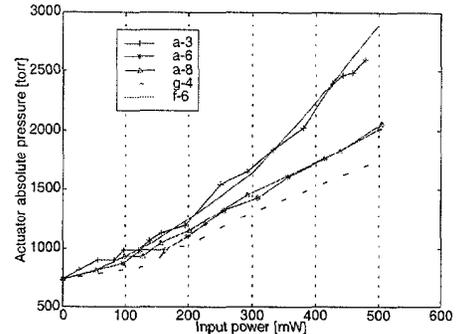
Figure 10. Steady-state pressure vs. power for pentane-filled actuators on insulating base.
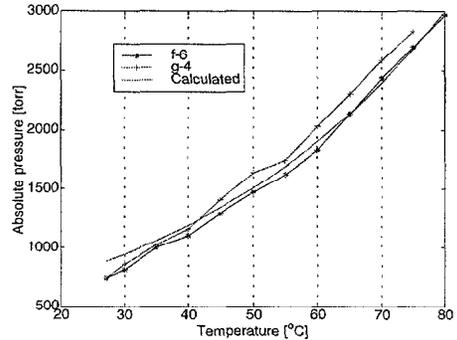
Figure 11. Steady-state pressure vs. temperature for pentane-filled actuators in temperature-controlled oven. Model prediction is also plotted.
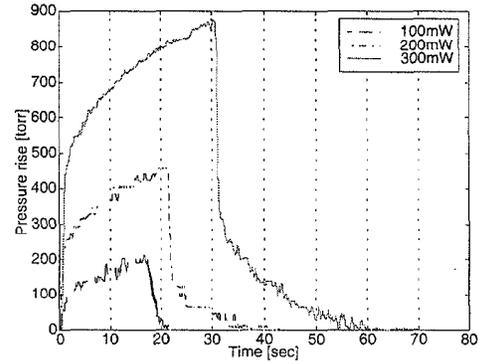
Figure 12. Transient response for one actuator on thermally-insulating base.
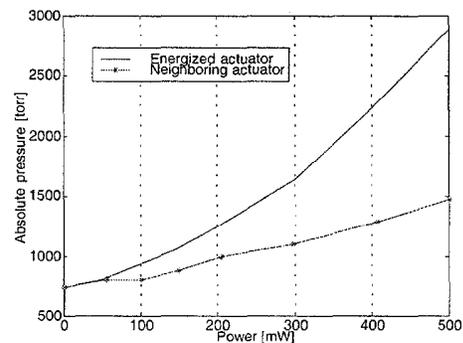
Figure 13. Cross-talk measurement: Steady-state pressure of "quiescent" actuator vs. power of neighboring, energized actuator. Pressure of energized actuator is also shown.

Table 2 summarizes the power performance of a valve *in-situ* at 1000torr inlet, venting to vacuum. With the substrate exposed to vacuum, heat loss diminishes dramatically.

Table 2. Valve performance at 1000torr inlet to vacuum.

Power [mW]	State	Flow [sccm]
0	Open	89
350	Closing	—
30	Closed	.034
50	Closed	.005

DESIGN OPTIMIZATION

Analysis of modeling and results suggests the optimized structure of Fig. 14, which would require 5-10% as much power to achieve the same performance. Features of the redesign include: minimizing the glass and silicon thermal masses (decreases substrate heat storage); surrounding the actuator cavity with an air- or vacuum-filled insulating buffer (decreases heat storage in rest of chip bulk and reduces crosstalk); and increasing the thickness of the insulating coating in the cavity (reduces bulk heat storage). Closure response time can also be improved with a closed-loop, spike-and-hold control scheme: quickly charge the thermal mass of the substrate with a high power spike, and then reduce power to maintain target pressure once the substrate equilibrates. An integrated pressure sensor is essential to such closed-loop control.
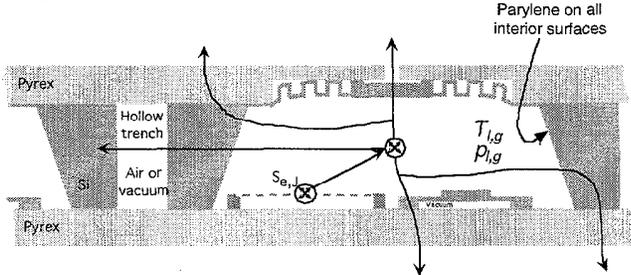
Figure 14. Revised valve structure for improved performance.

Simulation of the redesigned actuator predicts the response of Fig. 15. Table 3 compares the redsign to the experimental device. Furthermore, Fig. 16 shows the design space for the redesigned actuator. Device-to-ambient thermal resistance is the "tuning" factor: cooling time increases with resistance. From Fig. 16, a 2000torr pressure rise with 50mW input and a 1sec response time should be achievable.
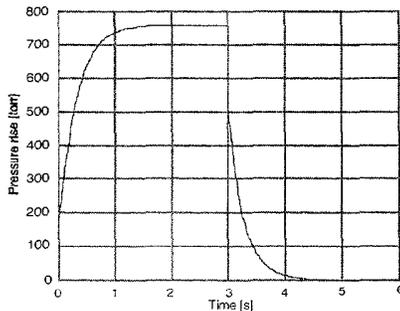
Figure 15. Simulated pressure response for the redesigned actuator with substrate mount similar to the fabricated device. Input power is 155mW. Ambient temperature is 24 °C.

Table 3. Existing and redesigned actuator performance.

Parameter	Existing	Redesign
Example pressure rise	750torr	750torr
90% rise/fall time	15sec	500ms
Required power	300mW	155mW
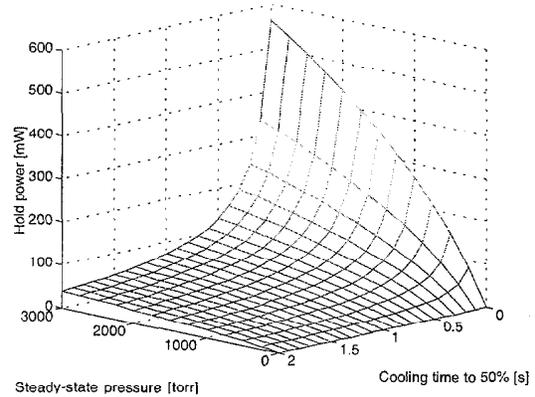
Figure 16. Required hold power vs. pressure rise and cooling (response) time for optimized actuator model. Cooling decay time is to 50% of steady-state pressure and is determined by thermal resistance of mount. $T_f=24°C$.

CONCLUSIONS

A thermopneumatic microvalve structure was designed, fabricated and tested. The structure demonstrated the feasibility of using raised heater grids and partially filled cavities to improve performance. The mechanical aspects of the valve (e.g. valve diaphragm) show good prospects for large dynamic range and long-term reliability. Real-time monitoring of actuator pressure was facilitated with integrated capacitive pressure sensors.

A model was also developed to predict performance for an optimized valve structure. The components of the model yield a good fit to experimental data. A redesigned actuator should achieve faster performance with an order of magnitude reduction in power consumption.

REFERENCES

1. C. Rich and K. Wise, "An 8-bit Microflow Controller Using Pneumatically-Actuated Microvalves," *Proc. 12th Int'l. Conf. Microelectromechanical Systems,* Orlando, FL, 1/99, IEEE (1999), pp. 130-134.

2. P. Bergstrom, et al., "Thermally Driven Phase-Change Microactuation," *J. Microelect. Systems, 4,* 10-17 (1995).

3. A. Chavan and K. Wise, "A Batch-Processed Vacuum-Sealed Capacitive Pressure Sensor," *Proc. 1997 Int'l. Conf. on Solid-State Sensors and Actuators,* 6/97, pp. 1449-1452.

4. M. Kaviany, *Principles of Heat Transfer,* Ann Arbor, MI, Fall 1999, pp. 717-718.

NOVEL PARAMETRIC-EFFECT MEMS AMPLIFIERS/TRANSDUCERS

Jean-Pierre Raskin, Andrew R. Brown and Gabriel M. Rebeiz
The University of Michigan, Radiation Laboratory, Ann Arbor, MI 48109-2122, USA

Butrus T. Khuri-Yakub
Stanford University, Edward L. Gizton Laboratory, CA 94305, USA.

ABSTRACT

A parametric effect amplifier has been built at 200 kHz (input) and 1.84 MHz (output) using a MEMS time-varying capacitor. The capacitor is composed of a thin low stress metallized silicon-nitride diaphragm and is pumped by a large signal voltage at 1.64 MHz. This results in a large change in the capacitance, and parametric amplification of an input signal at 200 kHz. To our knowledge, this device is the first-ever mechanical up-converter parametric-effect amplifier, with an up-conversion ratio of 9:1. Its main advantages are the absence of any resistive and *1/f* noise, and most importantly, it provides gain at the transducer level. Some advantages over CMOS based electronic amplifiers are the possibility to operate at very high temperatures (200-600°C), under high particle bombardment (nuclear applications), and very low noise operation due to the absence of thermal and shot noise in parametric amplifiers.

I. INTRODUCTION

Parametric amplifiers, based on non-linear reactances such as capacitors and inductors, have been extensively used in the 1960's for amplification, frequency up- and down-conversion and oscillators at microwave frequencies [1], [2]. Briefly, parametric devices transfer the power from the pump frequency (ω_p) to the signal frequency (ω_s) or to the up-conversion frequency (ω_u), thereby resulting in gain. Parametric amplifiers were commonly built using varactor diodes for microwave applications. Since parametric amplifiers are not resistive based, they do not suffer from the Johnson and *1/f* noise. Simple parametric amplifiers, using the negative resistance effect due to the non-linear reactance, resulted in narrowband amplifiers with a bandwidth of 1-6 %, depending on its gain. On the other side, parametric amplifiers using the parametric up-conversion effect, resulted in amplifiers with a bandwidth of 20-40 %. However, parametric up-conversion as not used since the frequency needed to be downconverted again using a microwave mixer, thereby introducing loss and noise in the system. With the introduction of MESFETs presenting excellent gain and noise characteristics at microwave frequencies, parametric amplifiers were abandoned and to our knowledge, they are not use anymore in any microwave applications. Frequency conversion (doublers, triplers) using a time-varying capacitance (varactor diodes) is still used today, but this effect results in a typical conversion loss of –4 dB to –10 dB (or a minimum conversion loss of 0 dB from the Manley-Rower equations) and is not called parametric amplification. This last decade, the parametric amplification has been used mostly in optical systems [3], [4]. In this paper, we present the interest of a novel MEMS parametric-effect for providing gain at the transducer level.

II. PARAMETRIC AMPLIFICATION THEORY

In 1956, Manley and Rowe derived a pair of equations governing the power balance of passive and lossless one-port devices with an arbitrary nonlinear characteristics [5]. As an example of such a device we may think of a deal varactor diode, of a Josephson junction or of a variable plate capacitor. If such a device forms a part of a parametric circuit is simultaneously fed by a signal source at frequency f_s and a pump source at frequency f_p. Due to its non-linear characteristics, time varying currents and voltages are generated, in principle, at all combination frequencies: $f_{n,m} = n\,f_s + m\,f_p$ with $n,m = -\infty \ldots \infty$. Manley and Rowe have shown that the following pair of equations holds:

$$\sum_{n=0}^{\times}\sum_{m=-\infty}^{\times} \frac{nP_{n,m}}{nf_s + mf_p} = 0 \quad (1) \text{ and } \sum_{m=0}^{\times}\sum_{n=-\infty}^{\times} \frac{mP_{n,m}}{nf_s + mf_p} = 0 \quad (2)$$

A general understanding of parametric circuits would be very difficult if all these frequencies were unconditionally necessary for the different modes of operation. This is not the case, however, and on the contrary, really not desirable. Actually, in practice, only a single combination frequency (f_2) is kept besides the signal frequency f_s and the pump frequency f_p, the other combination frequencies are filtered. The two most simple choices for f_2 are: *frequency-noninverting case*: $f_2 = f_u = f_p + f_s$ and *frequency-inverting case*: $f_2 = f_l = f_p - f_s$. In these cases, from the general results (1)-(2), the available power gain for both cases can be expressed respectively by:

$$G_u = \frac{P_u}{P_s} = -\frac{f_p + f_s}{f_s} \quad (3) \text{ and } G_l = \frac{P_l}{P_s} = -\frac{f_p - f_s}{f_s} \quad (4)$$

Relations (3)-(4) show that a parametric amplifier fed by an input signal f_s and pumped at a frequency f_p can present a large gain if $f_p \gg f_s$.

Considering the time-varying capacitance as a linear four-pole Rowe [6] has calculated the gain, the matching conditions, the bandwidth and the sensitivity for the inverting and noninverting cases. Because the inverting down- and up-converters are potentially unstable and the inverting negative resistance amplifier (degenerate case: $f_p = 2f_s$) requires an isolator (circulator) for operation since the source and the load are at the same port, the noninverting up-converter has been preferred as demonstrator for our study. In fact, as described by Rowe, the noninverting up-converter (signal frequencies f_s and $f_u = f_p + f_s$) is stable device and yields maximum gain with matched source and load. Under matched conditions the power gain is equal to the ratio of output to input frequencies (3); for widely separated signal frequencies the up-converter has substantial gain. Only a relatively small amount of nonlinearity is required to attain a bandwidth equal to the low-signal frequency. Since the source and load are matched the sensitivity is zero. Fig. 1 represents the scheme of the parametric up-converter which will be analyzed and built using the latest MEMS technologies.

Since we consider ideal input and output filters, all other sidebands are terminated with an open-circuit. In Fig. 1 $C(t)$ is the time-varying capacitance, R_c is the parasitic series resistance of the time-varying capacitor, R_G and R_L the source and the load resistances, respectively, and, $R_s + j\omega L_s$ and $R_u + j\omega L_u$ are, respectively, the terminal impedances added for matching purposes. For conceptual purposes, the circuit model of Fig. 1 can be analyzed as a linear four-terminal scheme. The transducer gain is defined as the ratio of the actual power output ($|I_L|^2 . R_L$) to the

available power input ($|V_G|^2/4R_G$) and is calculated by the standard linear circuit techniques.
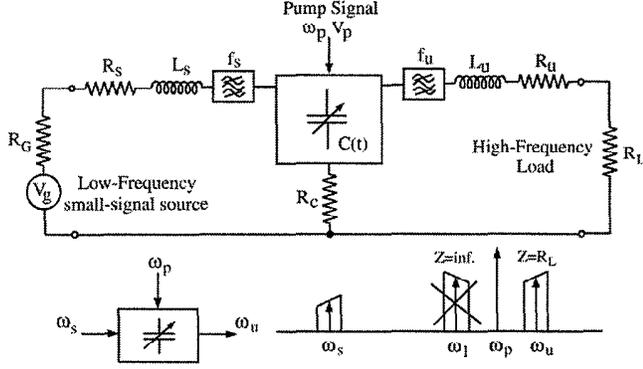
Figure 1. *Equivalent model and amplitude spectra for a generic parametric up-converter.*

Blackwell and Kotzebue in [7] have solved the transducer gain corresponding to equivalent circuit presented in Fig. 1.

$$G_t = \frac{4R_G R_L |I_L|^2}{|V_G|^2} = \frac{4R_G R_L |Z_{21}|^2}{|(Z_{11}+Z_{T1})(Z_{22}+Z_{T2})-Z_{12}Z_{21}|^2} \quad (5)$$

with Z_{ij} the two-port Z parameters of the time-varying capacitor $C(t)$, Z_{T1} = total external circuit impedance at f_s: $Z_{T1} = X_s + R_s + R_G + R_c$ and Z_{T2} = total external circuit impedance at f_u: $Z_{T2} = X_u + R_u + R_L + R_c$. Neglecting the losses of the matching inductances L_s and L_u ($R_s = R_u = 0$), the maximum transducer gain and bandwidth are reached under the matched conditions defined as follows:

$$R_G = R_L = R_c \sqrt{1 + \frac{\gamma^2}{\omega_s \omega_u C^2 R_c^2}} \quad (6)$$

$$X_s = \frac{1}{\omega_s C} \quad (7)$$

$$X_u = \frac{1}{\omega_u C} \quad (8)$$

with $\gamma = \gamma_1(1+\gamma_1^2)(1-\gamma_2)$ and $C = C_0(1-2\gamma_1^2)$
And it is given by:

$$G_t = \frac{f_p + f_s}{f_s} \frac{x}{\left[1 + \sqrt{1+x}\right]^2} \quad (9)$$

where $x = (f_p + f_s)/f_s \, (\gamma Q)^2$ with Q the effective quality factor of the time-varying capacitor, $Q = 1/(\omega_s C R_c)$. It is interesting to note that if we consider a lossless time-varying capacitor ($Q \to \infty$, i.e. $R_c = 0$), the transducer gain tends to the ratio $(f_p + f_s)/f_s$ which is in accordance with the power gain defined from the Manley and Rowe relationships (3) for the noninverting case.

Under the matching conditions defined above, the maximum bandwidth is given by:

$$BW = \gamma \sqrt{2 f_u f_s} \quad (10)$$

III. MEMS PARAMETRIC UP-CONVERTER

The parametric effect results from a time-varying capacitor which consists of a metallized membrane (top electrode) suspended above a heavily doped silicon bulk (bottom electrode). By pumping the movable top electrode with a large signal voltage at the pump frequency (ω_p), one generates a time-varying capacitance:

$$C(t) = C_0 \left(1 + 2\gamma_1 \cos(\omega_p t + \varphi_1) + 2\gamma_2 \cos(2\omega_p t + \varphi_2) + \ldots\right) \quad (11)$$

The fabrication process is summarized in Fig. 2 and more details can be found in [8].
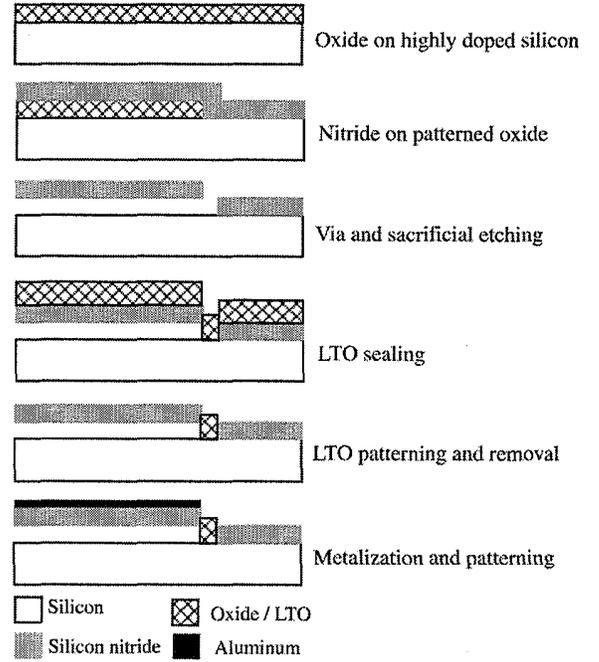
Figure 2. *Fabrication steps of the MEMS capacitor array.*

A. Static Analysis

The external load in that case is an electrostatic force F_e induced by the source voltage V_{dc} applied between the two electrodes of the MEMS capacitor. Fig. 3 shows the capacitance variation of the MEMS capacitor versus dc bias conditions. The suspended nitride membrane is 0.6 μm thick, 50 μm radius and with a residual stress of 170 MPa. The total device area is 1 cm². The calculated collapse potential for the measured structure is 85 V. It is important to work under that limit dc potential to avoid the collapse of the membrane.

B. Small-Signal Analysis

In the linear case, we assume that the microstructure displacements (x) are small compared to the physical gap dimensions of the structure (g_0). With other words, if the applied external force has a sine wave form, the movement of the mass will also be sinusoidal at the same frequency (linear system). The linear analysis is useful for determining the mechanical resonant frequency and quality factor of the movable structure versus frequency.

Under small-signal conditions, the capacitor is driven by a small ac voltage v_{ac} superimposed on the bias (V_{dc}) which induces a small ac displacement x. The dynamics of the resonator under small signal conditions are approximately determined by the second order ordinary differential equation:

$$m\ddot{x} + b\dot{x} + kx = f_e \quad (12)$$

where m is the mass of the suspended diaphragm, b the damping coefficient, k the spring constant of the system and f_e the small signal ac electrostatic force expressed by:

$$f_e = \frac{\partial F_e}{\partial V_{dc}} v_{ac} = \frac{C_0 V_{dc} v_{ac}}{g_0} \quad (13)$$

Since the resulting equation is linear, we can assume that all variables can be represented as phasors yielding:

$$\hat{x} = \frac{-C_0 V_{dc}}{g_0 m \omega_0^2} \frac{\hat{v}_{ac}}{\sqrt{1 - \omega^2/\omega_0^2 + j\omega b/m\omega_0^2}} \quad (14)$$

where $\omega_0 = \sqrt{k/m}$ is the resonant frequency.
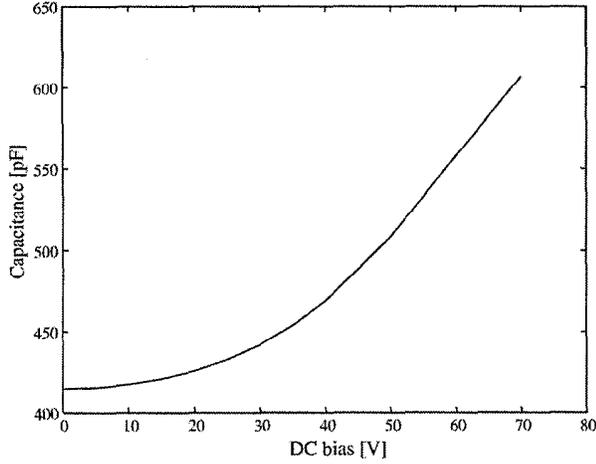
Figure 3. *Measured capacitance variation versus dc voltage.*

The device resonance can be detected by sensing the device impedance. In fact, by analogy the spring-mass-damper system can be represented by the equivalent small-signal electrical circuit in Fig 4. In this scheme R_c represents the conductor losses in the capacitor plate. The capacitor current is expressed as $\hat{i}_{ac} = (Y_0 + Y_x)\hat{\nu}_{ac}$ where

$$Y_0 = j\omega C_0 \tag{15}$$

$$Y_x = \frac{j\omega C_x}{1 - \omega^2/\omega_0^2 + j\omega C_x R_x} \tag{16}$$

From the measurement of its electrical transfer function, the mechanical properties of the resonant MEMS capacitor such as the quality factor Q and the resonant frequency are extracted. Under a dc bias of 60 V, the capacitor membrane has a resonant frequency of 1.64 MHz and a Q of 18. From the plot of the real (Fig. 5) and imaginary (Fig. 6) parts of the parameter Y_{12}, the series parasitic resistance R_c, the parallel capacitance C_0 and the equivalent electrical lumped elements of the resonator L_x, C_x and R_x are extracted (see Figs. 5 and 6).
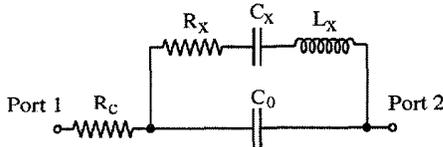
Figure 4. *Small-signal electrical model of the MEMS capacitor.*

C. Large Signal Analysis

If a large pump signal voltage is applied, the displacements of the suspended diaphragm become non-negligible compared to the default gap (g_0), then non-linear behaviors appear and a time-varying capacitance $C(t)$ is generated. In order to minimize the magnitude of the pump signal voltage and then the power consumption, the pump frequency f_p is chosen equal to the nature resonant frequency of the capacitor plate. More the displacements are large more the values of the Fourier series coefficients γ_n are important. Their extraction is an essential step for designing a parametric amplifier/up-converter. Two extraction methods have been developed depending on the importance of the parasitic conductor series resistance R_c of the MEMS capacitor plate. In fact, if R_c is small such as $j\omega R_c C_0 << 1$, coefficients φ_n tend to zero and γ_n can be extracted directly from the current spectrum

magnitude given by a spectrum analyzer. If $j\omega R_c C(\omega) << 1$, with other words, if R_c can be neglected compared to $(j\omega C(\omega))^{-1}$, the current through the time-varying MEMS capacitor, when a voltage signal $v(t) = V_{dc} + v_{ac} \cos(\omega_p t)$ is applied, can be expressed in the frequency domain by:

$$I(\omega) = j\omega(C(\omega) * V(\omega)) \tag{17}$$

Solving the convolution product, we obtain the following expressions for the three first harmonics:

$$I(\omega_p) = v_{ac} j\omega_p \pi C_0 (1 + \gamma_2) + V_{dc} j\omega_p 2\pi\gamma_1 C_0 \tag{18}$$

$$I(2\omega_p) = v_{ac} j2\omega_p \pi C_0 (\gamma_1 + \gamma_3) + V_{dc} j2\omega_p 2\pi\gamma_2 C_0 \tag{19}$$

$$I(3\omega_p) = v_{ac} j3\omega_p \pi C_0 \gamma_2 + V_{dc} j3\omega_p 2\pi\gamma_3 C_0 \tag{20}$$

In that case, from the current magnitude measured with a spectrum analyzer and the analytical relations (18)-(20), the Fourier series coefficients γ_n are extracted. On the other hand, if R_c is not negligible the current spectrum is not sufficient to extract the coefficients γ_n and φ_n of $C(t)$ (11). Actually, in that case we need a measurement of the current phase through the MEMS capacitor. This current phase information can be provided by the measurement of the voltage $v(t)$ across $C(t)$ and the current $i(t)$ flowing through it with a digital scope as explained below.

The current $i(t)$ is obtained by measuring the voltage across a external resistor R_{ext} connected in series with the time-varying capacitor. Knowing the value of the parasitic resistor R_c (extracted from the small-signal analysis) the voltage across $C(t)$ is given by:

$$v(t) = v_t(t) - i(t)(R_{ext} + R_c) \tag{21}$$

Then the time-varying capacitance can be expressed as:

$$C(t) = \frac{i(t)dt + C_0 V_{dc}}{v(t)} \tag{22}$$

Finally, considering the extracted numerical function $C(t)$ as a periodic time function, it can be decomposed into Fourier series for determining the coefficients γ_n and φ_n. Fig. 7 shows the measured harmonics level at the 50 Ω input of the spectrum analyzer for the MEMS time-varying capacitance pumped with an ac large signal $v(t) = 60 + 15 \cos(\omega_p t)$ with $\omega_p = 1.64$ MHz. Unfortunately, due to the pretty high value of $R_c = 50 \Omega$ compared to the equivalent impedance $((j\omega C(\omega))^{-1} = 400 \Omega)$ of the MEMS capacitor, it is not possible to characterize accurately its non-linear behavior from the only measurement of its current spectrum. In that case, we have to use the second extraction method based on the measurements of voltages versus time with a scope. From the measured characteristic $C(t)$, the coefficients γ_n and φ_n extracted are: $\gamma_1 = 0.22$, $\gamma_2 = 0.05$, $\gamma_3 = 0.01$, $\varphi_1 = 7°$, $\varphi_2 = 3°$ and $\varphi_3 = 0.85°$.

From the measured characteristics $(\omega_0, Q, \gamma_n, \varphi_n,$ etc.) of the time-varying capacitor, the input and output matching networks and filters are designed to maximize the gain of the built MEMS parametric up-converter. Fig. 8 shows the measured and simulated transducer gain of the parametric up-converter versus the load resistance. A good agreement is obtained between measurements and simulations. It is interesting to note that the optimum transducer gain achieves effectively the available power gain G_u=1.84 MHz/200 kHz = 9.6 dB when R_c tends to zero. The measured optimum load of 140 Ω is in accordance with the value (133 Ω) calculated with the analytical relationship (6). Fig. 9 presents the measured transducer gain versus the coefficient γ_1 which depends on the pump signal magnitude. For these measurements, the matching networks and input/output filters has been optimized for $\gamma_1 = 0.22$. As predicted by the simulations, the transducer gain of the MEMS parametric up-converter increases with γ_1.

IV. CONCLUSION

Parametric amplifiers can be built for sensors (thermal, acceleration, chemical, etc.) without any CMOS electronics. This means that a pre-amplification stage ca be provided for non CMOS sensors before sending the signal to a CMOS wafer for more gain and processing.
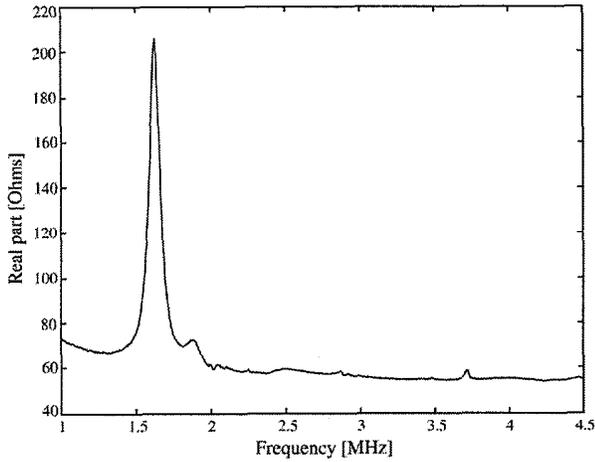
Figure 5. *Measured real part of 1/Y_{12} versus frequency. Extracted value of R_c is 52 Ω.*
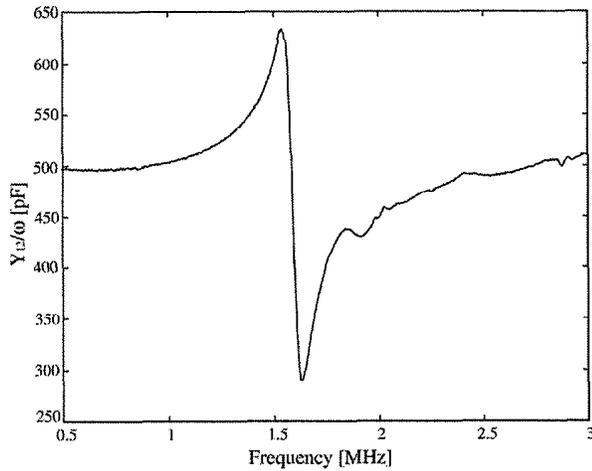
Figure 6. *Measured equivalent capacitance (Y_{12}/ω) around the resonant frequency. Extracted values of C_0, C_x, L_x and R_x are, respectively, 500 pF, 32.5 pF, 0.29 mH and 166 Ω.*

V. REFERENCES

[1] George L. Matthaei, *IRE Trans. on MTT*, pp. 23-28, Jan. 1961.

[2] E. Sard et al, *IEEE Trans. on MTT*, pp. 608-618, Dec. 1966.

[3] Y. J. Ding and J. B. Khurgin, *IEEE J. of Quantum Electronics*, pp. 1574-1582, Sept. 1996.

[4] G. M. Gal et al., *IEEE J. of Selected Topics in Quantum Electronics*, pp. 224-229, Feb. 1998.

[5] J. M. Manley and H. E. Rowe, *Proceedings of the IRE*, pp. 904-914, July 1956.

[6] H. E. Rowe, *Proceedings of the IRE*, pp. 850-860, May 1958.

[7] L. A. Blackwell and K. L. Kotzebue, "Semiconductor-diode parametric amplifiers", *Prentice-Hall, Inc.*, N.J., 1961.

[8] Xuecheng Jin, et al., *IEEE Journal of Microelectromechanical Systems*, vol. 8, no. 1, pp. 100-114, March 1999.
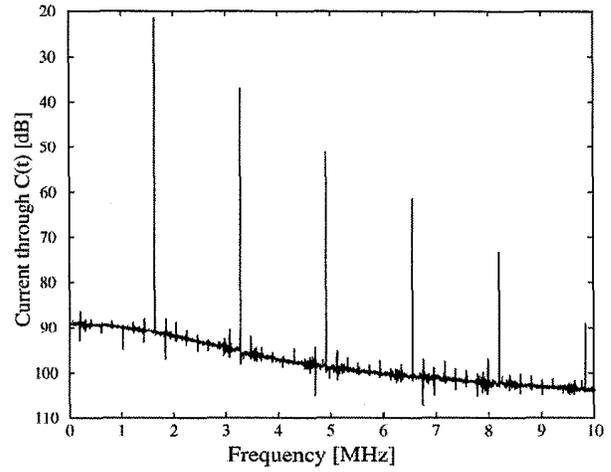
Figure 7. *Measured spectrum of the current flowing through the MEMS time-varying capacitor for V_{dc} = 60 V and v_{ac} =15 cos(2πf_p t), with f_p = 1.64 MHz.*
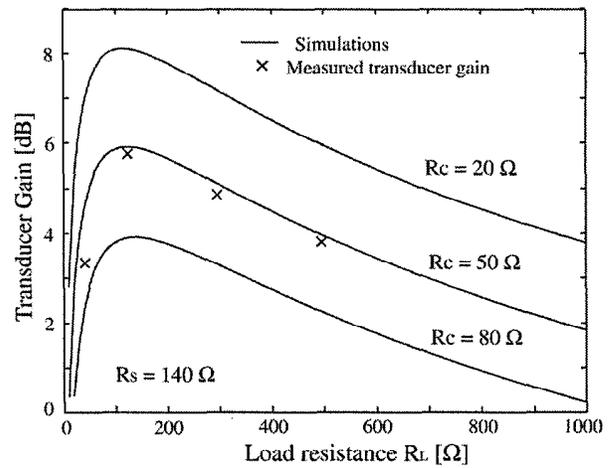
Figure 8. *Measured parametric up-converter transducer gain versus the load resistance value (f_s = 200 kHz and f_p = 1.64 MHz).*
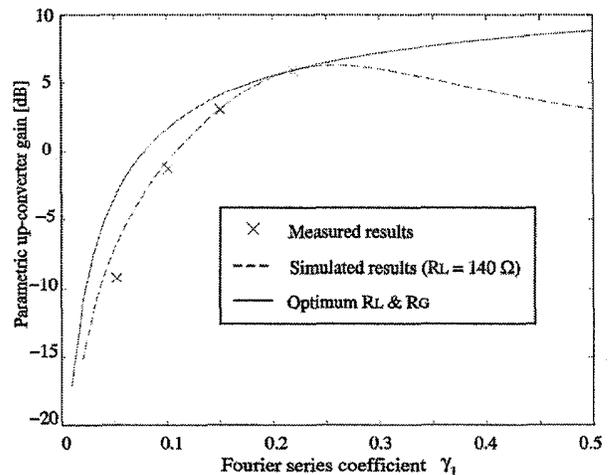
Figure 9. *Measured parametric up-converter transducer gain versus γ_1 (f_s = 200 kHz and f_p = 1.64 MHz).*

SURFACE MICROMACHINED PIEZOELECTRIC
RESONANT BEAM FILTERS

Brett Piekarski,[1,2] Don DeVoe,[2] Madan Dubey,[1] Roger Kaul,[1] John Conrad,[1] Robert Zeto[1,*]

[1] U.S. Army Research Laboratory Adelphi, Maryland 20783-1197

[2] Dept. of Mechanical Engineering and Institute for Systems Research
University of Maryland College Park, Maryland 20742

ABSTRACT

Piezoelectric actuation and sensing of suspended single-beam microelectromechanical system (MEMS) resonant filters is demonstrated. Resonant frequencies between 171 kHz and 9.2 MHZ were observed for resonators with dimensions of 6- to 20-μm wide and 25- to 400-μm long. Operation was demonstrated from a vacuum pressure of 20-mTorr to atmospheric with drive voltages as low as 2.5 mV and a dynamic range of greater than 40 dB. Loaded quality factors (Qs) greater than 1000 resulted at 9.2 MHz and a 20-mTorr pressure. The resonators utilize a clamped-clamped beam design with sol-gel lead zirconate titanate (PZT) as the piezoelectric actuator and sensor, PECVD silicon dioxide as the resonant beam structure, and platinum as the top and bottom electrode.

INTRODUCTION

Due to their small size, low power consumption, and ability to be integrated with microelectronics, MEMS resonators are becoming viable for use as filters in communication systems. To date, these efforts have focused primarily on using electrostatic transduction for actuation and sensing of polysilicon and silicon resonant beams [1-5].

The theoretical coupling strength of electrostatic devices has been shown to decay rapidly as resonant frequencies are increased. Piezoelectric transduction offers an alternative mechanism for micromachined resonant clamped-clamped beams that avoids this difficulty, since piezoelectric coupling strength is significantly higher than electrostatic coupling and decays more slowly as beam dimensions are decreased to increase the resonant frequency [6]. Reduction in coupling strength is slower because capacitive actuation results from a distributed force whereas piezoelectric actuation results from a distributed moment independent of the beam length. Improved coupling can potentially lead to improved performance and reduced noise.

Piezoelectric resonators demonstrated with ZnO as the piezoelectric actuation and sense material were performance limited by the low piezoelectric coefficients of ZnO [6]. PZT has much higher piezoelectric coefficients and while devices using PZT as the piezoelectric material have been used as actuators and force sensors [7-12], little work has been reported on thin-film PZT resonators for communication systems. Here, we demonstrate suspended single-beam piezoelectric clamped-clamped micromechanical resonators using sol-gel-deposited PZT as the active piezoelectric for both sensing and actuating beam resonance.

RESONATOR STRUCTURE AND FABRICATION

Figure 1 shows a top view and cross section of the clamped-clamped beam resonator design. The starting wafer is fabricated by first depositing a PECVD SiO_2 layer, underlies the entire structure, on a Si wafer.
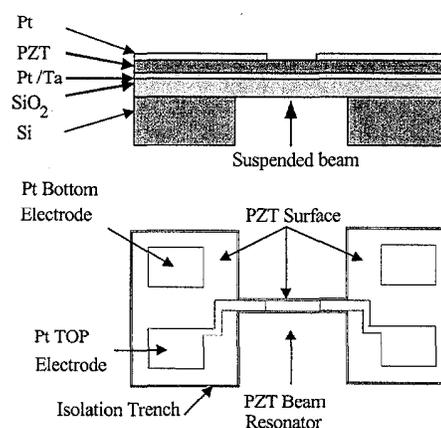
Figure 1. Resonator design.

A sputtered tantalum adhesion layer is then deposited on top of the oxide, followed by a sputtered platinum bottom electrode that also underlies the entire structure. The piezoelectric material is a sol-gel $PZ_{52}T_{48}$, spun on and pyrolyzed until the desired PZT thickness is achieved. A sputtered platinum top electrode layer is then deposited over the PZT. Table 1 shows the thickness of each layer used in the reported devices.

Table 1. Material thickness.

Material	Thickness (μm)
Silicon Dioxide	2
Tantalum	0.02
Bottom Platinum	0.17
PZT	0.5
Top Platinum	0.17

Resonator fabrication starts by defining the top electrode pattern through photolithography and ion milling away the unwanted platinum. An opening is then created in the PZT by wet etching to provide access to the bottom electrode. The beam length and width are then defined, along with an electrical isolation

trench, by ion milling through the PZT, bottom electrode, and SiO₂ to expose the underlying silicon. The final step is an isotropic silicon reactive-ion etch (RIE) that undercuts the beam geometry and results in the suspended clamped-clamped beam. Additional photolithography can be done prior to the isotropic silicon RIE step so that the isotropic etch undercuts only the photoresist and not the beam anchors while the beam is undercut. Figures 2 and 3 are SEM micrographs of a 50-μm long beam and show the isolation trench, probe contact pads, isotropically etched trench, and the quarter length top electrodes.
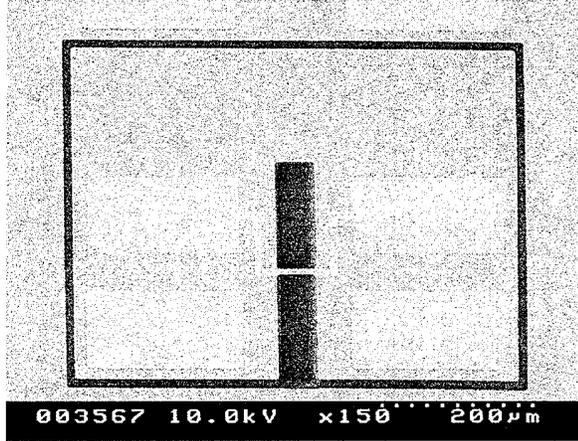
Figure 2. Top view of 50-μm long resonator.
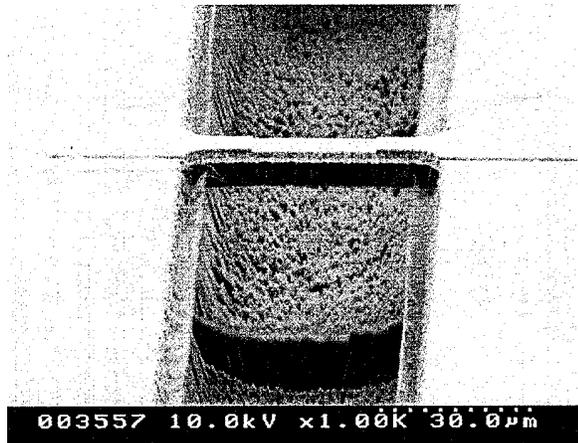
Figure 3. 50-μm beam over 30-μm deep trench.

A detailed description of the PZT material properties and resonator fabrication process used for the reported devices can be found in references [13-15]. The only significant change from previously reported fabrication techniques is that the devices reported here were suspended from the top by an isotropic RIE process, whereas previous techniques used Si deep reactive-ion etching to release the structures from the backside.

Resonator Operation

The constitutive equations for the plane stress case with an electric field applied in the z-axis can be shown in the following reduced tensor form [16].

$$S_1 = s^E_{11}T_1 + d_{31}E_3$$
$$D_3 = d_{31}T_1 + \varepsilon^T_{33}E_3$$

In this equation, S is the strain, s is elastic compliance, T is stress, d is piezoelectric coefficient, E is electric field, D is electric displacement, and ε is the permittivity.

The input drive voltage generates an electric field between the top and bottom electrode and induces a strain in the drive side of the resonator through the direct piezoelectric effect. This strain results in a concentrated moment applied at the terminal edge of the input electrode. A bending moment results from the offset of the piezoelectric from the beam neutral axis. When the input voltage frequency is near the fundamental frequency of the beam, resonance occurs in the beam amplifying the piezoelectric strain on the sense side of the resonator. This strain induces a current in the sense circuit at the resonant frequency of the beam through the converse piezoelectric effect.

Figure 4 shows the equivalent circuit [6] for the resonator structure which consists of three components: a drive circuit, a sense circuit, and a feed-through capacitance. The middle ground is the bottom electrode and the input to the drive circuit and the output to the sense circuit are the top electrodes. In the circuit, C_p is the static capacitance, C_f is the feed through capacitance, and the L_x, R_x, and C_x branch is the motional arm of the circuit that produces a series resonance caused by the mechanical vibrations. The mechanical vibration from the drive circuit induces a current in the sense circuit that is reduced by the mechanical coupling factor χ_{21}.
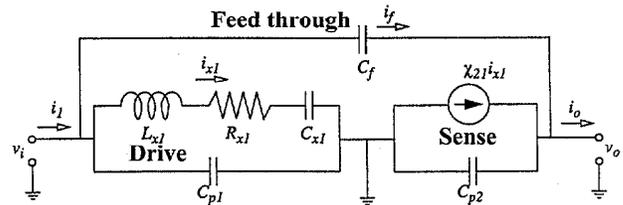
Figure 4. Equivalent circuit.

EXPERIMENTAL SETUP

The experimental setup included an HP3577 vector network analyzer, an MMR vacuum probe station, and a computer for electronic data storage. The HP3577 network analyzer provided the drive signal and recorded the scattering parameter S_{21} (ratio of the output to the input, assuming that the output is matched) magnitude and phase for the resonators. The MMR vacuum probe station was used for contacting the top and bottom electrodes. A single ground probe was used since the bottom electrode extends under the entire device. The two remaining probes were placed on the drive and sense top electrodes. The MMR probe station also included a vacuum kit that allowed pressures to be varied from 20 mTorr to ambient. A CLC 109 unity gain operational amplifier was connected at the output port of the MMR probe station due to the impedance mismatch between the 50-Ω test setup and the high impedance of the devices. HP VEE5.0 software was used to dump the data to a text file for electronic storage and analysis.

EXPERIMENTAL RESULTS

Table 2 shows the dimensions of the reported devices and their designed and actual operating frequencies. There was some variation from the designed theoretical value. We believe that much of the difference was caused by the residual stresses in the clamped-clamped beam structure and the fact that the platinum layers were not considered in the theoretical calculation of the designed frequency and some bulk material properties were used for the calculation.

Table 2. Theoretical and actual resonant frequencies.

Length (μm)	Width (μm)	Designed frequency	Actual frequency
400	20	59 kHz	171 kHz
200	20	236 kHz	435 kHz
80	16	1.5 MHz	1.9 MHz
50	10	3.8 MHz	3.6 MHz
25	6	15 MHz	9.2 MHz

The stress in the overall oxide/Pt/PZT/Pt stack prior to fabrication was measured to be 80 MPa tensile, but at the time of this paper, the actual stress gradient for the various length devices was unknown. An overall tensile stress in the beam should lead to an increase in the resonant frequency as given by the following relation for the fundamental bending mode of a beam subjected to an axial force N [17].

$$\omega_n = \left(\frac{4.73}{L}\right)^2 \sqrt{\frac{EI}{\rho L}\left(1 + 0.295\frac{NL^2}{12EI}\right)}$$

In this equation, ω_n is the natural resonant frequency, L is beam length, E is effective beam elastic modulus, I is effective beam moment of inertia, ρ is density, A is cross-sectional area, and N is the applied force. This does explain the lower than designed frequency of the 50- and 25-μm devices and studies are ongoing.

Figure 5 shows the S_{21} response of an 80-μm long resonator operating in a 20-mTorr vacuum. Its resonant frequency is 1.87 MHz with a 90 degree phase shift at the resonant peak and a measured loaded Q of 1212. When the pressure was raised to atmospheric, the loaded Q reduced to 430 and there was a flattening of the resonant peak due to the increased damping caused by the air flow around the beam. Figure 6 shows the same device operating in air but over a 200-kHz span.
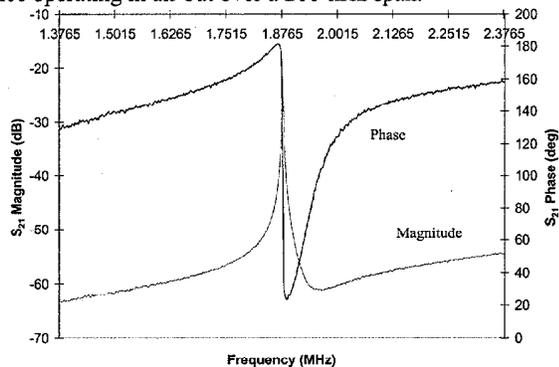
Figure 5. 80-μm resonator at 20 mTorr.
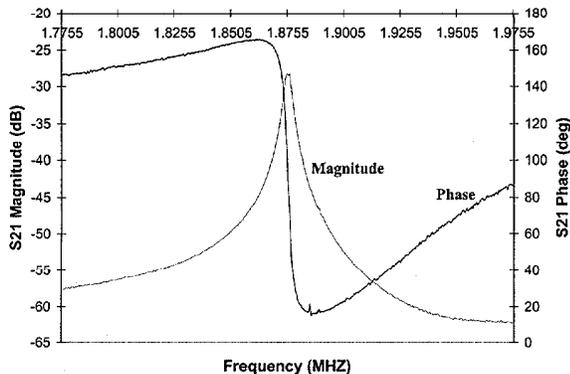
Figure 6. 80-μm resonator at atmosphere.

This same effect was observed for the 400-, 200-, and 25-μm long resonators. Table 2 shows the resonant frequencies of these devices and figure 7 shows the loaded Qs for both the in-air and under 20-mTorr vacuum measurements. The loaded Qs in air appear to be fairly constant between 400 and 500 over the range of devices tested. The loaded Qs for the devices tested under vacuum are much higher (3400 at 170 kHz). The Q for the in-vacuum measurements decreases rapidly at low frequencies but then levels off at approximately 1000 as the frequency is increased above 1 to 3 MHz.
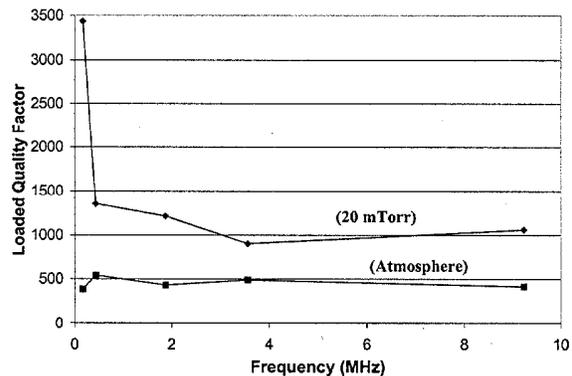
Figure 7. Loaded Qs for resonators measured in air and at 20-mTorr.

Figure 8 shows the effect of the vacuum pressure on the S_{21} measurements for the same 80-μm resonator. As the device is brought under vacuum, the Q and magnitude of the resonant peak increase until about 1 Torr is reached and then no further increase is observed through 20 mTorr. These data show that the resonators work well in air but that a vacuum of less than 1 Torr is required for optimum performance.
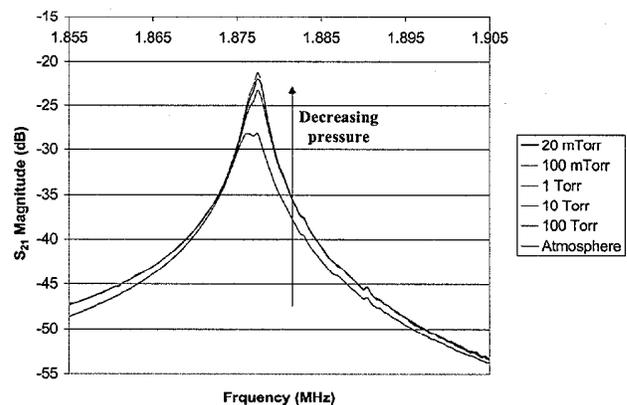
Figure 8. Effect of vacuum pressure on S_{21} Magnitude

The dynamic range was tested by varying the drive voltage from 2.5 mV (minimum voltage the test setup could apply) to 0.8 V. Figure 9 shows the results for the 80-μm resonator. Evidence of overdriving the resonators was not observed until above 0.25 V leading to a minimum dynamic range of 40 dB. The resonators do not require a bias voltage and they are not restricted by any geometrical configuration like the electrode gap for electrostatic devices.
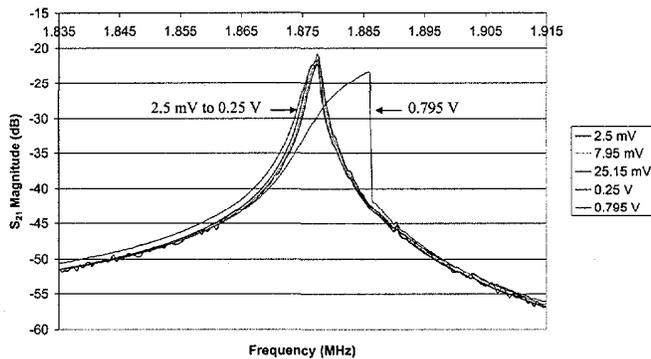
Figure 9. Effect of Drive Voltage on Resonance

Measurements were made without any signal amplification and losses ranged from −19 dB at 171 KHz to −48 dB at 9.2 MHz. The large probe pads (100 X 200 microns) used in the design resulted in a considerable shunt capacitance to ground. This capacitance induced a large parallel resonance due to finite conductance in the ground electrodes and reduced the signal to noise ratio and obscured the signal at high frequencies. To reduce this effect, large portions of the probe contact pads were removed by running the probe tips across them prior to testing. This resulted in a slight variation in the top electrode area from test to test and a variation in the noise floor for each test. Significant increase and uniformity in the off-frequency rejection is expected to be obtained in the next generation devices, especially at high frequencies, by reducing the contact area.

SUMMARY

Suspended single-beam microelectromechanical system (MEMS) resonant filters using sol-gel PZT for piezoelectric transduction were demonstrated up to 9.2 MHz with 90 degrees of phase shift at resonance and loaded quality factors greater than 1000 under a 20-mTorr vacuum and greater than 400 in air. Qs up to 3400 were obtained at 171 kHz and the minimum vacuum required for optimum operation was 1 Torr.

The resonators do not require a bias voltage and they are not restricted by any geometrical configuration like the electrode gap for electrostatic devices. The minimum drive voltage used was 2.5 mV which is the test setup minimum and they exhibited a minimum dynamic range of 40 dB.

ACKNOWLEGEMENTS

The authors would like to thank Dr. Matthew Ervin for his usual excellent SEM work and Eugene Zakar and Ronald Polcawich for their help in the starting wafer fabrication and bulk film-stack stress measurements.

REFERENCES

[1] C.T.-C. Nguyen, "Micromechanical resonators for oscillators and filters," *Proceedings IEEE Int. Ultrasonics Symp.*, Seattle, WA (1995) pp. 489-499.
[2] C.T. Nguyen, "Micromechanical Devices for Wireless Communications," *Proceedings of the 1998 IEEE Micro Electro Mechanical Systems Workshop*, Heidelberg, Germany, (1998).
[3] L. Lin, R.T. Howe, A.P. Pisano, "Microelectromechanical Filters for Signal Processing," *Journal of Microelectromechanical Systems*, 7, (1998).

[4] J.W. Weigold, S. W. Pang, "Fabrication of Thick Si Resonators with a Frontside-Release Etch-diffusion Process," *Journal of Microelectromechanical Systems*, 7, (1998).
[5] J.J. Yao, N.C. MacDonald, "A Micromachined, Single-Crystal Silicon, Tunable Resonator," *J. Micromech. Microeng.* 6 (1996).
[6] D.L. DeVoe, "Thin Film Zinc Oxide Microsensors and Microactuators," Ph.D. Dissertation, University of California, Berkeley, Fall (1997).
[7] P. Luginbuhl, *Sensors and Actuators*, 54 (1996).
[8] C. Lee, *Trans. Ultrasonics, Ferroelectrics, and Frequency Control*, 43 (1996).
[9] M.A. Dubois, *Trans. Ultrasonics, Ferroelectrics, and Frequency Control*, Sept. (1998).
[10] P. Muralt, *Sensors and Actuators*, 53 (1996).
[11] E. Zakar, M. Dubey, B. Piekarski, J. Conrad, R. Piekarz, R. Widuta, "Process and Fabrication of a Thin Film PZT Pressure Sensor," Presented at the American Vacuum Society 46th International Symposium and to be published in *J. Amer. Vac. Soc.*, (2000).
[12] E. Zakar, M. Dubey, B. Piekarski, R. Piekarz, J. Conrad, R. Widuta, "Study of PZT Film Stress in Multilayer Structures for MEMS Devices," *proceedings of MRS 1999 Fall Symposium*, Boston, MA (1999).
[13] S. Trolier-McKinstry, "Manufacturable Sol-Gel PZT Films for Microsensors and Microactuators," *Final Technical Report, U.S. Army Contract DABT63-95-C-0053*, (1998).
[14] R.J. Zeto, B.J. Rod, M.H. Ervin, R.C. Piekarz, S. Trolier-McKinstry, T. Su, and J.F. Shepard, "Dry Etching of Sol-Gel PZT," *Proceedings of Material Sci of MEMS Devices*, 546, pg 159-164, (1998).
[15] B. Piekarski, M. Dubey, D. DeVoe, E. Zakar, R. Zeto, J. Conrad, R. Piekarz, M. Ervin, "Fabrication Of Suspended Piezoelectric Microresonators," *Integrated Ferroelectrics*, (1999), 24, pp 147-154.
[16] T. Ikeda, *Fundamentals of Piezoelectricity*, Oxford Science Publications, New York NY (1996).
[17] H. Tilmans, M. Elwenspoek, J. Fluitman, "Micro resonant force gauges," *Sensors and Actuators A*, 30, 35-53, (1992).

A LOW POWER / LOW VOLTAGE ELECTROSTATIC ACTUATOR FOR RF MEMS APPLICATIONS

J. Jason Yao*, SangTae Park, Robert Anderson, and Jeffrey DeNatale
Rockwell Science Center, 1049 Camino Dos Rios, Thousand Oaks, CA 91360
*Voice: (805) 373-4471, FAX: (805) 373-4869, E-Mail: *jjyao@rsc.rockwell.com*

ABSTRACT

We have designed and fabricated a low power / low voltage electrostatic actuator, and demonstrated its application to a large tuning-ratio tunable capacitor for RF MEMS applications. Using adhesive bonding and deep silicon reactive ion etching techniques, the entire MEMS device is made of single crystal silicon, and is suspended over a glass substrate. A coat of aluminum is sputtered on after the dry release in an oxygen plasma to provide good electrical conductivity for the integrated RF devices. An electrostatic deflection of 23 μm has been demonstrated with an applied voltage of 5.2 V, and has resulted in the tunable capacitor having a maximum continuous tuning ratio in excess of 4.5 to 1. Alternative devices with a required low tuning voltage of 3 V have also been demonstrated to provide a tuning ratio of 2 to 1. The power consumption of this actuator is linearly proportional to the actuation frequency, and is in the range of 10's of nano-Watts when the device is actuated at a frequency of a few kHz.

INTRODUCTION

Electrostatics is one of the favorable means of actuation at the reduced size scale typical of MEMS. The major attractive characteristics include simplicity in operation and device construction, low power consumption, and relatively low sensitivity to temperature variations. Various alternative actuation mechanisms have been demonstrated in MEMS, and are being used in their targeted niche applications. However, tradeoffs are evident in each of those micro actuation mechanisms. Electromagnetic actuation provides a large force because of the high energy density in magnetic materials, but is often complex to construct and integrate, and requires a significant current level (typically mA's) to operate. Electro-thermal actuation also provides a relatively large force, but in general is highly temperature dependent and requires a non-negligible amount of current to operate, as well. Depending upon the size of thermal mass and various other design parameters, electro-thermal actuators can also be limited to an operation frequency below one kilo-Hertz. Electrostatic actuators require a very low amount of current (typically nA's), but, in the past, often required a relatively high actuation voltage (e.g., > 10 V) while still outputting a significantly smaller force. However, the low power-consumption nature of the electrostatic actuators makes it an attractive candidate for many applications, such as wireless communications, especially where the actuation force is not a critical requirement. Tunable capacitors or varactors used in many communications circuits are an example of such applications where low power consumption is a critical requirement. By massively integrating electrostatic actuators, we have been able to lower the required electrostatic actuation voltage to approximately 5 V, and for some designs to a voltage below 3 V, and still maintain a required current level in the nano-Amperes range.

Tunable capacitors or varactors are widely used in RF communications for low-noise parametric amplifiers, harmonic frequency generators, and frequency controllers such as voltage controlled oscillators (VCO). Conventional solid-state varactors are made using either p-n or Schottky-barrier junction structures. These solid-state components [1] often suffer from a small tuning ratio (< 30%), excessive resistive loss and thus a low Q (typically less than 10), and a low electrical self-resonance caused by large parasitics. Drastic measures are often required for the solid-state components to improve on one single parameter; for example, the use of light power to increase the tuning ratio [2]. Recent efforts within the MEMS field [3-9] have shown promising results in the realization of MEMS-based high performance tunable capacitors. Many shortcomings of solid-state varactors have indeed been overcome in these MEMS-based devices. Parameters already improved using the MEMS approaches include equivalent series resistance and Q factor, capacitance tuning-ratio, associated parasitic capacitance, and intrinsic inductance and electrical self-resonance. This paper addresses the low actuation power and voltage requirement of the tunable capacitor device, and discusses in detail a massively parallel electrostatic actuator for RF MEMS applications.

DEVICE DESIGN AND FABRICATION

There are two general approaches in the MEMS-based tunable capacitor design – area-tuning interdigitated "finger"[3] and gap-tuning parallel plate [4-9] design. Each of the designs has advantages and disadvantages. A general comparison between the two was presented in [3], and the design selection is heavily guided by their targeted applications, which have very specific and often conflicting requirements on the device parameters. In general, the gap-tuning capacitors can be constructed using parallel plates stacked on top of a substrate using surface micromachining techniques. This approach provides an easy means to integrate the tunable capacitors with other conventional IC electronics. However, as in any electrostatic parallel plate system, the movable plate can only be deflected one-third of the initial gap distance before the two plates snap together as the electrostatic attracting force becomes greater than the mechanical spring back force. This deflection limitation results in a theoretical limit of 50% on the capacitance tuning range for a parallel plate tunable capacitor. Efforts have been reported in [5] to increase this 50% tuning limit to 100% by stacking three plates vertically, so that there is an actuator on each side of the middle movable plate to both increase and decrease the overall capacitance value. Nevertheless, the tuning range still comes short for many RF applications. Another inherent characteristic in a gap-tuning capacitor is the dependence of the device's RF power handling capability on the gap spacing between the two electrodes. As the capacitor is tuned higher, the gap is reduced; RF breakdown in the capacitor air-gap is thus more likely to occur. This aspect of the device must be taken into consideration when designing and using the device.

The area-tuning interdigitated "finger" capacitors do not have any theoretical tuning range limitation, as long as the fingers are designed long enough for the tuning to occur, and the springs are deflected within the material boundaries. The RF power handling capability of an area-tuning capacitor is independent of the tuned capacitance because the gap between the finger electrodes remains constant. Therefore, the area-tuning design is selected in the tunable capacitor project presented here. The interdigitated "finger" structures can be fabricated using surface micromachining technologies such as the polysilicon process. Although some thick polysilicon (e.g., 7 μm) processes have been demonstrated, typical polysilicon thickness is limited to 2 μm. To achieve a higher capacitance density per unit area, higher aspect

ratio structures must be used. A silicon-on-insulator (SOI) based process was presented in [3] to allow fabrication of high aspect-ratio single crystal silicon structures on silicon substrate. However, substrate losses and various parasitic effects due to the use of silicon substrates have introduced a significant and negative impact to the device's RF performance. In this work, silicon substrates are consequently replaced by low-loss glass substrates, resulting in higher RF performance for the signal capacitor while maintaining similar or improved performance characteristics for the electrostatic actuator.

The device structure is a 20 μm silicon layer bonded to glass substrate using epoxy as an adhesive. Figure 1 schematically illustrates the process sequence. The silicon layer is originated from a SOI substrate (fig. 1(a)) after removing the handle wafer side of an SOI along with the buried oxide. On top of the silicon device layer, 2 μm aluminum is deposited (Fig. 1(b)), mainly to improve the quality factor or equivalent series resistance. Both aluminum and silicon are patterned using a single mask. The aluminum is etched by reactive ion etch (RIE) using chlorine chemistry and silicon is etched by Inductively Coupled Plasma (ICP) using the Bosch process (Fig. 1(c)). The device structure is then released in an oxygen-plasma etcher (Fig. 1(d)). After the structures are fully released, a thin layer (~ 0.25 μm) of aluminum is sputtered to further improve the quality factor (fig. 1(e)). Due to the isotropic etching of epoxy, electrical isolation is naturally created between structures during the aluminum sputter deposition, and no additional mask is needed for metal patterning.

Figure 2 contains three scanning electron micrographs, showing perspective views of a MEMS-based tunable capacitor on a glass substrate. The device is comprised of two mechanically joined suspended capacitors with the center suspended mechanical attachment grounded electrically. One movable capacitor is for electrostatic actuation and the other for RF signals. The mechanical attachment between the actuator and the RF capacitor enables the translation of movements provided by the actuator to the RF capacitor to realize the capacitance tuning by adjusting the physical "finger" overlapping dimension of the RF capacitor. It should be noted that the tuning can be conducted in a non-vacuum environment with a small tuning voltage of a frequency from DC up to the mechanical resonance of the structures (~10 kHz). No resonant mode of operation was used for the capacitance tuning.

Figure 1. A schematic illustration shows the process sequence for making a single crystal silicon tunable capacitor with an integrated electrostatic actuator on a glass substrate.
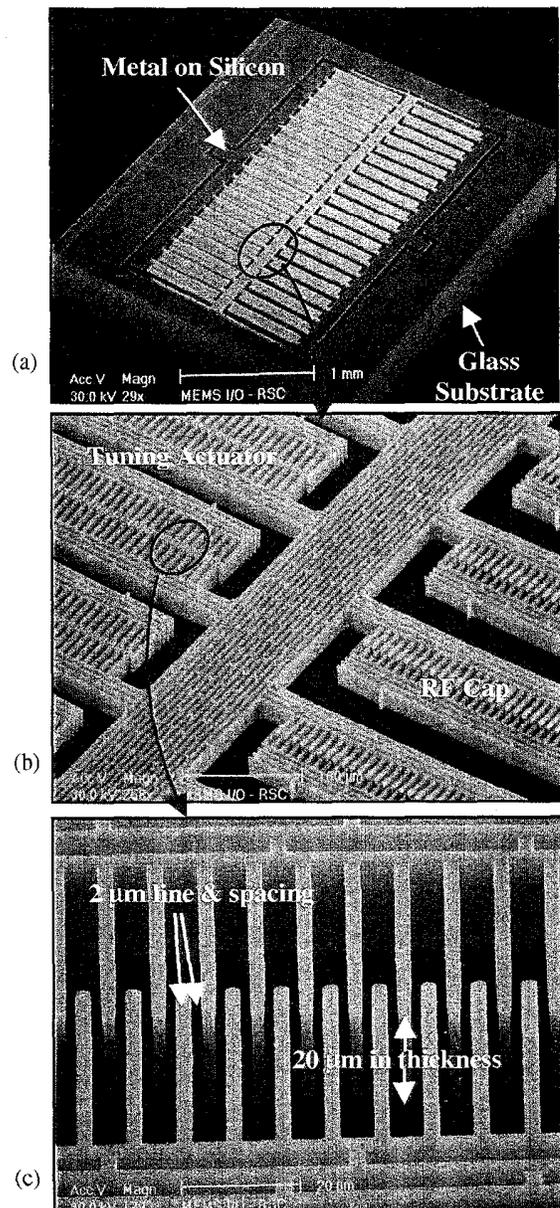
Figure 2. Scanning electron micrographs show the perspective views of a MEMS-based tunable capacitor at three different magnifications. The high aspect-ratio and the anisotropic side-wall profile provided by ICP are critical to the performance of the tunable capacitor.

247

The electrostatic actuator consists of an average of 1200 sets of interdigitated fingers, which are nominally 20 μm tall, 2 μm in width, and 2 μm spaced. These dimensions result in an actuator capacitance on the order of 2.2 pF at 0 V. The supporting spring designed for the device has a spring constant on the order of 0.5 N/m. The actuator is capable of providing a maximum travel distance of 30 μm, and occupies a nominal area of 2 mm².

EXPERIMENTAL RESULTS

The device was first tested with a DC tuning voltage to observe visual movements. Continuous tuning has been observed by varying the tuning voltage, and no resonant operational mode is required. Figure 3 shows a top view of a portion of the tunable capacitor device with the tuning voltage at (a) 0 V, (b) 3 V, and (c) 5 V. The tunable capacitor is designed to mimic a solid-state varactor diode, with its capacitance decreasing as the tuning voltage increases. At a zero tuning voltage, as shown in Fig. 3 (a),
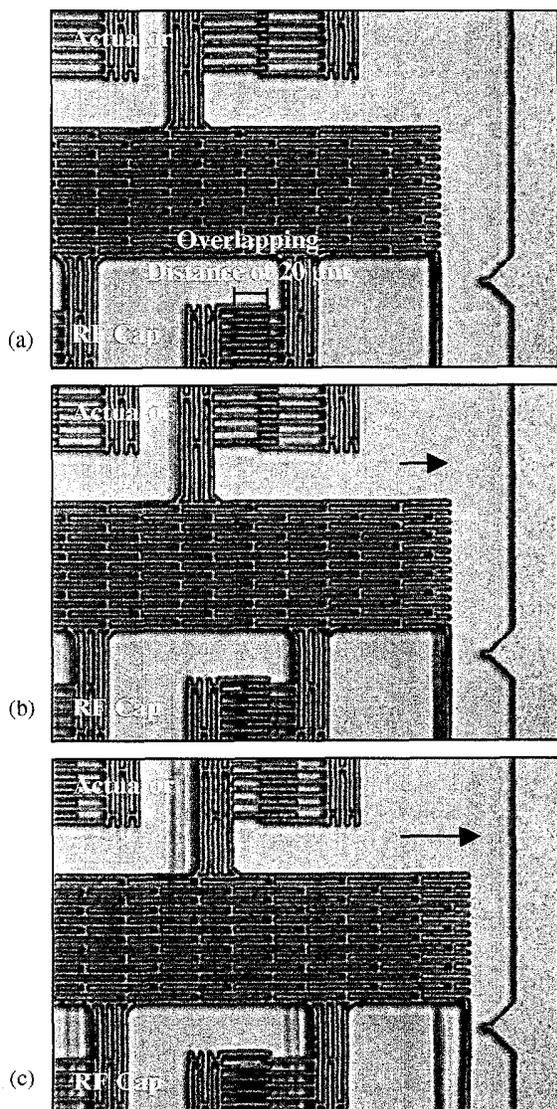
Figure 3. Top view images of a portion of the tunable capacitor (Dev. #33) with a tuning voltage of (a) 0 V, (b) 3 V, and (c) 5 V. The change in the overlapping distance between the fingers in the capacitor indicates a change in capacitance value.

the overlapping distance in the RF capacitor fingers is at its maximum of 20 μm. This overlapping distance decreases as the tuning voltage increases, as shown in Figs. 3 (b) and (c). At a tuning voltage of 5 V, the electrostatic actuator generates a total deflection about 20 μm, resulting in a zero overlapping distance in the RF capacitor fingers (Fig. 3 (c)). When further increasing the tuning voltage, the fingers in the RF capacitor become completely disengaged; the RF capacitance relies entirely on fringing fields, and decreases at a reduced rate.

The device is designed with ground-signal-ground probe pads spaced 150 μm apart. RF probes with the same pitch distance were calibrated to the probe tips, using short, load, and open configurations. The capacitance was then probed using a HP8753D network analyzer from 30 MHz to 6 GHz. All parasitic impedance associated with the tunable capacitor is included in the measurement. Figure 4 is a S11 plot in a Smith chart. The nominal equivalent series resistance is 1.3 Ω, and the electrical self-resonance is about 5 GHz at a tuned capacitance value of 2 pF. Figure 5 contains plots of the tuned capacitance values as functions of the tuning voltage for two of the device designs. Tuning ratios of 4.55 to 1 with a maximum tuning voltage of 5.2 V for Dev. #33 and 2.1 to 1 with 3 V for Dev. #35 have been achieved. These
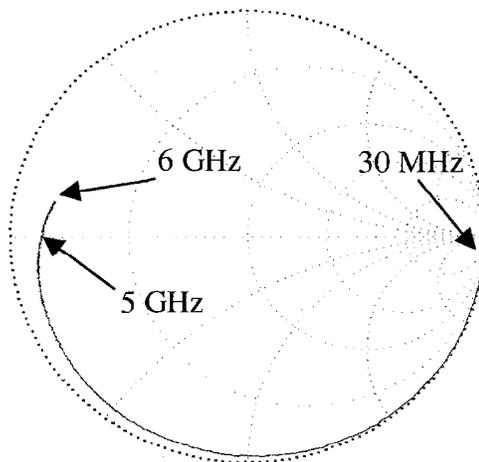
Figure 4. S11 plot in Smith chart shows a minimal parasitic effect, low equivalent series resistance, and a high self-resonance in the MEMS tunable capacitor.
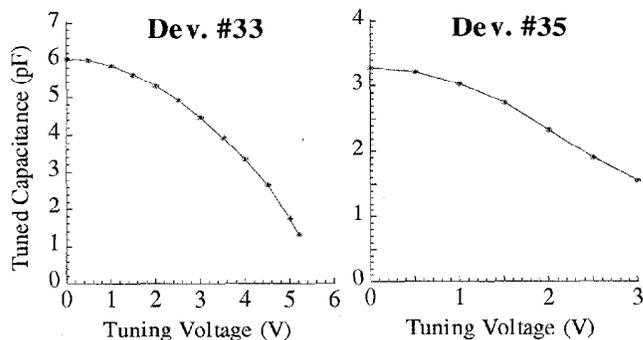
Figure 5. Tuned capacitance is plotted as functions of the applied tuning voltage. Tuning ratios of 4.55 to 1 with a maximum tuning voltage of 5.2 V for Dev. #33 and 2.1 to 1 with 3 V for Dev. #35 have been achieved.

capacitance-tuning results correspond to an approximate deflection of 23 μm and 8 μm, respectively. The approximate power required to actuate such an electrostatic actuator can be expressed as

$$P = \frac{1}{2}CV^2 f$$

where f is the actuation frequency. For an assumed actuation frequency of 1 kHz, the actuation power consumed by Dev. #33 is 61 nW, with an average actuator capacitance of 4.5 pF to achieve a full tuning range swing, which is the worst-case scenario for power consumption at a given actuation frequency.

CONCLUSIONS

A low-power and low-voltage electrostatic actuator was designed, fabricated, and demonstrated as an integrated part of a large tuning-ratio tunable capacitor for RF MEMS applications. The tunable capacitor device was made of single crystal silicon, coated with an aluminum metal thin film, on low loss glass substrate to utilize the superior mechanical properties of single crystal silicon while maintaining excellent electrical characteristics, including low parasitic effects and equivalent series resistance. The integrated electrostatic actuator has been demonstrated to provide a motion of 23 μm for an applied voltage of 5.2 V (61 nW), resulting in the tunable capacitor having a continuous capacitance tuning-ratio in excess of 4.5 to 1. Alternative devices with a required low tuning voltage of 3 V have also been demonstrated to provide a tuning-ratio of 2 to 1. The power consumption of this actuator is linearly proportional to the actuation frequency, and is in the range of 10's of nano-Watts when the device is actuated at a frequency of a few kHz. These low power and low voltage micro-actuators have a wide range of applications, especially in the wireless arena where minimization of power consumption brings significant impacts to the overall system performance.

ACKNOWLEDGEMENTS

REFERENCES

[1] M. Soyuer, K. A. Jenkins, J. N. Burghartz, and M. D. Hulvey, "A 3V 4 GHz nMOS Voltage-Controlled Oscillator with Integrated Resonator," *Technical Digest of IEEE International Solid-State Circuit Conference*, Feb. 1996, pp. 394-395

[2] D-W Kim and Y-S Kwon, "Photoresponse of Area-Variable Varactor Diode," *Electronics Letters*, vol. 32, no. 5, Feb. 29, 1996, pp. 466-467

[3] J. J. Yao, S. Park, and J. DeNatale, "High Tuning-Ratio MEMS-Based Tunable Capacitors for RF Communications Applications," *Proc. of 1998 Solid-State Sensor and Actuator Workshop*, Hilton Head Island, 1998

[4] D. J. Young and B. E. Boser, "A Micromachined Variable Capacitor for Monolithic Low-Noise VCOs," *Technical Digest of Solid State Sensor and Actuator Workshop*, 1996, pp. 86-89

[5] A. Dec and K. Suyama, "Micromachined Varactor with Wide Tuning Range," *Electronics Letters*, vol. 33, no. 11, May 22, 1997, pp. 922-924

[6] E. Hung and S. Senturia, "Tunable Capacitors With Programmable Capacitance-Voltage Characteristic," *Technical Digest of Solid State Sensor and Actuator Workshop*, 1998, pp. 292-295

[7] J-H Park, H-T Kim, Y. Kwon, and Y-K Kim, "A Tunable Millimeter-Wave Filter Using Coplanar Waveguide and Micromachined Variable Capacitors," *Tech Digest of The 10th International Conference on Solid-State Sensors and Actuators*, 1999, pp. 1272-1275

[8] C. L. Goldsmith, A. Malczewski, Z. J. Yao, S. Chen, J. Ehmke, and D. H. Hinzel, "RF MEMS Variable Capacitors for Tunable Filters," *Int. J. of RF & Microwave Computer-Aided Engineering*, vol. 9, no. 4, 1999, pp. 362-374

[9] Z. Feng, W. Zhang, B. Su, K. F. Harsh, K. C. Gupta, V. Bright, and Y. C. Lee, "Design and Modeling of RF MEMS Tunable Capacitors Using Electro-Thermal Actuators," *Technical Digest of 1999 IEEE MTT-S International Microwave Symposium*, 1999, pp. 1507-1510

A UNIVERSAL MEMS FABRICATION PROCESS FOR HIGH-PERFORMANCE ON-CHIP RF PASSIVE COMPONENTS AND CIRCUITS

Hongrui Jiang, Bradley A. Minch, Ye Wang, Jer-Liang A. Yeh, and Norman C. Tien
School of Electrical and Computer Engineering, Cornell University
Ithaca, NY 14850

ABSTRACT

We have developed a fabrication approach that allows us to integrate monolithically on silicon high-performance on-chip radio-frequency (RF) passive components, such as inductors, transformers and fixed and tunable capacitors. We applied two-layer polysilicon surface micromachining to construct the devices, which were suspended over 30-μm-deep cavities formed into the silicon substrate. We also performed electroless copper (Cu) plating to metallize the polysilicon structures for high conductivity. The inner surfaces of the cavities were Cu coated too, providing good RF ground and electromagnetic shielding. The deep cavities dramatically reduce the electric and magnetic couplings and parasitic capacitances between the devices and the substrate. High quality factors over 30 and resonant frequencies over 10 GHz have been achieved for inductors. We have designed and fabricated high-performance varactors, transformers and LC-passive filters as well.

INTRODUCTION

On-chip passive components, such as inductors, capacitors, and transformers are indispensible in radio-frequency (RF) circuits for wireless communication [1-2]. Today's on-chip inductors, however, generally have low quality factors (Q's), lack good RF grounds, have characteristics dependent on the substrate geometry and their ambient due to electromagnetic coupling, and have low self-resonant frequencies. Many techniques have been developed to reduce the substrate loss and/or parasitics between the inductors and the substrate, such as using a high-resistivity silicon substrate, silicon on sapphire, glass or quartz, etching away the substrate under the device [3], and building the inductor on a thick silicon-oxide layer [4]. These methods, however, cannot solve the ground and cross-talk problems. To provide an electromagnetic shield, a patterned metal shield can be inserted beneath the inductor [3], but the benefit is counterbalanced, more or less, by the loss induced within the inserted shield itself. These problems also apply to on-chip transformers. Tunable capacitors have also been reported using two metal layers [5] or two polysilicon layers [6]. Because the metal used is soft, it cannot form large-area suspended plates. Therefore, the capacitances achieved are small and multiple capacitors must be put in parallel to reach values of a few picofarads [5], which are commonly used in RF circuits. With polysilicon plates, on the other hand, only the top polysilicon layer can be metallized [6], which produces a large series resistance for the capacitor and low Q. Another issue stems from the closeness between the capacitor plates and the silicon substrate, which creates a large parasitic capacitance, on the same order of magnitude as the capacitance obtained from the parallel plates. Consequently, the tunable range of such a varactor is very small [5]. This will be a severe problem for a floating capacitor, which is commonly used. Moreover, these disparate approaches cannot simultaneously solve all the problems involved in the optimization of the elements discussed, let alone provide for the integration of various elements on one chip [3-7].

We previously reported a microelectromechanical-system (MEMS) technology to create on-chip high-Q suspended electromagnetically-shielded spiral inductors, using silicon micromachining and electroless copper (Cu) plating [8-9]. We have improved and extended this process to a universal one with which we can integrate most of the high-performance RF passive components just mentioned onto a silicon substrate. Circuits such as low-pass filters are also realizable through this method. The fabrication processes are also CMOS-compatible; therefore, it can potentially be integrated with CMOS technologies for wider applications.
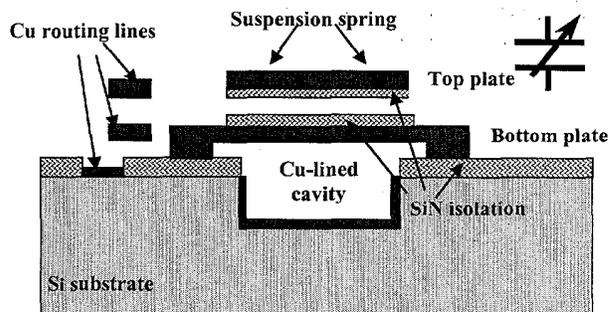
Figure 1. Schematic of the cross section of an electrostatically actuated parallel-plate varactor.

PRINCIPLE AND DESIGN

Our approach features the suspension of all of the devices over cavities formed in the silicon substrate and conformal encapsulation of selectively exposed silicon and polysilicon structures with Cu. The cavity, currently chosen to be 30-μm deep, dramatically reduces the electromagnetic coupling and the parasitic capacitance between the device and the substrate. To save die area, the cavity has vertical sidewalls and an opening slightly larger than the device. Polysilicon is used as the structural material of the suspended devices. The choice of polysilicon as the structural material is based on two factors. First, polysilicon is a stiff material [10] that can well withstand vibrations and shocks from the environment. Second, polysilicon surface micromachining is well developed and has the flexibility to produce complex structures [11]. Cu encapsulation of the polysilicon structures renders low resistance comparable to metal. The cavities beneath the devices are lined with Cu in the same plating procedure to provide both good RF ground and electromagnetic shielding that isolates the devices from their ambient. Hence, high device performances can be achieved in terms of high Q, good isolation, and, in the varactor's case, large tunability. Figure 1 shows schematically the cross section of a parallel-plate varactor fabricated by this method.

FABRICATION PROCESS

The fabrication process was carried out at the Cornell Nanofabrication Facility (CNF). The schematic of the process flow is shown in Figure 2, where the manufacture of a tunable capacitor is used as an example. The process started with the deposition of an 800-nm-thick low-pressure chemical vapor-deposited (LPCVD)
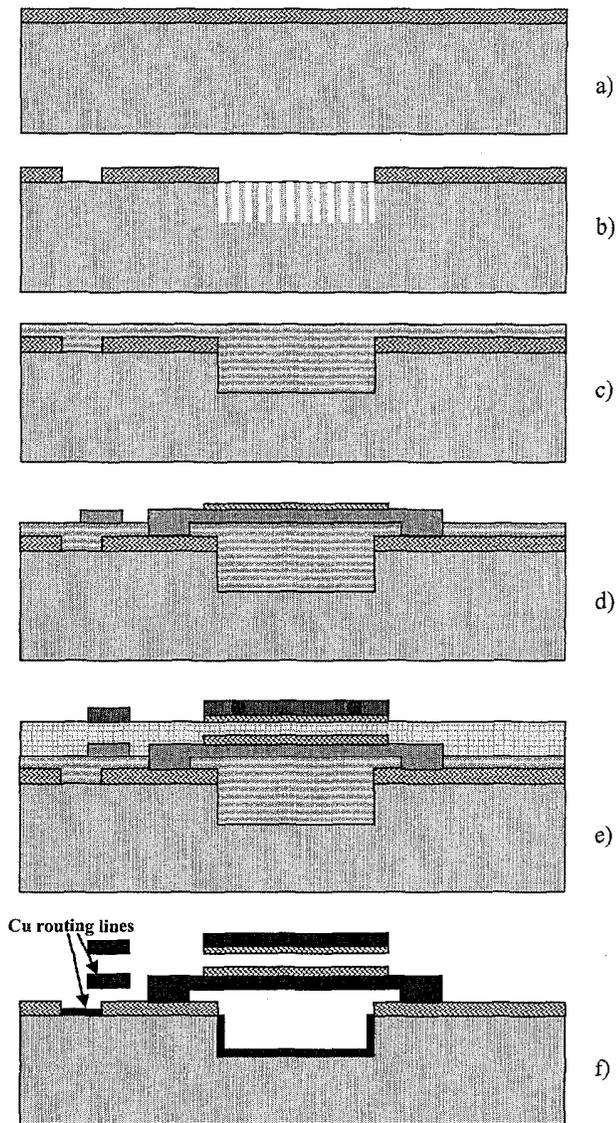
Figure 2. Schematic of the process flow: a) deposition and patterning of isolation silicon nitride; b) etching narrow beam-and-trench structures for the sacrificial silicon-oxide block by DRIE; c) thermal oxidation, silicon-oxide deposition and CMP to form the sacrificial block; d) deposition and patterning of the first polysilicon structural and silicon-nitride isolation layers; e) deposition and patterning of the second sacrificial silicon oxide, the second polysilicon structural and silicon-nitride isolation layers; f) HF release, RTA and electroless Cu plating.

low-stress silicon-nitride film as the isolation layer (Figure 2a). Then, the areas where the sacrificial silicon-oxide blocks for the cavity formation were to be defined were opened by etching away the silicon-nitride layer. The patterns of the first-level Cu routing lines were etched out as well in this step. Next, 30-μm-deep narrow beam-and-trench structureswere created in those opened windows for the sacrificial silicon-oxide blocks by deep reactive ion etching (DRIE) (Figure 2b). Afterwards, the narrow beams were thermally oxidized, followed by the deposition of an LPCVD low-temperature oxide (LTO) to completely seal any openings or gaps left after the thermal oxidation. The SEM image of the cross section of a silicon-oxide block thus formed is shown in Figure 3, where the ripples on the surface due to the sealing of the openings
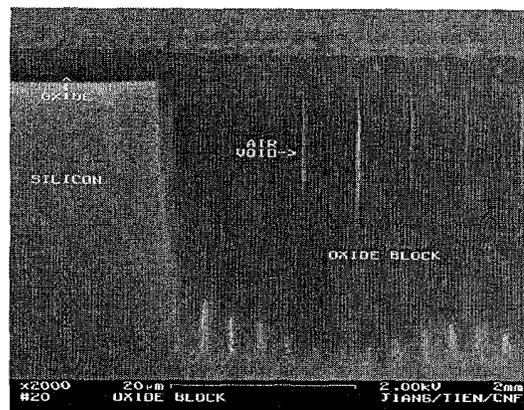
Figure 3. SEM image of the cross section of a 30-μm-deep sacrificial silicon-oxide block.

are clearly shown. Chemical mechanical polishing (CMP) was applied to provide a flat surface for the later steps (Figure 2c).

After the formation of the sacrificial silicon-oxide block, a two-polysilicon-layer surface micromachining process was applied to build the devices. These two n-type polysilicon layers were doped in situ using phosphine (PH_3) as the phosphorus source. The first polysilicon film was used to form the following structures: the inductor spirals, the transformer coils, the bottom plates of the capacitors and the second-level Cu wirings (Figure 2d). The second polysilicon layer was employed to build the following: the overpasses of the inductors, the transformer coils, if necessary, the top plates of the capacitors, the suspension springs of the variable capacitors and the third-level Cu wirings (Figure 2e). The function of the overpasses of the inductors is to connect the input and output ports across the spiral traces and the cavity edges to probing pads outside. A second LPCVD LTO sacrificial layer of a thickness of 3.2 μm was grown and separated the polysilicon layers. This silicon-oxide film was planarized and thinned by a second CMP step. Its thickness was determined by the designed air gap between the two plates of the variable capacitors, which was 1.2 μm in this run. Two thin 250-nm-thick LPCVD low-stress silicon-nitride films were grown and lithographically patterned, one onto the first polysilicon layer and the second onto the second sacrificial silicon-oxide layer, immediately under the second polysilicon layer, wherever overlapping or crossover between the two polysilicon layers occurred. These two silicon-nitride films served as isolation and Cu-plating stoppage layers, because it was found that the electroless Cu deposition favored the areas where the structures were dense, which might cause shorting problems [9]. These two silicon-nitride layers were also used as the dielectric materials between the two plates of fixed capacitors.

The post-processing began with a rapid thermal annealing (RTA) step at the temperature of 1100 °C for 90 seconds to relieve the internal and interfacial stress. The structures were finally released in hydrofluoric (HF) acid and electroless Cu plating was performed (Figure 2f). The process was selective: all the exposed silicon and polysilicon structures were plated with Cu, including the inductors, transformers, capacitors and metal routing lines, while those structures covered with silicon nitride remained as they were, providing good isolation. Figure 4 shows the focus-ion-beam (FIB) image of the cross section of a polysilicon coil fully encapsulated with Cu. As is clearly demonstrated, the plating is conformal. The resistivity of the plated Cu was measured to be 2.1 μΩ-cm [9].
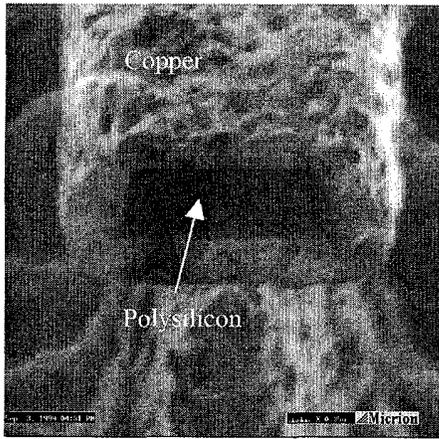
Figure 4. Focus ion beam micrograph of the cross section of a polysilicon coil encapsulated with Cu [9].

a)

b)
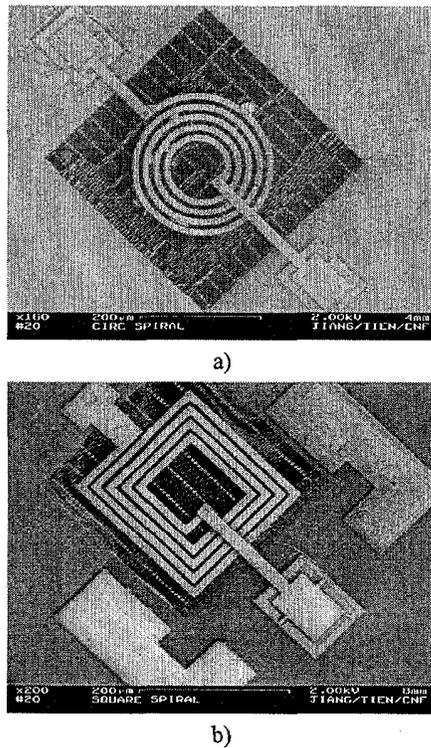
Figure 5. SEM image of spiral suspended inductors: a) circular and b) rectangular.

TEST DEVICES

Figure 5 shows the SEM images of suspended spiral inductors, in both circular (Figure 5a) and rectangular shapes (Figure 5b). The fabricated inductors have inductances ranging from about 1 nH to about 10 nH, and occupy areas, including the cavities beneath, from about 150 × 150 μm² to about 350 × 350 μm². High Q over 30 and self-resonance frequency above 10 GHz have been accomplished [8-9]. Circular spirals suffer less bending and warping after HF release than rectangular ones, probably because the right-angle corners of the rectangular spirals have more stress. Nevertheless, rectangular inductors are studied more comprehensively because they are easier to simulate in a finite-element-method (FEM) simulator such as Microcosm MEMCAD. Simulation predicts that, owing to the Cu-lined cavity, the mutual

inductance between two adjacent inductors drops by a factor around 5 from that without the cavity [9], indicating good shielding effect.
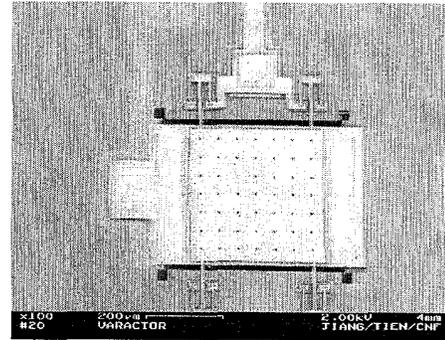
Figure 6. SEM image of an electrostatically actuated parallel-plate tunable capacitor.

Figure 6 shows an electrostatically actuated parallel-plate varactor. The overlapping area of the two plates are 400 × 400 μm² and the air gap between them is 1.2 μm, which gives a nominal capacitance of 1.1 pF if no DC voltage is applied across the two plates. The nominal tunability of this varactor is close to 1.5:1, the maximum predicted by theory [6], owing to the negligible parasitic capacitance between the plates and the cavity inner surfaces. With the designed total spring constant of 4.3 N/m, the maximum capacitance can be achieved under the maximum bias of 3 V.

a)

b)
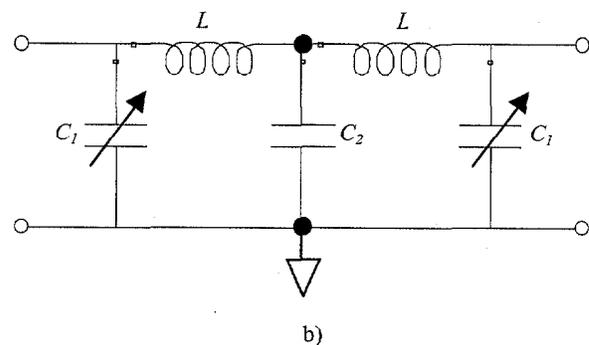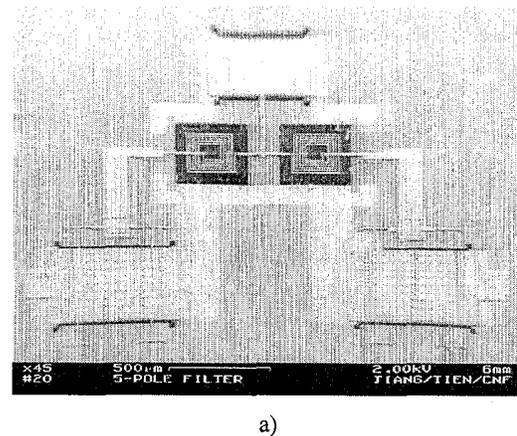
Figure 7. a) SEM image of a five-pole tunable low-pass LC-ladder filter and b) its simplified circuit diagram.

Passive *LC*-filters can be built incorporating the inductors and capacitors described above. A prototype five-pole tunable low-

pass *LC*-ladder filter was fabricated and is demonstrated in Figure 7a and its simplified circuit diagram is given in Figure 7b. The filter consists of two inductors (Figure 7a middle) with nominal inductances of 8 nH, two varactors (Figure 7a bottom) with nominal capacitances varying from 1.1 to 1.65 pF, and one fixed capacitor (Figure 7a top) with nominal capacitance of 10 pF. Simulation predicts a 3-dB frequency ranging from 800 to 900 MHz and a roll-off slope of 30 dB/octave.
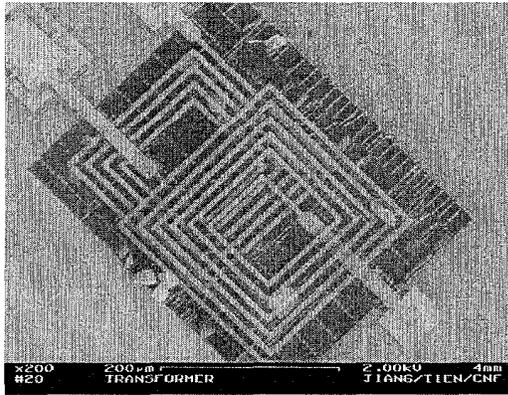
a)
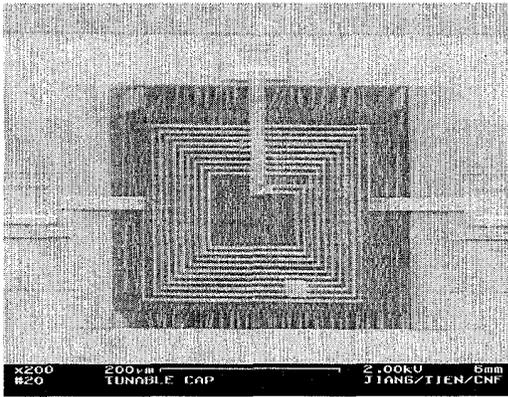
b)

Figure 8. *SEM images of a) a stacked and b) an interleaved three-terminal transformer.*

Figure 8 gives two configurations, stacked (Figure 8a) and interleaved (Figure 8b), of three-terminal transformers. The stacked one is composed of two overlapping spirals built in two polysilicon layers. An overpass and an underpass, respectively, are needed to tap the centers of the coils to outside, as in the case of inductors, and are also made out of the polysilicon. The parameters of these two transformers at 2 GHz, the self-inductances of the primary and secondary coils, L_{pri} and L_{sec}, the mutual inductance, M, and the coupling coefficient, k, defined by $k = M / \sqrt{L_{pri}L_{sec}}$, were simulated by ASITIC [12] and are listed in Table 1. The interleaved configuration offers a large coupling coefficient, k, between the primary and secondary coils, while the stacked structure provides a lower k. This phenomenon is due to the restriction on the overlap area between the two coils to avoid the touching of the overpass/underpass and the coils. Nonetheless, the stacked configuration offers flexibility in k because it allows varying overlap between the primary and secondary coils.

Table 1. *Simulated parameters of transformers at 2 GHz.*

	L_{pri} (nH)	L_{sec} (nH)	M	k
stacked	7.0	6.2	2.1	0.32
interleaved	6.0	5.6	4.6	0.79

CONCLUSION

We have developed and employed a universal MEMS fabrication method to build monolithically on silicon high-performance on-chip passive components, including inductors, transformers and electrostatically actuated parallel-plate tunable capacitors. The devices were constructed using two-layer polysilicon micromachining and were suspended over 30-μm-deep cavities formed in the silicon substrate. Electroless Cu plating was performed to metallize the polysilicon device structures for low series resistance. The same Cu deposition process coated the inner surfaces of the cavities, which formed good RF ground and electric and magnetic shielding. The deep cavities diminish the electric and magnetic couplings and the parasitic capacitances between the devices and the silicon substrate. Therefore, high Q's and small cross-talks for the devices, high self-resonant frequencies for the inductors and large tunability of 1.5:1 for varactors, respectively, can be achieved. An *LC*-ladder filter was designed and fabricated as well. The process can be potentially integrated with conventional CMOS technologies for wider applications.

ACKNOWLEDGEMENTS

REFERENCES

1. P. R. Gray and R. G. Meyer, "Future Directions in Silicon IC's for RF Personal Communication," *Proc. IEEE Custom Integrated Circuts Conference (CICC)*, Santa Clara, CA, May 1995, pp. 83-90.

2. L.E. Larson, "Integrated Circuit Technology Options for RFIC's – Present Status and Future Directions," *IEEE J. Solid-State Circuits*, vol. 33, no. 3, pp. 387-399, Mar. 1998.

3. J. N. Burghartz, "Progress in RF Inductors on Silicon – Understanding Substrate Losses," *Proc. IEEE International Electron Devices Meeting (IEDM)*, San Francisco, CA, Dec. 1998, pp. 523-526.

4. H. B. Erzgräber, T. Grabolla, H. H. Richter, P. Schley, and A. Wolf, "A Novel Buried Oxide Isolation for Monolithic RF Inductors on Silicon," *Proc. IEEE International Electron Devices Meeting (IEDM)*, Sna Francisco, CA, Dec. 1998, pp. 535-539.

5. D. J. Young and B. E. Boser, "A Micromachined Variable Capacitor for Monolithic Low-Noise VCOs," *Technical Digest of the 1994 Solid-State Sensor and Actuator Workshop*, Hilton Head Isl., SC, June 1996, Transducer Research Foundation, Cleveland (1996), pp. 86-89.

6. A. Dec and K. Suyama, "Micromachined Electro-Mechanically Tunable Capacitors and Their Applications to RF IC's," *IEEE*

Trans. Microwave Theory Techn., vol. 46, no. 12, Dec. 1998, pp. 2587-2596.

7. L. Fan, R. T. Chen, A. Nespola, and M. C. Wu, "Universal MEMS Platforms for Passive RF Components: Suspended Inductors and Variable Capacitors," *Proc. IEEE International Conference on Micro Electro Mechanical Systems (MEMS),* Heidelberg, Germany, Jan. 1998, pp. 29-33.

8. H. Jiang, J.-L. A. Yeh, Y. Wang, and N. C. Tien, "Electromagnetically-Shielded High-Q CMOS Compatible Copper Inductors," *Digest of Technical Papers of IEEE International Solid-State Circuits Conference (ISSCC),* San Francisco, CA, Feb. 2000, pp. 330-331.

9. H. Jiang, Y. Wang, J.-L. A. Yeh, and N. C. Tien, "Fabrication of High-Performance On-Chip Suspended Spiral Inductors by Micromachining and Electroless Copper Plating," to appear in *IEEE MTT-S International Microwave Symposium Digest,* Boston, MA, June 2000.

10. K. E. Petersen, "Silicon as a Mechanical Material," *Proc. IEEE,* vol. 70, no. 5, 1982, pp. 420-457.

11. M. Rodgers and J. Sniegowski, "5-Level Polysilicon Surface Micromachine Technology: Application to Complex Mechanical Systems," *Technical Digest of the 1998 Solid-State Sensor and Actuator Workshop,* Hilton Head Isl., SC, June 1998, Transducer Research Foundation, Cleveland (1998), pp. 144-149.

12. A. M. Niknejad and R. G. Meyer, "Analysis, Design, and Optimization of Spiral Inductors and Transformers for Si RF IC's," *IEEE J. Solid-State Circuits,* vol. 33, no. 10, Oct. 1998, pp. 1470-1481.

MEMS-BASED VARIABLE CAPACITOR FOR MILLIMETER-WAVE APPLICATIONS

Zhiping Feng[1], Huantong Zhang[1], Wenge Zhang[2], Bingzhi Su[2],
K. C. Gupta[1], Victor M. Bright[2], and Y. C. Lee[2]
NSF Center for Advanced Manufacturing and Packaging of
Microwave, Optical, and Digital Electronics (CAMPmode)
[1]Department of Electrical and Computer Engineering
[2]Department of Mechanical Engineering
University of Colorado, Boulder, CO 80309-0427
http://mems.colorado.edu

ABSTRACT

We have demonstrated a MEMS-based variable capacitor with outstanding RF performance acceptable for millimeter-wave applications. Actuators were used to move the MEMS capacitor plate vertically to change the air gap between the MEMS plate and the signal line of a coplanar waveguide (CPW). The air gap change resulted in a capacitance change with a ratio of 2.7: 1. The MEMS-based capacitor achieved a very impressive quality factor (Q) that was close to 300 at 0.1 pF and 10 GHz. In addition, the capacitance showed no resonance and the equivalent resistance was less than 0.5 Ω in the frequency range up to 40 GHz. The assembly technology and the RF and actuator designs critical to the capacitor demonstration are discussed.

INTRODUCTION

Microelectromechanical systems (MEMS) allow precise positioning and repositioning of suspended membranes and cantilevers that can be integrated with radio frequency (RF) circuits for tuning and switching. One of the important applications of MEMS in RF circuits and antennas is in the form of tunable capacitors. High Q-factor capacitors are needed in microwave communication systems to replace the semiconductor varactors. In MEMS, mechanical tuning avoids the high losses associated with semiconductors at high frequencies. Polysilicon as the structure material in MEMS with a gold layer and air instead of other dielectric materials make high-Q capacitors possible. Also, it is possible to design MEMS with linear movement (with respect to the control signal), thus allowing linearly tunable capacitance.

Research on the MEMS tunable capacitors is underway [1-4]. Larson et al. designed a MEMS tunable interdigitated capacitor using an electrostatic actuator [1]. The capacitance could be changed from 0.035 pF to 0.1 pF with the bias voltages of 80 V to 200 V. Dec et al. [2,3] have reported a MEMS tunable parallel plane capacitor using, again, an electrostatic actuator. A Q-factor of 9.6 at 1GHz for a 4 pF capacitor has been reported. The capacitance value changed from 4.0 pF to 4.4 pF for bias voltage 0 to 0.8V. Young and Boser [4] have described a MEMS tunable capacitor with electrostatic actuators using aluminum as the structural material and obtained a Q value of 62 at 1GHz for a capacitance of 2.11 pF (4 shunt capacitors in parallel) up to 1.2 GHz. However, all these MEMS devices were developed for microwave applications. It is equally important to develop MEMS variable capacitors for millimeter-wave applications.

In this study, we have demonstrated a MEMS-based variable capacitor with outstanding RF performance acceptable for millimeter-wave applications. The capacitor is shown in Figure 1. More importantly, we have developed the variable capacitor using existing silicon-MEMS foundry MUMPS (Multiple User MEMS Processes). The MEMS structure on the lossy host silicon substrate was transferred to a low-loss ceramics substrate. Such a flip-chip assembly and transfer process was the key technology to assure the excellent millimeter-wave performance of our devices. In addition, the RF and the actuator designs were critical considerations. The assembly technology and the MEMS designs are described in the following sections.
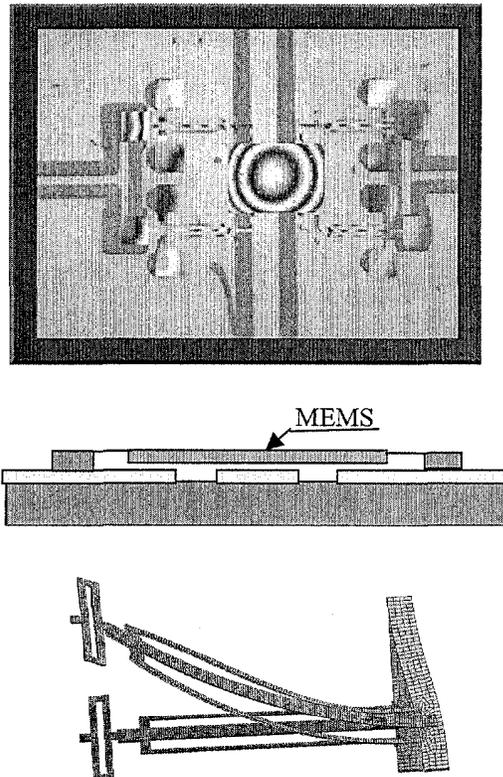
Figure 1: MEMS-based variable capacitor driven by vertical electro-thermal actuators.

ASSEMBLY

Figure 1 shows the capacitor driven by four vertical electro-thermal actuators. The actuators move the MEMS plate vertically to change the air gap between the MEMS and the signal line of a coplanar-waveguide (CPW). For good manufacturability, we decided to use foundry-fabricated MEMS. Unfortunately, a typical MEMS foundry process, e.g. MUMPS for this study, uses low-resistivity silicon substrates that are lossy at high frequencies. As a result, a new assembly technology has been developed to transfer the MEMS from the host silicon substrate to a new ceramics substrate. Referring to Figure 2, after flip-chip bonding, hydrofluoric acid (HF) is used to removed SiO_2 sacrificial layers that are placed between the MEMS and the host silicon substrate. The process was reported in detail in [5,6]. After transferring, the MEMS device made of polysilicon and gold layers can perform well up to 40 GHz if the actuator and the RF designs are appropriately considered.
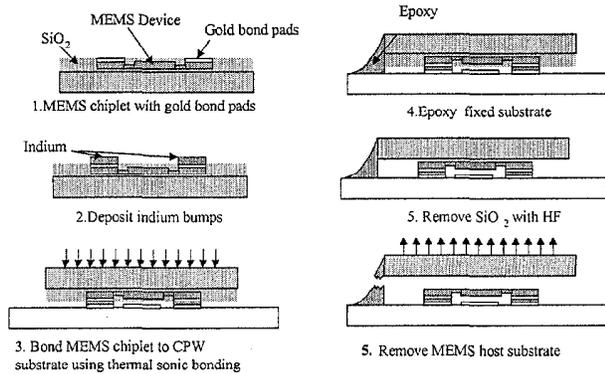
Figure 2: Flip-chip assembly and silicon removal for RF MEMS

ACTUATOR DESIGN

The vertical displacement of the electro-thermal actuator is achieved by the thermal mismatch between the hot and the cold arms. After the flip-chip assembly and transfer, the MEMS vertical displacement should be controllable from 2 µm to 0.5 or 0.2 µm. Such a displacement is achieved by the differential expansions of the arms resulting from different temperature distributions. Figure 3 illustrates the temperature distributions of a design with the cold arm placed in the middle between the two hot arms. The distribution is presented along each arm from the top intersection point to the bottom point anchored to the ceramics substrate. Clearly indicated, the cold arm is really not cold. In fact, its temperature can be higher than the corresponding hot arm when the length is longer than 150 µm. When the temperature differentials are non-linear, repeatable vertical displacements are difficult to achieve. The solution is to place the hot arm in the middle for a consistent temperature differential for repeatable up-and-down plate movements. The improved temperature distributions are shown in Figure 4.
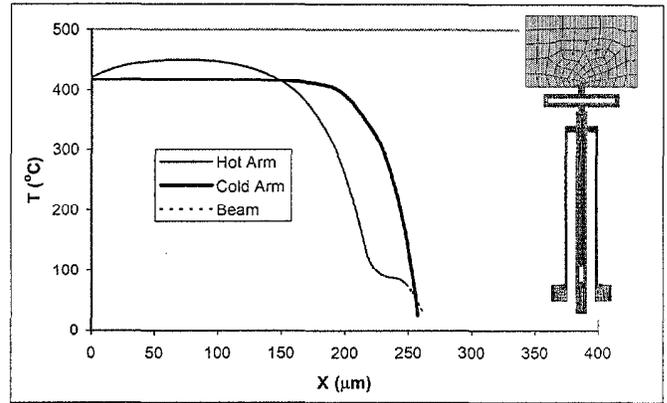
Figure 3: Temperature distributions of an actuator with the cold arm in the middle
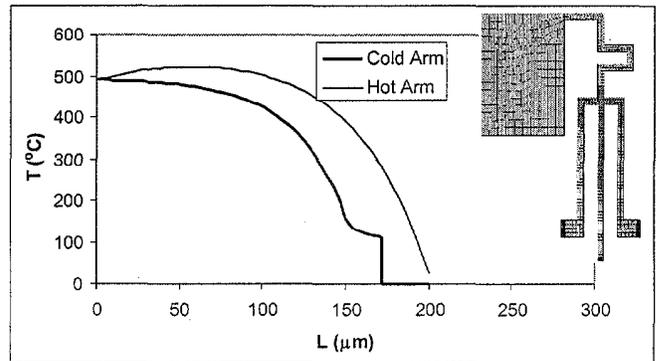
Figure 4: Temperature distributions of an improved actuator with the hot arm in the middle

RF DESIGN AND CHARACTERIZATION

There were two objectives for the RF design: controlling the self-resonant frequency and de-coupling the actuators from the capacitor plate for low RF losses. In addition to capacitance, there is always spurious inductance associated with the plate, the actuator and the connections between the two. With a poor inductance design, the MEMS could reach the undesirable self-resonance around 30 GHz (see Figure 5)[6]. The design was improved to solve this resonance problem. As shown in Figure 6, for low inductance and low loss, a shunt-mounted configuration replaced the series-mounted configuration reported in [6].
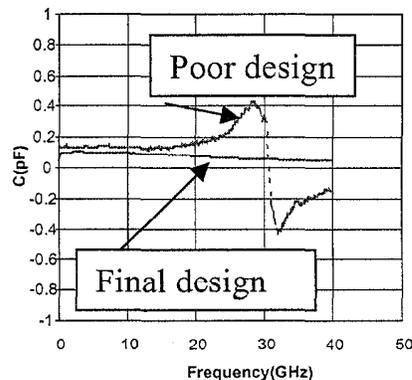
Figure 5: Measured capacitance up to 40 GHz
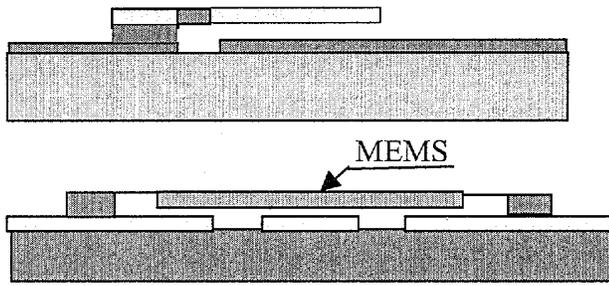
Figure 6: Series-mounted configuration (top) with left-to-right RF transmission and shunt-mounted (bottom) configuration with out-of-plane RF transmission

With the shunt-mounted configuration, the inductance of the actuators was further designed through simulations. Figure 7 illustrates the configuration simulated, and Figure 8 presents the resonant frequencies corresponding to different actuator widths. Wider actuators resulted in low inductance that increased the resonant frequency. In addition to the width, the length and the number of actuators could be changed to control the resonant frequencies.
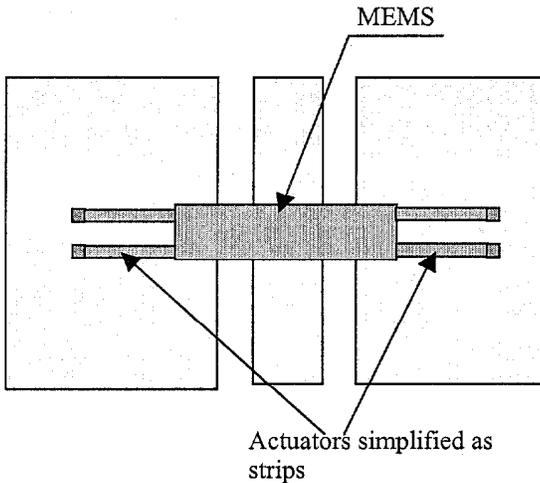
Figure 7: MEMS with actuators simulated
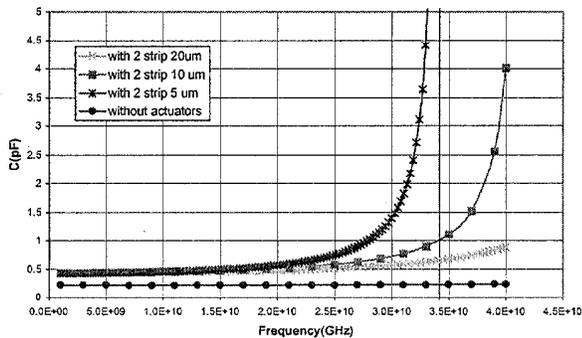
Figure 8: Effects of actuators on resonant frequencies

With the new configuration and enhanced MEMS design, the resonant frequency was increased from 30 GHz to a much higher

value. As shown in Figure 5, there is no resonance up to 40 GHz. The new MEMS variable capacitor could be designed with low inductance for millimeter-wave applications.

However, low inductance was not always preferred. As shown in Figure 9, the actuators should be isolated from the MEMS plate in order to reduce the RF losses. As shown in Figure 10, the de-coupling would become effective when the actuator inductance (La) is high. With high La, the RF current would be limited to the MEMS plate region and could avoid unnecessary losses in actuators. For good design, we want La to be large but L_1 to be small.
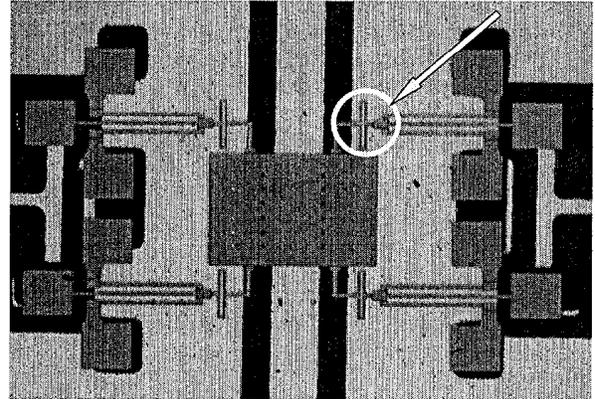
Figure 9: Coupling between the capacitor plate and the actuator
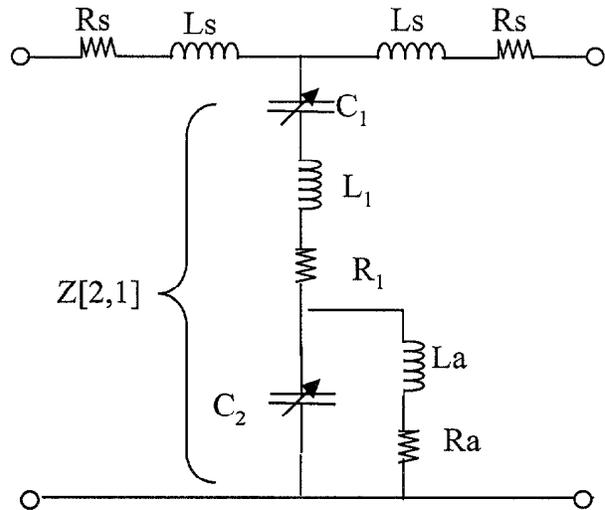
Figure 10: Equivalent circuits for the shunt-mounted MEMS variable capacitor

The S-parameters of the MEMS variable capacitor were measured using HP 8510B Network Analyzer. Figure 11 shows the S-parameters at different capacitance values controlled by the voltages applied to the thermal actuators. The capacitance change achieved a ratio up to 2.7:1 for the case measured. The ratio was directly related to the displacement of the thermal actuator, and could be improved further. Air's dielectric loss is negligible; therefore, the MEMS-based capacitor achieved a very impressive quality factor (Q) that was about 300 at 0.1 pF and 10 GHz. Figure 12 presents the S-parameters measured. These parameters were used to derive the Z-parameters, which were used to establish the equivalent lump elements in the T-circuit as shown in Figure 10.

The equivalent resistance was the Z[2,1] + Rs/2. From the resistance, the quality factor (Q) could be estimated.

Figure 13 shows the equivalent resistances in the frequency range up to 40 GHz. Three interesting observations are listed below.

1. The resistance dropped substantially when the frequency was above 10 GHz. Above that frequency, the actuators were decoupled from the MEMS plate and the RF losses were reduced.
2. The equivalent resistance's error bars were close to the mean values. Thus, the resistances were so low that they were in the "noise" region. It is very difficult to measure these resistance values precisely. But, clearly indicated, the resistances, i.e. RF losses, were very very low even at 40 GHz. It should be noted that the gold thickness of the MEMS plate was only 0.5 µm that is the standard thickness offered by the MUMPS. With a thicker gold layer, the RF losses could be reduced further.
3. The quality factor could be estimated using $Q=1/(\omega RC)$. With $R = 0.5$ Ω, the Q was about 300 at 0.1 pF and 10 GHz.

The achieved quality factor of 300 at 0.1 pF and 10 GHz is very impressive. The MEMS variable capacitor demonstrated has the outstanding RF performance even in the millimeter-wave range.

(a)

(b)
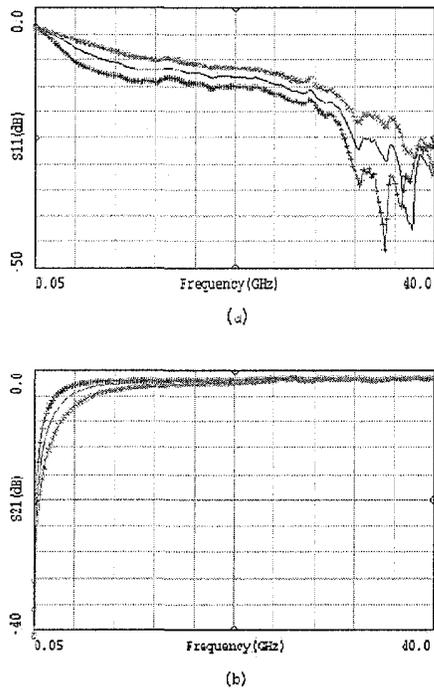
Figure 11: S-parameters at different capacitance values controlled by different biased voltages

SUMMARY

We have demonstrated a MEMS variable capacitor with outstanding RF performance up to millimeter-wave range. More importantly, these MEMS were fabricated using standard silicon-MEMS foundry service, MUMPS. The excellent performance of

our devices was achieved by using a novel flip-chip assembly with silicon removal technology and by appropriate actuator and RF designs.
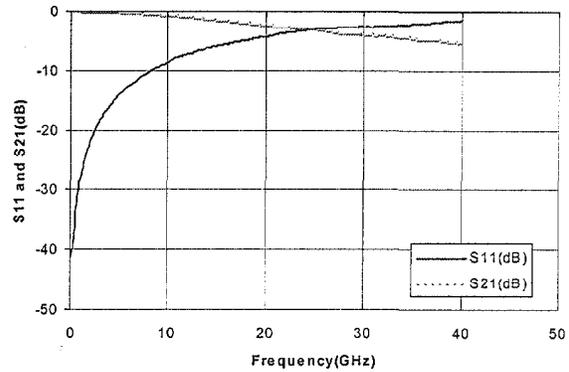
Figure 12: S-parameters for the MEMS variable capacitor
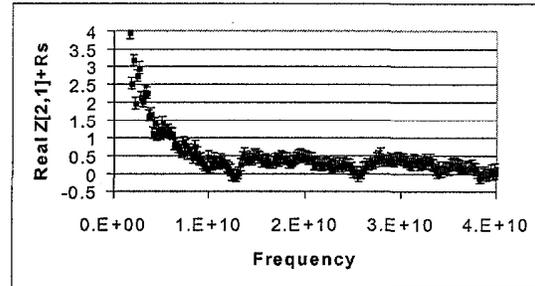
Figure 13: Measured resistance values up to 40 GHz

ACKNOWLEDGMENTS

REFERENCES

1. L. Larson, et al., "Micromachined microwave actuator (MIMAC) technology - a new tuning approach for microwave integrated circuits," IEEE Microwave and Millimeter-wave Monolithic Circuits Symposium, 1991, pp. 27-30.
2. A. Dec and K. Suynama, "Micromachined varactor with wide tuning range," Electronics Letters, Vol. 33, No. 11, May 1997, pp. 922-924.
3. A. Dec and K. Suynama, "A 2.4 GHz CMOS LC VCO using micromachined variable capacitors for frequency tuning," 1999 IEEE MTT-S International Microwave Symposium Digest, pp. 79-82,
4. D. Young and B. Boser, "A micromachined variable capacitor for monolithic low-noise VCOs," Solid-state Sensor and Actuator Workshop, June 2-6, 1996. pp. 86-89.
5. K.F. Harsh, et al., "Flip-chip assembly for Si-based RF MEMS," Technical Digest of *12th IEEE International Conference on MicroElectroMechanical Systems – MEMS '99*, pp. 273-278, 17-21 January 1999, Orlando, FL.
6. Z. Feng, et al., "Design and modeling of RF MEMS tunable capacitors using electro-thermal actuators," 1999 IEEE MTT-S International Microwave Symposium Digest, pp. 1507-1510.

SCANNING MICROLENS SHUTTLES FABRICATED USING SILICON-ON-INSULATOR WAFERS

Decai Sun, Chingwen Yeh and Michel Rosa
Xerox Palo Alto Research Center, Palo Alto, CA 94304.
Email: mrosa@parc.xerox.com ; Fax: 650-812-4105

Abstract

Scanning microlens shuttles fabricated using silicon-on-insulator (SOI) wafer material are presented in this paper. Refractive polymer microlens arrays are integrated on top of a single crystal silicon shuttle, which is actuated by a set of comb drives fabricated in the wafers' single-crystal-silicon (SCS) device layer. The lens shuttles have been dynamically operated at resonant frequencies ranging from tens of hertz up to ~10kHz. With a maximum displacement of ~40μm demonstrated at a fundamental resonant frequency of ~4.7kHz, with a quality factor (Q) of ~149. Upon the SCS shuttle(s), a microlens was fabricated using a photoresist reflow technique. The lenses fabricated using this technique have a calculated focal length of ~808 μm.

Introduction

Technological development and progress in optical MEMS is driving the desire to make low cost miniaturized optical scanners for applications in the areas of display, printing and more recently optical telecommunications. Such devices can be easily integrated with light sources such as edge emitting lasers and vertical cavity surface emitting lasers to form compact laser scanners. Previously a scanning diffractive Fresnel microlens has been demonstrated using a poly-Si surface MEMS fabrication technique [1]. In this paper we present a refractive polymer microlens array integrated on top of a comb-drive actuated SCS shuttle fabricated using SOI. We believe this to be the first demonstration of such a device ever reported.
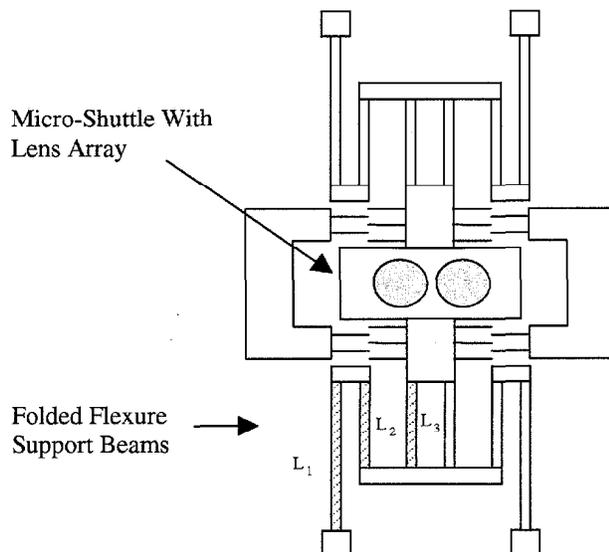
Figure. 1. A comb drive actuated shuttle with a polymer microlens array.

Device Design

The device shown in Figure 1 has been prototyped for use in laser scanning applications. A SCS stage of ~545μm x ~300μm suspended by folded springs on either side and actuated by a set of interdigitated comb fingers, is used as the platform for the microlens array. Device actuation is achieved electrostatically by the 100 comb fingers on either side of the suspended stage. The large number of comb fingers is used to generate a large electrostatic force for a relatively low applied voltage, which in combination with a reduced spring constant can more readily facilitate the large displacements required for beam scanning applications [2]. Each of the interdigitated comb fingers is ~55μm long and ~2μm wide, with an interdigitated comb finger gap of 2 μm. The schematic of Figure 1, depicts the entire device with L_1, L_2 and L_3 representing the lengths of each of the beams in the folded flexure design used to support the lens shuttle. The spring constant calculation used for this support beam configuration is :

$$k_s = \left[\frac{\dfrac{k_1 * k_2}{k_1 + k_2} * k_3}{\dfrac{k_1 * k_2}{k_1 + k_2} + k_3} \right] * 4 \; ; \; \text{where } k_i = \frac{12EI}{L^3_i} \quad (1)$$

The young's modulus used for the calculations was 150Gpa and for devices such as that of Figure 1, the dimensions of beams L_1, L_2 and L_3 were ~350μm, ~300μm and ~300μm respectively for a support beam width of ~5μm. The resonant frequency of the shuttle is :

$$f = \tfrac{1}{2\pi}\sqrt{\frac{k_s}{M}} , \quad (2)$$

where M is the effective mass of the lens shuttle.

The focal length of the refractive polymer lens is,

$$f = \frac{R}{n-1} \quad (3)$$

where n is the refractive index of the polymer, and R is the radius of curvature of the spherical photoresist lens. The thickness and the diameter of the photoresist lens before reflow determine its radius of curvature and therefore the final focal length.

Fabrication

Typically, comb-drive based devices are fabricated using surface micro-machining techniques on deposited polysilicon layers [3], [4]. Previously, it has been found that the use of SOI allows simple device fabrication requiring only one masking step [5]. Such techniques were applied here to reduce the fabrication complexity of the microlens devices. The wafer material used for the fabrication process consisted of 4-inch (100) BESOI (Bond and Etch Back Silicon-On-Insulator) wafers with a device layer thickness of 10µm, a buried oxide thickness of 0.5µm and a handle wafer thickness of ~420µm. The main processing steps toward the fabrication of the microlens devices tested are illustrated in the process flow of Figure 2.
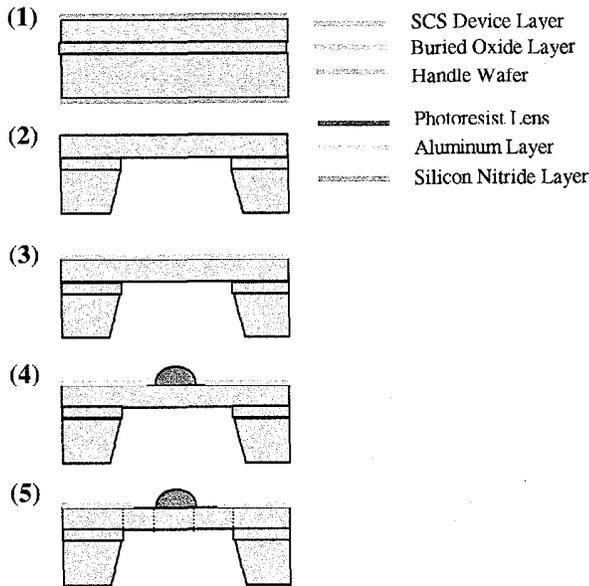
Figure 2. The fabrication of the microlens array on the shuttle of a comb-drive device involved 6 main processing steps.

Step (1) of the process proceeds a standard wafer cleaning procedure and involves the deposition of a LPCVD SiN$_x$ masking layer on the front and backsides of the wafer in preparation for KOH etching. Having lithographically defined open areas in the SiN$_x$ masking layer on the back side of the substrate, the wafer is immersed in an anisotropic etch bath consisting of KOH 30% for step (2) of the process. The SiO$_2$ etch stop layer is removed in buffered HF following the KOH etch. Once this is complete, the deposition of an Al layer on top of the wafer is performed in

step (3) after the SiN$_x$ is removed. This layer serves to pattern the electrical contacts for the device. In step (4), photolithography is performed toward the formation of the microlens array. Each lens is measured to be around 5µm in height and ~130µm in diameter after reflow. Reflow of the microlens array after photoresist patterning, takes place at a temperature of ~200°C [6]. Each lens in the array is spaced at 150µm. In a final processing step, step (5) involves the Deep Reactive Ion Etch (DRIE) of the 10µm thick SCS device layer. For this step, photoresist is used as a masking layer to protect both the lens array and the Al regions deposited earlier in the process, while patterning the comb drive device. Using a STS (Surface Technology Systems) Multiplex ICP (Inductively Coupled Plasma), the DRIE process was performed to form and free the shuttle array. Finally the photoresist etch mask is removed in an O$_2$ plasma etcher. The etch rate during this process was ~1.3µm/min.
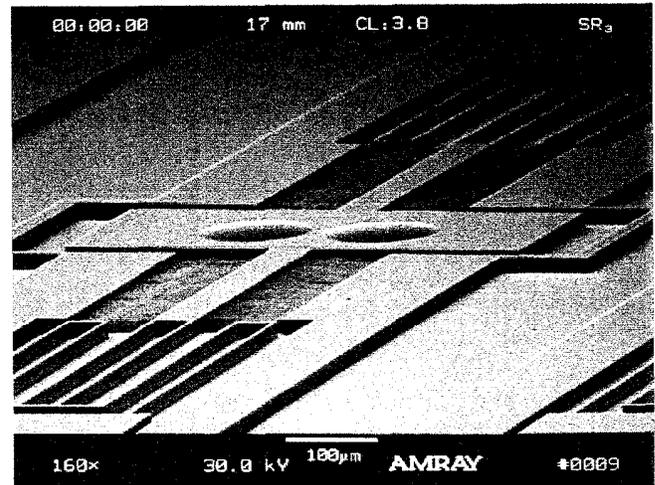
Figure 3. A SCS comb drive-actuated shuttle with a polymer microlens array fabricated on top prior to final release of the shuttle.

Figure 3 shows an SEM photo of a completed device with a set of two microlenses formed on the shuttle mass. Using similar fabrication steps as those previously outlined, the basic device design shown in Figure 3 can be expanded upon to create a super lens shuttle which can be used to provide longer and more seemless scan lines by incorporating a larger number of microlenses.

DEVICE	L₁, L₂, L₃ (μm)	Beam Width (μm)	Ks (Spring Const)	Voltage (AC)	f_r (Analytical)	f_r (Experimental)	Displacement (a.u.)
2-Lens Array	540,500,500	4.2	1.09096	25	2268 Hz	~ 4550 Hz	~ 30
2-Lens Array	350,300,300	4.2	4.58548	25	4651 Hz	~ 4700 Hz	~ 30

Table 1. For the micro-shuttle devices on which an array of two lenses were fabricated, the analytical and experimental characteristics are listed here for each design where the lengths of the support beams L₁, L₂ and L₃ are different.
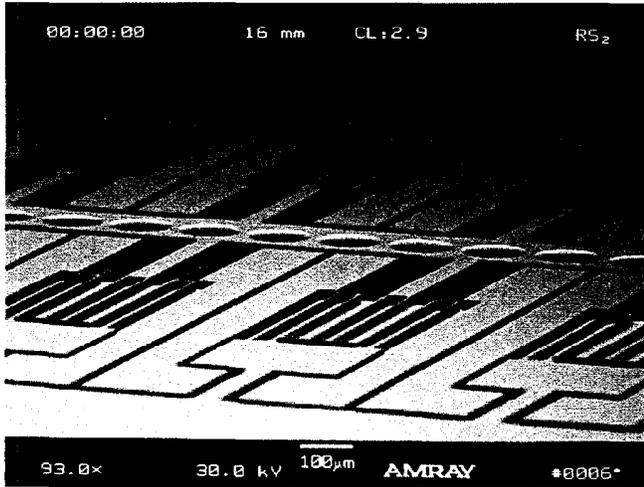
Figure 4. An SEM photograph showing part of a fabricated 13 lensed device.

Figure 4 shows an SEM photo of one such device – a super lens shuttle actuated by multiple parallel comb drive actuators. With a total of 13 microlenses, this is the largest actuated array fabricated in this series of experiments. Other devices which contain an array of 6 microlenses have also been successfully fabricated and tested.

Device Characterization

After microlens reflow, the thickness of the microlens is measured used a profiler. As shown in Figure 5, the thickness of the spherical lens is 5 μm. The radius of the lens is about 130 μm. It is estimated that the focal length of the refractive length is ~808 μm. As with most comb drive based actuators, testing involved both static and dynamic actuation measurements. A typical device as shown in Figure 3, where L1, L2 and L3 are 350μm, 300μm and 300μm respectively, demonstrated a displacement close to ±40 μm in resonance with a peak voltage of 25 V. The displacement of the lens shuttle was observed in slow motion with a strobe illumination system. Dynamic testing of this 2-lens device displayed a fundamental resonant frequency of ~4700Hz. Measurements carried out at and around the fundamental frequency of this device resulted in a Q of ~ 145. As such these devices display quality factors well in excess of those readily attainable by typical surface micromachined resonator devices fabricated using polysilicon or other
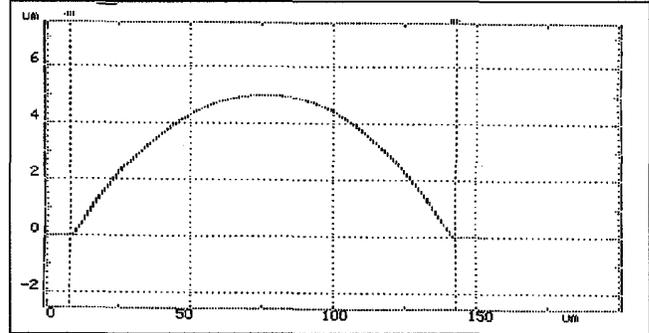
Figure 5. Surface profilometry measurement of a micro-lens after thermal reflow.

such deposited materials. It is believed that this feature is due in principle to the low air damping and high mechanical quality of the SCS layer in which the comb drive actuator is fabricated. Damping is greatly reduced due to the cavity over which the device is suspended. Formed using the backside KOH etch described earlier, this cavity enables the device to reach very high Q levels, an artifact that has made the experimental determination of device displacement at and around resonant peaks more complex.

Frequency Response Of Micro-Lens Shuttle
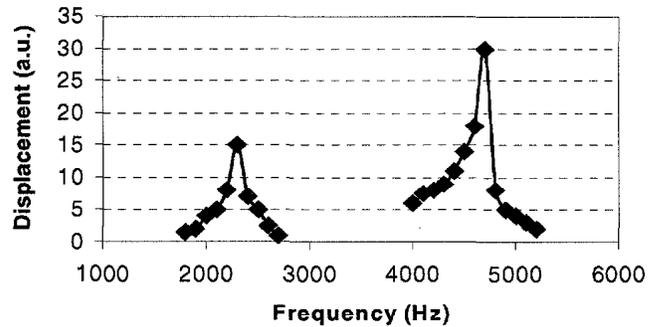
Figure 6. This graph depicts the experimentally determined fundamental frequency peak and its nearest sub-harmonic at ~4.7 kHz and ~1.15kHz, respectively.

Table 1 compares the analytical and experimental fundamental frequencies of both device designs in which for the 2-lens array design, only the support beam length varies. As the data in Table 1 indicates, an increased support beam length and reduced width (due to slight

DRIE over etch), combine to decrease the spring constant for the device(s), conversely increasing the frequency for each at which the fundamental resonant peak resides. The first of the two devices listed in Table 1 shows a large, almost 2-fold discrepancy between the analytical and experimental fundamental frequencies. Repeated measurement of multiple devices across the wafer resulted in a similar result, leaving the authors unclear as to why this phenomenon has occurred. Figure 6 displays two resonant peaks experimentally derived from a 2-lens array device where L_1, L_2 and L_3 are 350µm, 300µm and 300µm respectively. Both displacement curves are measured in absolute units with the devices fundamental frequency having the larger of the two peaks at ~30 (a.u.) and its sub-harmonic at ~1.150kHz having a peak of ~15 (a.u).
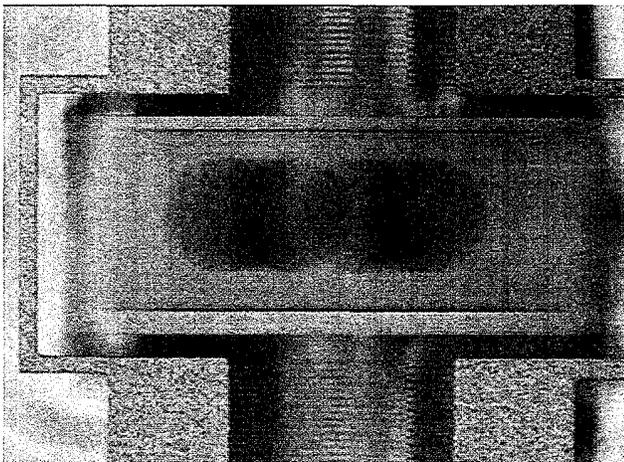
Figure 7. A micro-shuttle with a 2-lens polymer array, at resonance on the probe station.

A typical device with two polymer microlenses is shown at resonance in Figure 7. From the photo in Figure 7, the large displacement achieved while in resonance can easily be seen. This particular shot is taken while at a frequency of ~4700Hz and with an applied voltage of around 25volts (peak). All electrical device testing for these devices was performed using a four-point probe station and both analogue and digital function generators.

Conclusion
This paper has presented a device concept which would allow the seamless scanning of a laser beam across a given area. Such devices have been fabricated and shown experimentally to provide large displacements for relatively low actuation voltages at high resonant frequencies and Q values. Subsequent to this, the design and fabrication methodologies employed to realize these devices can be expanded upon to allow further device integration which would see these devices perform beam scanning in applications ranging from printing to image scanning.

References

[1]. L. Fan, M. C. Wu, K.D. Choquette, CME2, CLEO'97.
[2]. R. Legtenberg, A.W. Groeneveld and M. Elwenspoek, "Comb-drive actuators for large displacements", J. Micromech and Microeng, no. 6, pp 320-329, 1996.
[3]. W.C.Tang, T.H.Nguyen, M.W. Judy and R.T. Howe, "Electrostatic-Comb Drive of Lateral Polysilicon Resonators", Sensors And Actuators, A21-A23, pp 328-331, 1990.
[4]. J.J. Sniegowski and E.J. Garcia, "Surface-Micromachined Gear Trains Driven by an On-Chip Electrostatic Microengine", IEEE Electron Device Letters, pp 366-368. 1996.
[5]. M.A. Rosa, S. Dimitrijev and H.B. Harrison, "Improved Operation of Comb-Drive Actuators through the use of a New Comb Finger Design", J. Inl't Material Systems and Structures, vol. 9, no. 4, pp 283-290, 1998.
[6]. Z.D. Popovic, R.A. Sprague and G.A. Neville Connell, "Technique for monolithic fabrication of microlens arrays", Applied Optics, vol. 27, no. 7, pp 1281-1284, 1988.

MICROMACHINED SCANNING REFRACTIVE LENS

Daniel A. Fletcher, Kenneth B. Crozier, Gordon S. Kino, Calvin F. Quate, and Kenneth E. Goodson
E. L. Ginzton Laboratory, Stanford University
Stanford, CA 94305

ABSTRACT

A refractive lens placed close to a surface can be used to obtain resolution below the optical diffraction limit in air. The improvement in resolution is proportional to the index of refraction of the lens, called a Solid Immersion Lens (SIL). Small lenses (diameter ~ 10 µm) made with micromachining techniques offer several advantages over large SILs (diameter ~ 1 mm) including reduced aberrations and, when integrated onto cantilevers, control of lens-sample separation. We report here the fabrication and demonstration of the first micromachined scanning refractive lens. The integrated lens and cantilever is made from single-crystal Si and operated in transmission mode.

INTRODUCTION

Scanning probes such as the STM and AFM characterize surfaces through local electric, magnetic, and other interactions. Optical imaging below the diffraction limit can be obtained by scanning a sub-wavelength source or collector such as a tapered optical fiber above the surface, a technique known as scanning near-field optical microscopy (SNOM) [1]. Sub-wavelength sources can be created by forcing light through small apertures like those on tapered fibers. This approach suffers from low optical efficiency and difficulty manufacturing reliable apertures. An alternate and more efficient technique for high-resolution imaging is solid immersion microscopy, which uses the index of refraction of a lens held against the surface to reduce spot size [2].
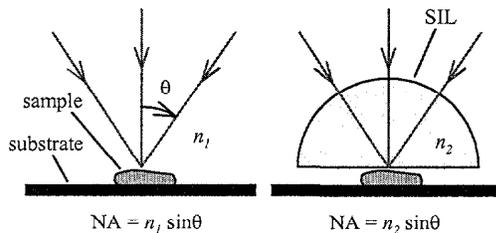
Figure 1. Focussing through air and through a hemispherical SIL. The minimum resolvable feature size is improved by n_2 / n_1 with use of the SIL.

The spatial resolution of a lens without aberration is limited by diffraction of light from the lens aperture. The minimum full-width half-maximum (FWHM) resolution, w, in the scalar approximation is $w = \lambda / 2NA$, where λ is the free space wavelength, NA $= n \sin \theta$ is the numerical aperture, θ is the maximum angle of incidence, and n is the index of refraction at the focus. Figure 1 shows light focused on a sample through a medium of index n_1 and through a hemispherical SIL of index n_2 and radius a onto the same sample. In the case of imaging in air, $n_1 = 1$, the improvement in resolution through the hemispherical SIL is $n_2 / n_1 = n_2 = n$. If the lens is a supersphere, where the light

is focussed to a point a / n from the center of the spherical surface, refraction at the lens further increases resolution for angles in air below the maximum θ_m given by $\sin \theta_m = 1 / n$. The highest resolution for both hemisphere and supershpere SILs is the same and occurs when $\sin \theta$ in the lens is 1.0.

Terris et al. [3] were able to write 317-nm bits in a magneto-optic material with a $n = 1.83$ SIL at $\lambda = 780$ nm. Ghislain et al. [4] exposed 190-nm lines in photoresist using a $n = 2.2$ SIL at $\lambda = 442$ nm. Other work has used SILs for high-resolution microscopy and phase-change data storage [5-6]. Further improvements in resolution may be possible with short wavelengths and high-index materials, such as Si_3N_4 ($n = 2.0$), $SrTiO_3$ ($n = 2.4$), or Si ($n = 3.4$).

The sample surface must be positioned within the near field of the SIL for optimum resolution, approximately 100 nm for visible light. If topographical variations in the sample cause a greater separation between lens and sample, resolution is reduced because of the exponential fall-off of the focussed fields in air [7]. Lenses can be made with tips on the order of 1 µm in diameter to localize contact between the SIL and sample, and images can be constructed by scanning the lens across the sample as with the AFM. Integrating a small lens onto an AFM cantilever allows scanning across high spatial-frequency topography with force feedback control. Traditional methods for manufacturing SILs cannot be used to make lenses smaller than approximately 1 mm in diameter. Micromachining offers a relatively inexpensive and repeatable method for batch fabrication of integrated lenses and cantilevers.
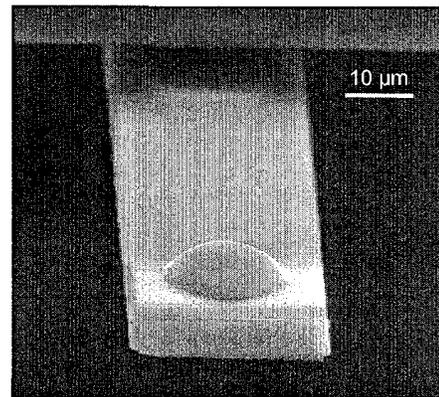
Figure 2. Integrated refractive microlens and cantilever fabricated from single-crystal Si.

Micromachined lenses also benefit from reduced aberrations and sensitivity to focal position in the propogation direction. As lens diameter is reduced, spherical aberration is reduced, and lenses smaller than 10 µm have been shown through an extension of Mie theory [8] to focus both plane waves and converging waves of $\lambda = 0.633$ µm to nearly the same spot size [9]. Lenses with diameters as small as one wavelength have also been shown to
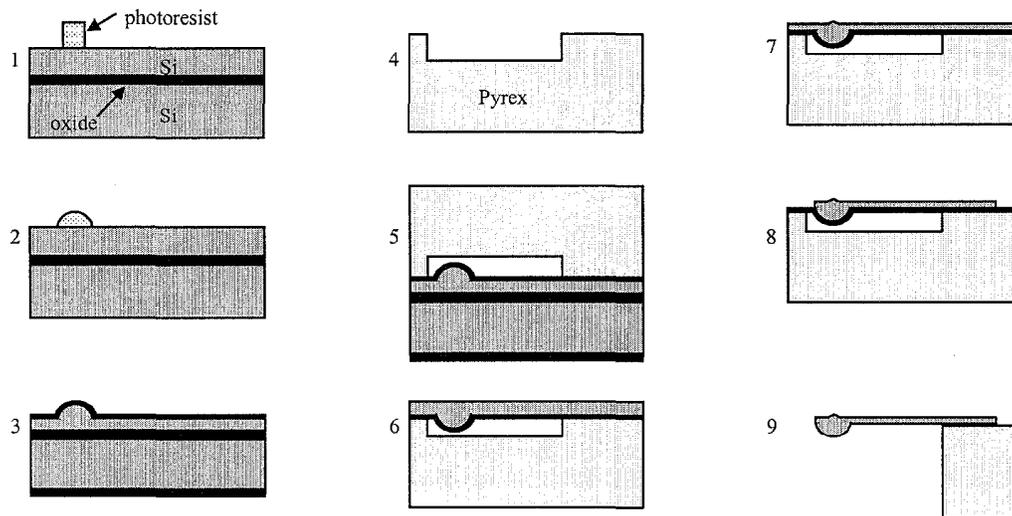
Figure 3. *Process flow for fabrication of the integrated cantilever and lens. (1) Pattern photoresist pillars. (2) Reflow pillars in acetone vapor. (3). Transfer shape into upper Si and grow thermal oxide. (4) Etch pits in Pyrex. (5) Anodically bond Pyrex and upper Si. (6) Etch substrate Si in TMAH and clear middle oxide in BOE. (7) Pattern tip with photoresist and isotropically dry etch. (8) Pattern cantilever with photoresist and anisotropically dry etch. (9) Release with wafer saw.*

focus light, though with a field distribution that differs from vector diffraction theory.

MICROLENS FABRICATION

Cantilevers with lenses on the order of 10 μm in diameter are fabricated from the upper single-crystal Si of a silicon-on-insulator (SOI) wafer. An integrated microlens and cantilever is shown in Fig. 2. The fabrication process begins by defining a pillar in photoresist, as shown in Fig. 3. The pillar is reflowed in acetone vapor to achieve the desired lens shape. The shape is a function of reflow time, acetone vapor concentration in the sealed container, and aspect ratio of the original photoresist pillar. After baking at 110 °C for 20 minutes to drive off solvents, the shape is transferred into the upper Si with a non-selective CF_4 / O_2 reactive ion etch [10]. The shape of the lens can be further adjusted during the transfer etch by varying the relative Si:photoresist etch rate. For example, a hemispherical shape transferred with an etch preferential to photoresist will result in a flattened Si lens shape. A 0.3-μm thick thermal oxide layer is grown on the lens to reduce roughness and serve as an etch stop for the cantilever etch.

A tip is needed opposite the lens at the focal point to localize the contact between lens and sample for scanning. To access the lower face of the upper Si layer, a Pyrex handle wafer is anodically bonded to the top Si at 350 °C. The Pyrex wafer is patterned with pits and etched in a HF / HNO_3 / H_2O mixture. The pits protect the lenses and define cantilever length. The bulk substrate Si is removed with a 25% TMAH etch, and the middle oxide is cleared with 6:1 BOE. Tips are then patterned on the suspended Si membrane and etched in an isotropic reactive ion etch. The cantilevers are patterned and isotropically etched, stopping on the thermal oxide layer grown earlier. The oxide is removed in 6:1 BOE to release the cantilevers. Finally, chips with integrated lenses and cantilevers are cut by wafer saw. A one-dimensional array of integrated lenses and cantilevers is shown in Fig. 4.
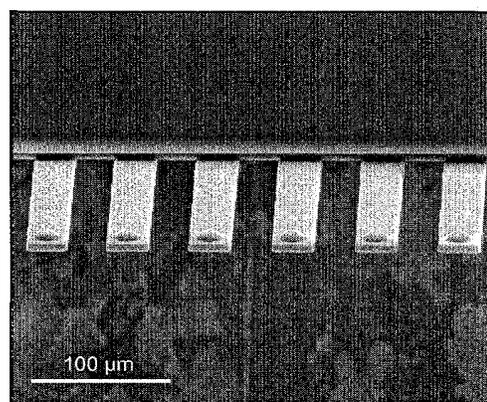
Figure 4. *One-dimensional array of integrated refractive microlenses and cantilevers.*

EXPERIMENTS

We have used the micromachined scanning lens to focus 0.633-μm light and to collect a two-dimensional image of a 0.2 μm line-width grating in the transmission mode. Focussing was demonstrated by imaging a laser spot with and without the micromachined lens. Figure 5A shows a HeNe laser spot (λ = 0.633 μm) created by a low-angle objective (sin θ = 0.04) and imaged with a CCD camera. When the micromachined lens is inserted near the focal point, as shown in Fig. 5B, the spot size is reduced by a factor of 15.7.

The improvement in spot size is due to the index of refraction of Si (n = 3.8 at λ = 0.633 μm) and refraction at the lens surface. Refraction occurs because the SIL is operating like a supersphere, in which the tip is below the center of the sphere and light focussed on the tip enters at an angle to the radius of the sphere. The maximum angle of incidence increases such that sin θ is a
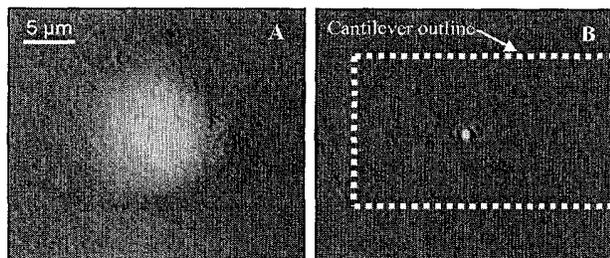
Figure 5. *Focussing by the microlens. (A) HeNe laser spot ($\lambda =$ 0.633 µm) created by a low-NA objective (NA = 0.04). (B) HeNe laser spot focussed through the micromachined Si lens.*

factor of 4.0 greater with the lens than without. A comparison of the spots in Fig. 5 fit to a Guassisan profile is presented in Fig. 6.

The RMS surface roughness of the lens is estimated from AFM scans to be less than 2 nm. Total integrated scattering due to the roughness is less than 0.5%. Note that bulk Si absorbs at $\lambda =$ 0.633 µm, but the lens is partially transparent. If micromachined lenses are made thin enough, it is possible to use absorbing, high-index materials such as Si and achieve good transmission efficiency.
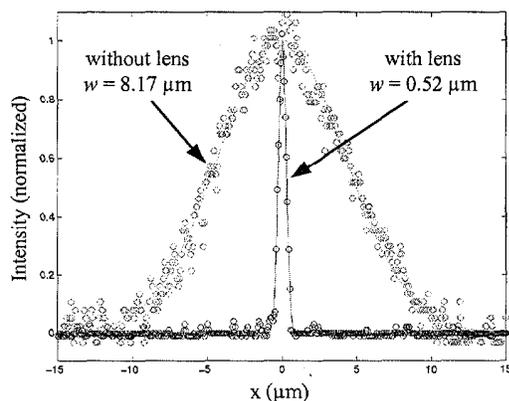
Figure 6. *Comparison of the FWHM size, w, of the focussed spots shown in Fig. 5 fit to a Guassian profile.*

Conventional lenses form images of adjacent points within their field of view. In contrast, a micromachined lens integrated onto a scanning probe with a tip to localize contact forms an optical image serially in a manner similar to an AFM. We have demonstrated scanned image formation by moving an amplitude grating beneath a single lens in contact mode and collecting the transmitted light. The grating is 0.1-µm thick Ti on quartz with a line-width of 0.2 µm, and the cantilever was operated without force feedback. A HeNe laser (λ = 0.633 µm) was chopped and focused onto the micromachined lens with a high-angle objective (sin θ = 0.8). Transmitted light was collected with a second high-angle objective (sin θ = 0.8) and measured with a photodetector and lock-in amplifier. A two-dimensional scan of the 0.2-µm line-width grating is shown in Fig. 7, with the normalized transmitted intensity ranging from 0.35 over a line to 0.65 over a space.
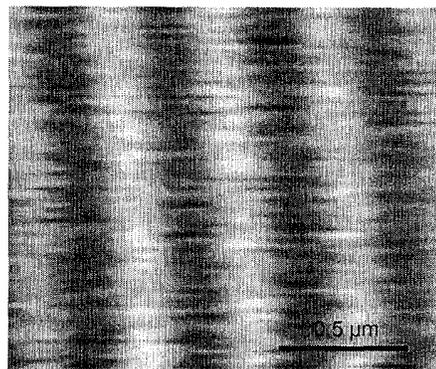
Figure 7. *Two-dimensional image of a 0.2 µm line-width Ti-on-quartz grating taken in transmission mode by scanning the grating beneath a HeNe laser focussed through the micromachined Si SIL.*

The resolution of the image taken with the SIL was estimated from the amplitude variation to be a factor of 2.8 greater than the same image taken without the SIL. When a high-angle objective is used to illuminate the SIL, refraction at the lens surface does not contribute significantly to a reduction in spot size. Aberrations in the incident beam or offsets in alignment may account differences from the maximum factor of 3.8 improvement from index of refraction. The minimum feature size that can be resolved with this technique is currently being investigated. Recent experiments at λ = 9.2 µm indicate that resolution as high as λ / 6 can be achieved with the Si SIL.

CONCLUSIONS

Solid Immersion Lenses micromachined from high-index materials will improve optical resolution in applications such as imaging and lithography. Micromachining offers a repeatable method for batch fabrication, and small lenses have reduced spherical aberration and improved tolerance to lens shape and wavefront errors. We have demonstrated the fabrication and operation of micromachined refractive lenses on the order of 10 µm in diameter. The lenses are integrated onto a cantilever for scanning and made with a tip to localize lens-sample contact. Results indicate that improvements in minimum resolution by a factor of almost 3 can be achieved with Si lenses.

ACKNOWLEDGEMENTS

REFERENCES

1. E. Betzig and J. K. Trautman, "Near-Field Optics: Microscopy, Spectroscopy, and Surface Modification Beyond the Diffraction Limit", *Science, 257,* 189 (1992).

2. S. M. Mansfield and G. S. Kino, "Solid Immersion Microscope", *Applied Physics Letters, 57 (24)*, 2615 (1990).

3. B. D. Terris, H. J. Mamin, D. Rugar, W. R. Studenmund, and G. S. Kino, "Near-field optical data storage using a solid immersion lens," *Applied Physics Letters, 65 (4)*, 388 (1994).

4. L. P. Ghislain, V. B. Elings, K. B. Crozier, S. R. Manalis, S. C. Minne, K. Wilder, G. S. Kino, and C. F. Quate, "Near-field photolithography with a solid immersion lens," *Applied Physics Letters, 74 (4)*, 501 (1999).

5. Q. Wu, G. D. Feke, R. D. Grober, and L. P. Ghislain, "Realization of numerical aperture 2.0 using a gallium phosphide solid immersion lens," *Applied Physics Letters, 75 (26)*, 4064 (1999).

6. K. Kishima, I. Ichimura, K. Yamamoto, K. Osato, Y. Kuroda, and K. Saito, "Near-Field Phase-Change Optical Recording over 1.2-Numerical-Aperture," *Proceedings of the 1999 International Symposium on Optical Memory and Optical Data Storage*, Hawaii, 7/11-15/1999, SPIE (1999).

7. G. S. Kino, D. A. Fletcher, K. B. Crozier, K. E. Goodson, and C. F. Quate, "An Efficient Near-field Microscope for Thermal Measurements," *Proceedings of the Seventeenth Symposium on Energy Engineering Sciences*, Argonne National Labs., 5/13-14/99, Department of Engergy (1999).

8. M. Born and E. Wolf, *Principles of Optics*, 6th edition, Cambridge: Cambridge University Press, 1980.

9. D. A. Fletcher, K. B. Crozier, K. E. Goodson, C. F. Quate, and G. S. Kino, "Optical characterization of microfabricated solid immersion lenses," *Proceedings of the 2000 International Symposium on Biomedical Optics, Photonics West*, San Jose, CA 1/22-28/2000, The International Society for Optical Engineering (SPIE), Bellingham, WA (2000).

10. D. A. Fletcher, K. B. Crozier, G. S. Kino, C. F. Quate, and K. E. Goodson, "Fabrication of High-Index Refractive Microlenses for Near-Field Optics", *Proceedings of the 1999 ASME International Mechanical Engineering Congress and Exposition, MEMS Symposium,* Nashville, TN, 11/14-19/99, American Society of Mechanical Engineers, New York (1999).

ATOMIC FORCE MICROSCOPE FOR PLANETARY APPLICATIONS

T. Akiyama, S. Gautsch, N.F. de Rooij, U. Staufer
Institute of Microtechnology, Univ. of Neuchâtel, Jaquet-Droz 1, 2007 Neuchâtel, Switzerland.
Ph. Niedermann
CSEM, Jaquet-Droz 1, 2007 Neuchâtel, Switzerland.
L. Howald, and D. Müller
Nanosurf AG, Austrasse 4, 4410 Liestal, Switzerland.
A. Tonin, and H.-R. Hidber
Insitute of Physics, Univ. of Basel, Klingelbergstr. 82 4056 Basel, Switzerland
W. T. Pike, M. H. Hecht
Jet Propulsion Laboratory, 4800 Oak Grove Dr. Pasadena, CA 91109-8099, USA

ABSTRACT

We have developed, built and tested an atomic force microscope (AFM) for planetary science applications, in particular for the study of Martian dust and soil. The system consists of a controller board, an electromagnetic scanner and a micro-fabricated sensor-chip. Eight cantilevers with integrated, piezoresistive deflection sensors are aligned in a row and are engaged one after the other to provide redundancy in case of tip or cantilever failure. Silicon and molded diamond tips are used for probing the sample. Images can be recorded in both, static and dynamic operation mode. In the latter case, excitation of the resonance frequencies of the cantilevers is achieved by vibrating the whole chip with a piezoelectric disk.

INTRODUCTION

Recent data [1, 2] from both Mars Pathfinder and Mars Global Surveyor indicate a vigorous recirculation of dust between the Martian surface and atmosphere. Dust devils scour the ground, leaving myriad dark trails, lofted material colors the sky, and global dust storms envelop the planet. Although properties of the dust have been inferred from remote sensing, there has been no imaging of individual dust and soil particles to determine their size distribution and shape. Such information is essential in both understanding the contribution of the particles to the Martian dynamics, and assessing the harmful effects of the dust on both robotic and human missions to the surface of Mars. NASA's Mars Environmental Compatibility Assessment (MECA) payload, which

will be sent to Mars in the next three years, contains a microscopy station to produce images of dust and soil particles. Mars Pathfinder data indicates that the mean particle size of Martian atmospheric dust is less than 2 micrometers. Hence MECA's microscopy station, in addition to an optical microsope capable of taking color and ultraviolet fluorescent images, includes an AFM to image far below optical resolution (fig. 1). The sample-handling system consists of an external robot arm, for delivery of surface and subsurface soil samples, and a two-degree-of-freedom stage. The stage contains 69 substrates that can be rotated in turn to the field of view of the two microscopes, and then translated for focussing and AFM approach in 0.25μm steps.

The AFM is well-established instruments for imaging conductive and insulating samples below the resolution limit of optical microscopes in laboratories and production environment. A sharp tip mounted on a thin cantilever beam is brought into close proximity to the sample surface. Forces acting between the sample and the probe deflect the cantilever. This deflection is monitored and provides the topographic image of the sample. Alternatively, the cantilever is excited at its resonance frequency. This resonance is de-tuned when exposed to the force gradient above the sample surface. This signal can be used for imaging the sample in a comparable way as the lever bending in the above-mentioned static mode. This second, dynamic mode has the advantage that lateral forces between the tip and the sample are minimized. In the imaging of particles dynamic mode is therefore preferred as it has been found that the tip does not push the particles around during scanning.

The requirements on an AFM for space applications and
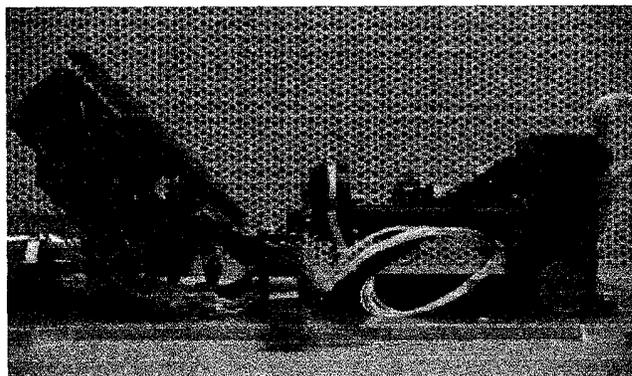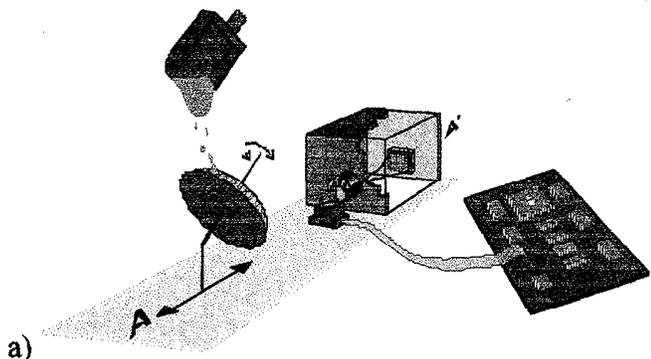
Figure 1 Set-up of the microscopy experiment of NASA's 2001 lander(a). A rotating sample stage delivers the soil collected by the robotic arm and brings it in front of the optical microscope and the AFM. 1b shows a photograph of the actual sample stage, AFM scanner and optical microscope.
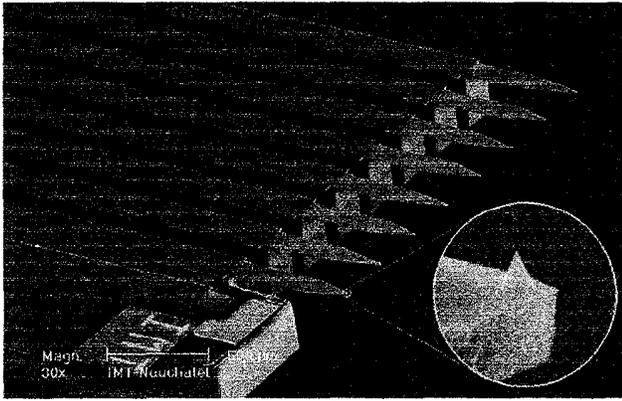
Figure 2. SEM picture of the microfabricated AFM chip with support beams etched by DRIE. The inset shows the silicon tip of cantilever 8. The CVD molded diamond tip can be viewed in fig 5. On the bottom part of the picture, the thermal compensation or with its protection bar is shown..

Figure 3. Method for removal of damaged cantilevers and support beams.

planetary research are quite different compared to those for laboratory use. Most important are weight and volume limitations, operation temperatures and shock and vibration conditions during launch and landing of the space vehicle. The electronic controller must be radiation hard or, at least, radiation tolerant. Moreover, there will be no operator for optimizing measuring parameters or tip exchange. Hence the AFM should be capable of self initialization, operate well under a broad range of conditions, and be capable of autonomous tip exchange, all within tight payload constraints.

AFM DESIGN AND FABRICATION

Integration from the initial design of the AFM with the optical microscope and sample-handing hardware, together with the development of an intelligent control system were necessary in achieving the required performance within the system constraints. In addition, maximum functionality was transferred to a microfabricated implementation of the AFM (fig. 2). First, the stress sensor is integrated into the micro-fabricated silicon cantilever. Second, means for exchanging blunt tips or broken cantilevers is provided through the use of a cantilever array. Third, diamond was incorporated as a low-wear material for selected tips in the array. Through this concentration of functionality into the microfabricated component of the instrument, constraints could be relaxed for the other subsystems; the scanner and servomotor as well as the electronic controller could be build with conventional technology, albeit tailored to perform reliably in the expected environments of the mission.

The cantilever deflection is measured by means of implanted piezoresistors in a Wheatstone bridge configuration [3] A special reference resistor is incorporated on an ultra-short cantilever for compensating thermal drifts. This reference cantilever is protected against mechanical damage by a surrounding, rigid safety bar. Electrical isolation between the resistors and the cantilevers is achieved by reverse biasing the pn-junction formed between the p-type resistors and the n-type bulk of the cantilever. This limits, to some extent, the range of usable tip- potential which can only be
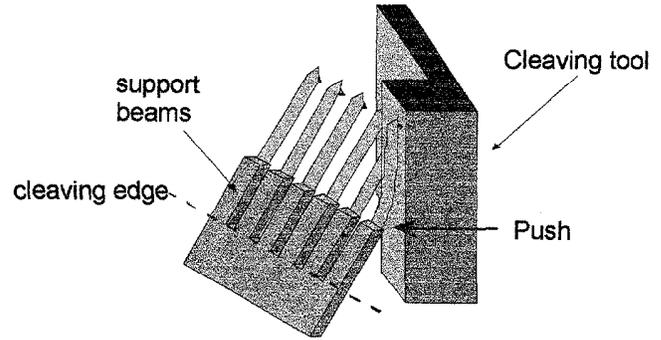
applied via the cantilever. It is expected that in the dry Martian atmosphere electrostatic forces between the sample and the tip will be present. The ability to change the tip potential is therefore important to at least partially compensate for these forces.

Broken tips will be permanently removed by means of a cleaving tool mounted in one of the substrate positions on the sample wheel. Therefore the AFM-chip features eight cantilevers for redundancy. The piezoresistors are contacted by wire-bonding and can be individually addressed via an external multiplexer. The AFM-chip is mounted on the scanner with two tilt angles relative to the substrate plane to ensure that only one tip at a time is in the lowest, imaging position. The cantilever tips are alternately equipped with monolithic silicon tips and CVD molded diamond tips.

The length of the cantilevers is alternately 580μm and or 610μm, their width is 160μm and thickness about 8μm. The spacing between two neighboring levers is $350\,\mu$m. Thus the whole chip has a width of at least 2.6 mm. This implies directly, that simply breaking of the cantilevers for removing blunt tips is not sufficient: the edge of the chip at the first cantilever would touch the sample when measuring with the last few cantilevers. Therefore, the levers reside on rigid support beams rather than on the bulk of the chip. It is this support beam that is cleaved for cantilever removal (see fig. 3).

The space between two support beams, which have the full wafer thickness, is about 90μm. Therefore, the standard KOH etching technique cannot be used for bulk-machining the AFM-chip and must be replaced by anisotropic, deep reactive ion etching (DRIE) [4].

The electronic circuit for dynamic mode imaging is based on a PLL frequency-shift detection. The frequency shift of the oscillation is detected by keeping a constant phase difference between detection and driving signal. Thus, the phase transition value has to be specified during autonomous initialization of the AFM for dynamic mode operation. During this process cross-talk between cantilevers is of concern since the whole chip and therefore all cantilevers are excited at the same time. We expected a relatively strong coupling between neighboring levers. Therefore, two different lever lengths are used to separate their respective resonances by 6 kHz. This allows electronic detection of the desired peak within a reasonable frequency span without interference due to cross-talk by the neighboring cantilever. The choice of the correct phase transition can then be determined
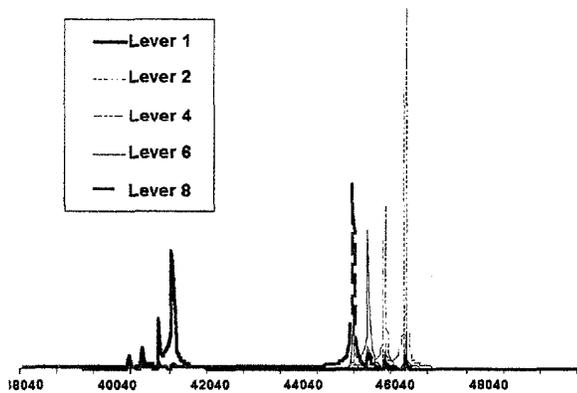
Figure 4. Simulated frequency response of the piezoresistors for cantilever 1 (low frequency) and cantilevers 2, 4, 6, 8 (high frequency). Cross-talk peaks induced by each cantilever on each cantilever.


Figure 5. Process-flow chart of AFM chip:
a) Thermal oxide, lithography, oxide etching, KOH etching.
b) CVD PSG deposition, CVD oxide deposition, Lithography, Oxide etching, Implantation, removing resist, CVD oxide deposition, thermal treatment (950°C, 30 min.).
c) Oxide etching.
d) CVD oxide deposition.
e) Lithography, Contact hole opening, Al evaporation, Lithography, Al etching.
f) Plasma SiN deposition, Lithography, SiN Plasma etching.
g) Lithography, Oxide etching, Topside Si etching.
h) Front side lithography, backside lithography, backside DRIE etching.
i) Oxided etching.

autonomously by the software by simply cycling through all possible phase values and detecting the highest signal amplitude.

Finite element analysis of the array and experimental evaluation showed, however, that cross-talk is important not just between neighboring cantilevers but between all even numbered cantilevers (2, 4, 6 and 8). We attribute this characteristic to the chip's asymmetry caused by the compensation sensor on the far right of the chip. Since these 4 cantilevers have approximately the same resonance frequency, the spectral response for each of them shows multiple peaks within a very narrow frequency span. It has been shown and tested that each of the peaks seen on the frequency response of a certain cantilever can be used as a reference for dynamic mode measurements.

However, changing pressures and temperatures on Mars might temporarily influence the cross-talk behavior and modify the amplitude distribution over the 4 peaks. Thus, the autonomous phase detection algorithm might detect a cross-talk induced phase shift that corresponds to a signal too low in amplitude, rendering dynamic measurements impossible. This problem could be overcome by having a more powerful autonomous algorithm for initializing the AFM in dynamic mode. For future designs, however, it is better to distribute the resonance frequencies of all 8 cantilevers evenly across the whole array. Figure 4 depicts the spectrum obtained by finite element analysis of a lever array. The analysis chip model has been created without a thermal compensation sensor and is symmetric. The cross-talk is present on all cantilevers.

The details of microfabrication are illustrated in the process-flow chart in fig. 5. Silicon-on-insulator wafers are used as base material. The device layer has a thickness of 20 μm and the support wafer is 380 μm thick. In a first step (fig. 5a), 8μm high silicon tips are fabricated for every second cantilever by wet etching the device layer in KOH and using square thermal-oxide masks. The other four cantilevers are equipped with diamond tips later. Then phosphor-silicate glass (PSG) and silicon-oxide films are deposited in a chemical-vapor deposition (CVD) process. These films are then structured by photolithography and wet etching. The photoresist and the oxide were used as mask during

ion-implantation (fig. 5b). After stripping the mask, CVD oxide was again deposited, followed by a thermal treatment at 950 C for 30 minutes fto activate the dopants. The oxide is replaced by a fresh layer for passivation and protection during subsequent processing (figs. 5c and d respectively). A third lithography structures a resist layer used as mask for wet etching the contact holes to the piezoresistors. A layer of aluminum is deposited and wet etched after the next lithography to define the contact paths and pads (fig. 5e). Figure 5f depicts the passivation process of the chip through deposition of a plasma nitride layer and a fifth lithography to structure it. This nitride structuring is used to open the contact pads. The oxide layer remaining on the silicon is structured by lithography and wet etching and used as a mask to define the top shape of the cantilevers (fig. 5g). Fig. 5h shows the last top lithography for removing the remaining oxide on the cantilevers. The top resist is also used as a protection layer for the backside deep reactive ion etching (DRIE) which defines the support beams and the final chip shape. For this backside etching, lithography on the back of the wafer is performed. DRIE allows etching through the full wafer with almost vertical side walls, which is important for maintaining a small pitch of 350 μm in the lever array. The oxide layer underneath the cantilever is used as an etch stop. As a final step, the oxide remaining underneath the

269
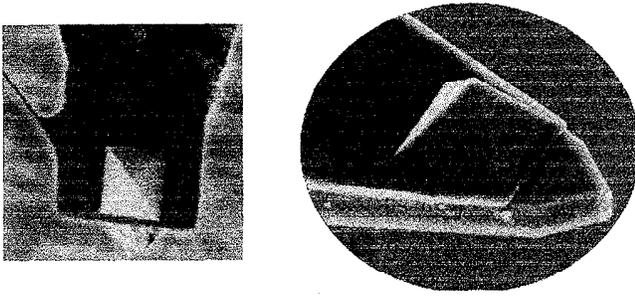
Figure 6. *FIB Trimming of diamond tip to fit cantilever shape. The left image shows the diamond tip with its support frame glued on the cantilever before FIB trimming. The left image shows the diamond tip after having been trimmed by focussed ion beam.*

cantilever is removed by wet etching (fig. 5i). After this step, the individual AFM chips are still connected to a silicon frame via thin bridges. They are easily broken for removing the chips.

The diamond tips are fabricated by CVD into pyramidal silicon molds [5]. The tips are then bonded to a special pyrex substrate. The Si mould can then be removed in KOH. Mounting of the diamond tips on the cantilevers is performed next, working on a chip-by-chip base and using a micro-manipulator. Finally, parts of the diamond tip which jut the border of the cantilever are trimmed by a focused ion beam (Fig. 6). The chip is directly glued to the scanner. The electrical contacts are wire bonded.

The mounted chip has withstood shocks up to 3000 g acceleration at 3 kHz, and the anticipated vibrations and thermal cycling of the mission.

RESULTS

Figures 7 and 8 show the image of a calibration grid (10 μm pitch) and a Mars-analog sample (Fe_2O_3-particles), both measured with the Mars AFM system.
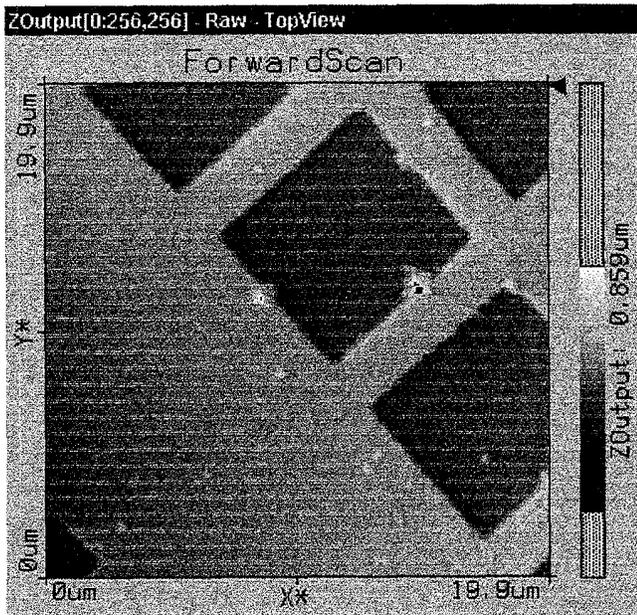
Figure 7. *AFM dynamic mode image of a 10 um pitch calibration sample. The step of the grid is 200 nm.*
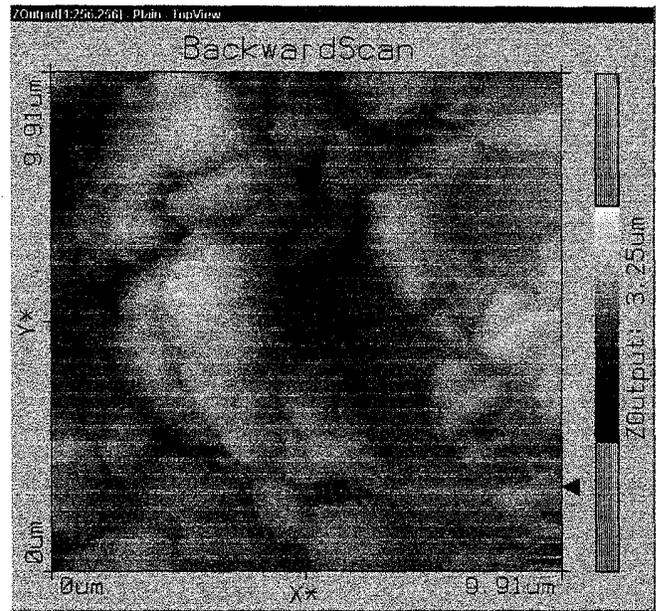
Figure 8. *AFM dynamic mode image of Mars equivalent dust particles (Fe_2O_3).*

ACKNOWLEDGEMENTS

REFERENCES

1 W. J. Markiewicz, R. M. Sablotny, H. U. Keller, N. Thomas, D. Titiv, P. H. Smith, "Properties of Martian aerosols derived from Imager for Mars Pathfinder midday sky brightness data", J. Geophys Res - Planets, **104**, 9009 (1999).

2. M. Malin http://www.msss.com/mars-images/moc (2000).

3 M. Tortonese, R.C. Barett, C.F. Quate, "Atomic resolution with an atomic force microscope using piezoresistive detection" *Appl. Phys. Lett.* **62**, 834, (1993).

4. P.-A. Clerc, L. Dellmann, F. Grétillat, M.-A. Grétillat, P.-F. Indermühle, S. Jeanneret, Ph. Luginbühl, C. Marxer, T.L. Pfeffer, G.-A. Racie, S. Roth, U. Staufer, C. Stebler, P. Thiébaud, and N.F. de Rooij, "Advanced deep reactive ion etching: a versatile tool for microelectromechanical systems", *J. Micromech. Microeng.* **8**, 272 (1998).

5. Niedermann, W. Hänni, S. Thurre, M. Gjoni, A. Perret, N. PH Skinner, P.-F. Indermühle, U. Staufer, N. F. de Rooij, "Mounting of micromachined diamond tips and cantilevers", *Surf. and interface anal.* **27**, 296 (1999).

SHIELDED-TIP/CANTILEVER PROCESS AND INTERFACE FOR MULTIFUNCTIONAL SCANNING PROBE MICROSCOPY

Toralf Bork*, Vivek Agrawal*, Björn Rosner, Patrick Gustafson, and Daniel van der Weide
Department of Electrical & Computer Engineering
University of Wisconsin
Madison, WI 53706
*Department of Electrical & Computer Engineering
University of Delaware
Newark, DE 19716

ABSTRACT

Along with measuring a sample's topography, coaxial scanning probe microscope (SPM) tips enable electric field imaging [1-3] with completely shielded tips, and both optical [4-6] and microwave [7] detection when diode junctions are formed at the tips. Integrating such tips with appropriate cantilevers and interfaces to external instruments will further promote their applications in scanning capacitance and thermal microscopy, and in electrochemistry and biology. We have batch-fabricated both coaxial and junction tips integrated with matching cantilevers on silicon. We describe the process, show initial results from junction tips as photodetectors, and discuss an electrical and mechanical interface from the probe body to external instrumentation.

INTRODUCTION

Near-field scanning optical microscopy [8-11] (NSOM) is a technique by which visible or near-visible light (free-space wavelength generally less than one micrometer) is confined within a sub-wavelength aperture (or is scattered by a sub-wavelength tip [12]) to excite or detect evanescent waves on a sample. A propagating wave incident on a small aperture or scattering center creates evanescent fields. By reciprocity, the aperture, when in proximity to such fields, should radiate propagating waves. To circumvent the far-field resolution Rayleigh limit of $\lambda/2$ requires that we scan the aperture over the sample. NSOM in the visible has been applied to a growing variety of imaging applications, from localized spectroscopy on semiconductor quantum wells [13] to observing fast thermal transport through conductors on integrated circuits [14] to observing ion transport through cell membranes [15].

Obtaining sub-Rayleigh resolution by trading the multiplex advantage of conventional optics for serial image acquisition is, however, a general principle, and need not be restricted to visible radiation. Ash and Nichols demonstrated this in 1972 using 3 cm wavelength microwaves confined to a sub-wavelength aperture to achieve 1/60 wavelength-relative resolution [16]. Since resolution is no longer limited by the wavelength of illumination, it is now possible to build microscopes for frequencies below the visible regime. There are myriad electromagnetic phenomena with characteristic frequencies below the visible but above the radio. Examples include integrated circuit (IC) electromagnetic fields (between 100 MHz and 300 GHz), biological membrane absorption in the far-infrared (FIR) range of 100 GHz to 30 THz (10 μm), molecular rotational or vibrational absorption (also primarily in FIR), and conductive or dielectric properties of materials. Just as far field optics enable traditional spectroscopy, which measures the collective response of large numbers of elements (such as molecular electric dipoles), near field techniques can now enable localized spectroscopy, though momentum is no longer conserved, changing the absorption and excitation spectra, especially in the visible and near-infrared.

COAXIAL PROBES FOR NEAR-FIELD MICROSCOPY IN THE SUB-VISIBLE

One approach to microscopy at wavelengths longer than visible light is to use near field antennas–primarily coaxial probes for electric field sensitivity. Seminal work demonstrating sub-wavelength field resolution with a coaxial tip was done by Fee, et al. [17], and using open-ended coaxial waveguide for measuring bulk microwave dielectric properties of materials is a common technique [18].

The advantage of using coaxial waveguide for near field imaging in the sub-visible regime is that the shield greatly reduces interference from stray radiation while enabling a non-cutoff transverse electromagnetic (TEM) mode of wave propagation up to or away from the tip. Furthermore, the tip itself can be sharpened to increase the "lightning rod" effect of field concentration, improving resolution. Finally, if the tip is sufficiently sharp, it can also function independently as an SPM topography probe. This multifunctional ability is particularly important in submicrometer field resolution, since the tip-sample distance should be stabilized independent of the field measurement.

Near field antenna probes can be used in different modes, depending on the image to be acquired. These include reflection, transmission, excitation and detection. Reflection and transmission rely on external instruments (lasers or signal generators) for both generating and detecting spectroscopic energy, while excitation might stimulate a photoconductive response, like a change in current through a semiconductor or cell membrane. Detection measures an excited or active sample, such as a waveguide or a circuit; for this mode, integrated photodetectors have been developed [4-6].

In reflection mode, the most widely used by those interested in high spatial resolution of microwave material properties [19-22], the same coaxial probe tip conducts energy both to and from the sample. Although it is necessary to have a closed loop SPM platform to acquire sub-micrometer topography and field images, open loop scanning with sharpened coaxial tips has produced quantitative spatially resolved material measurements such as surface resistivity, especially when the tips are coupled to resonant cavities to enhance the signal.

Although resonant measurements offer high sensitivity, they are also limited in frequency coverage, stimulating the

development of broadband probes for localized spectroscopy. With sufficiently sensitive microwave instrumentation, such as a vector network analyzer or bridge detector, broadband probes would offer the widest range of applications. Broadband capability is equally important for the transmission mode of sub-visible microscopy [23], in which one probe acts as a source while another on the opposite side of a transmitting sample acts as a detector.

To achieve the highest spatial resolution, both in topography and field, we have developed multifunctional SPM tips with integral near field antennas, both magnetic field loop probes [24] and silicon coaxial probe tips [1]. When integrated with broadband transmission lines and SFM cantilevers as shown here, these tips can provide field and topographic resolution commensurate with their tip radii, which can be ~ 10 nm.

The final two modes, excitation and detection, can be realized either with external microwave instruments or, in the case of detection, with a sensitive element integrated directly into the antenna. By modifying a commercial SFM tip/cantilever to have coaxial shielding, and connecting this probe to a 50 GHz sampling oscilloscope, we have measured picosecond waveforms and with both submicrometer field and topographic resolution on an IC [2], and have employed a microwave detector diode directly at the coaxial center conductor [7]. With this structure, using a lock-in amplifier we were able to detect the magnitude and phase of a 2 GHz signal on an interdigitated capacitor sample, as shown in Figure 5. This type of integrated detector has also been used for visible NSOM [4], and it considerably simplifies the instrumentation needed for localized microwave detection.

PROCESS AND INTERFACE DETAILS

The process for both coaxial and junction tips is similar, but the interface from the cantilever to the outside world is different: for low-loss microwave propagation from coaxial tips we suspend a controlled-impedance coplanar waveguide (CPW) transmission line on a dielectric membrane using XeF_2 to undercut the structure [25] as shown in Figs. 1 and 2. The common steps of the process for both types of tips, using a 350μm, (100) 3" double-side polished, n-doped, 1-20 Ohm-cm, TTV ~ 2μm wafer, are:

1. LPCVD Nitride deposition at 850°C for 30min as a mask for KOH etching
2. FSC spinning to protect the front side of wafer
3. Lithography: Pattern the nitride as a mask for backside KOH etch to form the chip body
4. Nitride etch using CF_4 plasma process
5. Body formation. 30% KOH at 80°C to etch to ~270μm deep
6. Removal of all nitride on front side using CF_4
7. 1.2μm PECVD oxide deposition to mask for cantilever and tip formation
8. Lithography: Pattern the oxide, which acts as a mask for the silicon etch to form the cantilevers
9. Oxide etch to form the cantilever using BHF 6:1
10. Cantilever formation: Using anisotropic plasma process to form Si cantilevers
11. Lithography: Pattern the oxide as a mask for the silicon etch to form the tips
12. Oxide etch to form the cantilever using BHF 6:1
13. Cantilever formation: Using anisotropic plasma process to form Si tips, resulting in a pillar with a ~ 1μm waist diameter.
14. Removal of all oxide

15. Over-oxidization of pillars to form and sharpen tips. The oxide is grown for 5h at 1100°C, yielding ~1.5μm. All oxide can be removed (HF 48% ~ 3min) and applied again to form sharper tips
16. Growth of ~ 270nm oxide (1.5h at 925°C)
17. Lithography: Pattern the insulating oxide
18. BHF 6:1 to pattern the insulating oxide
19. Ion implantat ohmic contacts between the ground metal planes and the Si body

Steps 20 and 21 for photodiode tips only:
20. Lithography: Open the oxide at the tip by spinning resist and opening it at the top of the tips using O_2 plasma
21. BHF 6:1 to open the oxide at the tip

22. Thermal evaporation of 1000Å metal (Au or Al)
23. Lithography: Patterning the metal waveguide
24. Metal etch to form the waveguide.
25. Lithography: Opening of the metal at the tip only by applying resist and opening it with O_2 plasma at tip
26. Metal etch and oxide etch in case of coax tips
27. PECVD Oxide deposition to protect front side of wafer in release step

27a-c for coaxial tips:
27a. Lithography: Pattern front side oxide to form XeF_2 etch pattern
27b. BHF 6:1 oxide etch to form the XeF_2 etch holes
27c. XeF_2 etch to form suspended center waveguide (see more detailed explanation following this list)

28. Silicon etch to release back side using anisotropic plasma process
29. Oxide etch to remove protective oxide on front side
30. Al evaporation on backside to enhance reflectance from cantilever
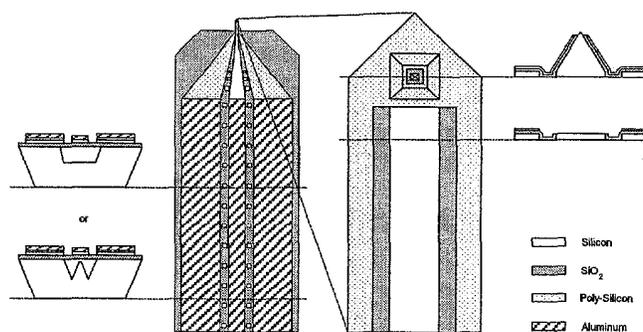
Figure 1. Coaxial SPM tip/cantilever layout. Coplanar waveguide (CPW) transmission line has windows in its gap for XeF_2 etching and suspension on oxide membrane.

To undercut the center waveguide, we built and used a small XeF_2 etching chamber. It incorporates the chamber, mechanical vacuum pump, etchant reservoir, nitrogen purge line, and computer-controlled valve actuation and pressure acquisition.

A long-working-distance microscope is necessary to observe the etch process since its speed depends on sample geometry, exposed area and etchant pressure.

In order to etch, a single wafer is loaded into the chamber and pressure is reduced to 10 mT. Exposing the XeF_2 reservoir to the

same pressure will sublimate some of the XeF_2 powder until a pressure of about 2 Torr is reached. This gas is introduced into the etching chamber where it etches any exposed Si surface. For this reason, the backside and rim of the wafer has to be covered with a layer of non-etching tape. The XeF_2 is allowed to stay in the etching chamber for about 90-120 seconds after which the chamber is purged and new etchant is introduced. After about 15 cycles, slight warping of the center waveguide can be observed (Figure 2). This is a sign that it has been completely underetched, connected only to the oxide above and the 3-4 suspension posts along the waveguide which are necessary for mechanical stability.
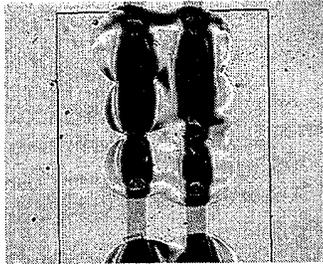
Figure 2. Center waveguide after XeF_2 etching.

As an important addition to microfabricating the probe tip, we have designed an interface to quickly exchange scanning tips on the SFM scanner and reconnect its electrical connections without adding height that could hamper SFM operation (Figure 3).
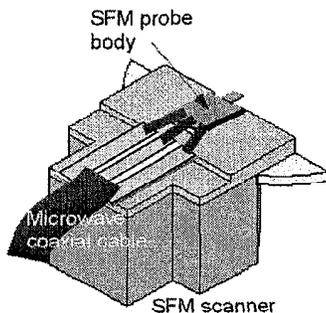
Figure 3. Interface between coaxial tip/SFM body and microwave cable, shown on SFM scanner.

The interface consists of a clamp holder for the SFM chip and scanner, a waveguide transition and a permanently attached microwave coaxial cable. The connection between SFM chip and waveguide transition can be established using gold wire bonding or small clamping springs.

TOPOGRAPHY AND OPTICAL RESULTS

As shown in the following scans, the coaxial nature of the SFM tip (Fig. 4) does not affect images of typical IC-like structures. Further tests of the probe's optical response showed that it is single-aperture with no parasitic junctions.
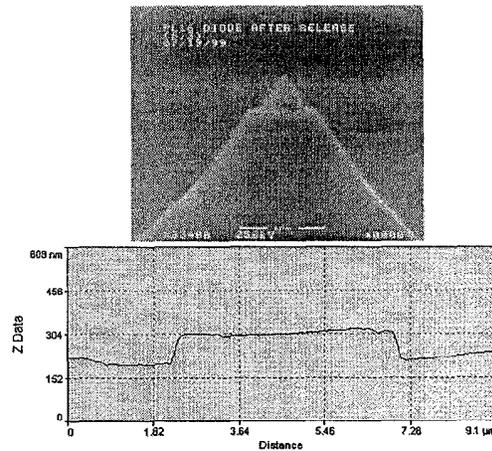
Figure 4. Coaxial tip (top) and line scan using tip (bottom)
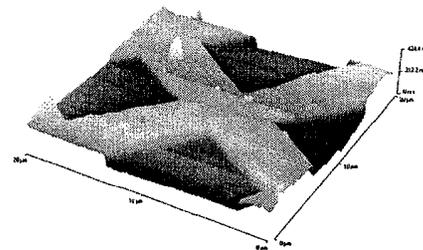
Figure 5. Topographical image of test sample.
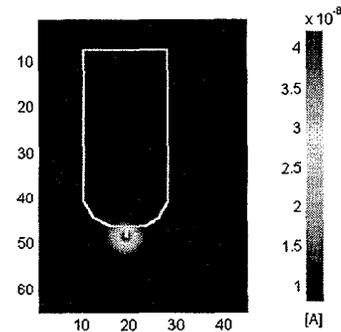
Figure 6. Scan with HeNe laser over photodetector tip and body (20 x 30 mm area).

The process we have discussed enables complete integration of coaxial (and photodiode) probe tips, low-loss microwave transmission lines and SFM cantilevers. These will enable a wide range of near-field microwave and optical investigations.

REFERENCES

[1]   D. W. van der Weide and P. Neuzil, "The nanoscilloscope: Combined topography and AC field probing with a micromachined tip," *Journal of Vacuum Science & Technology B, 14,* 4144 (1996).

[2] D. W. van der Weide, "Localized picosecond resolution with a near-field microwave/scanning-force microscope," *Applied Physics Letters, 70,* 677 (1997).

[3] D. W. van der Weide, "Microscopes for the sub-visible: scanning the near field in the microwave through infrared," *Optics and Photonics News, 9,* 40 (1998).

[4] R. C. Davis, C. C. Williams, and P. Neuzil, "Micromachined submicrometer photodiode for scanning probe microscopy," *Applied Physics Letters, 66,* 2309 (1995).

[5] T. Leinhos, M. Stopka, and E. Oesterschulze, "Micromachined fabrication of Si cantilevers with Schottky diodes integrated in the tip," *Appl. Phys. A, 66,* 65 (1998).

[6] D. Drews, W. Ehrfeld, M. Lacher, K. Mayr, W. Noell, S. Schmitt, and M. Abraham, "Nanostructured probes for scanning near-field optical microscopy," *Nanotechnology, 10,* 61 (1999).

[7] D. W. van der Weide, V. Agrawal, T. Bork, and P. Neuzil, "Localized circuit probing with a combined Schottky diode/scanning force microscope," *Technical Digest of the IEEE Microwave Theory and Techniques Symposium,* Baltimore, IEEE (1998), pp. 1341.

[8] U. Dürig, D. Pohl, and H. Rohrer, "Near-field optical scanning microscopy," *Journal of Applied Physics, 59,* 3318 (1986).

[9] E. Betzig, A. Lewis, A. Harootunian, M. Isaacson, and E. Kratschmer, "Near-field scanning optical microscopy (NSOM)," *Biophys. J., 49,* 269 (1986).

[10] E. Betzig and J. K. Trautman, "Near-Field Optics: Microscopy, Spectroscopy, and Surface Modification Beyond the Diffraction Limit," *Science, 257,* 189 (1992).

[11] D. Courjon and C. Bainier, "Near field microscopy and near field optics," *Rep. Prog. Phys., 57,* 989 (1994).

[12] F. Zenhausern, M. P. O'Boyle, and H. K. Wickramasinghe, "Apertureless near-field optical microscope," *Applied Physics Letters, 65,* 1623 (1994).

[13] P. A. Crowell, D. K. Young, S. Keller, E. L. Hu, and D. D. Awschalom, "Near-field scanning optical spectroscopy of an InGaN quantum well," *Applied Physics Letters, 72,* 927 (1998).

[14] K. E. Goodson and M. Asheghi, "Near-field optical thermometry," *Microscale Thermophysical Engineering, 1,* 225 (1997).

[15] P. G. Haydon, S. Marchese-Ragona, T. A. Basarsky, M. Szulczewski, and M. McCloskey, "Near-field confocal optical spectroscopy (NCOS): Subdiffraction optical resolution for biological systems," *Journal Of Microscopy, 182,* 208 (1996).

[16] E. A. Ash and G. Nicholls, "Super-resolution aperture scanning microscope," *Nature, 237,* 510 (1972).

[17] M. Fee, S. Chu, and T. W. Hansch, "Scanning electromagnetic transmission line microscope with sub-wavelength resolution," *Optics Communications, 69,* 219 (1989).

[18] E. Tanabe and W. T. Joines, "A nondestructive method for measuring the complex permittivity of dielectric materials at microwave frequencies using an open transmission line resonator," *IEEE Transactions on Instrumentation and Measurement, 25,* 222 (1976).

[19] C. P. Vlahacos, R. C. Black, S. M. Anlage, A. Amar, and F. C. Wellstood, "Near-field scanning microwave microscope with 100 mu m resolution," *Applied Physics Letters, 69,* 3272 (1996).

[20] D. E. Steinhauer, C. P. Vlahacos, S. K. Dutta, B. J. Feenstra, F. C. Wellstood, and S. M. Anlage, "Quantitative imaging of sheet resistance with a scanning near-field microwave microscope," *Applied Physics Letters, 72,* 861 (1998).

[21] T. Wei, X. D. Xiang, W. G. Wallace-Freedman, and P. G. Schultz, "Scanning tip microwave near-field microscope," *Applied Physics Letters, 68,* 3506 (1996).

[22] Y. Lu, T. Wei, F. Duewer, Y. Lu, N.-B. Ming, P. G. Schultz, and X. D. Xiang, "Nondestructive imaging of dielectric-contrast profiles and ferroelectric domains with a scanning-tip microwave near-field microscope," *Science, 276,* 2004 (1997).

[23] F. Keilmann, D. W. van der Weide, T. Eickelkamp, R. Merz, and D. Stockle, "Extreme sub-wavelength resolution with a scanning radio-frequency transmission microscope," *Optics Communications, 129,* 15 (1996).

[24] V. Agrawal, P. Neuzil, and D. W. van der Weide, "A microfabricated tip for simultaneous acquisition of sample topography and high-frequency magnetic field," *Applied Physics Letters, 71,* 2343 (1997).

[25] N. H. Tea, V. Milanovic, C. A. Zincke, J. S. Suehle, M. Gaitan, M. E. Zaghloul, and J. Geist, "Hybrid postprocessing etching for CMOS-compatible MEMS," *Journal of Microelectromechanical Systems, 6,* 363 (1997).

A MICRO HEAT EXCHANGER WITH INTEGRATED HEATERS AND THERMOMETERS

**Lian Zhang, Shilajeet S. Banerjee, Jae-Mo Koo, Daniel J. Laser, Mehdi Asheghi,
Kenneth E. Goodson, Juan G. Santiago, Thomas W. Kenny**
Department of Mechanical Engineering, Design Division, Stanford University
Stanford, CA 94305

ABSTRACT

Two-phase microchannel flow has the potential to remove large heat fluxes from high-performance integrated circuits. More research must study the behavior of microchannel heat exchangers, in particular the impact of small dimensions on bubble formation and the critical heat flux. This work develops microchannel test devices with integrated heaters and thermometers for studying single and two-phase convection. Fabrication techniques include DRIE and silicon-glass anodic bonding. DRIE is used to fabricate rectangular microchannels with a hydraulic diameter less than 100 μm in silicon chips. These devices include integrated semiconductor heaters and thermometers and are enabling detailed experiments on single and two-phase forced convective heat removal from microchannels. The heat exchangers remove 25 W from 2 cm^2 area with approximately 50 C average chip temperature rise using 5 ml/min of DI water, which remains in the liquid state. We observe water boiling in microchannels with a hydraulic diameter of 55 μm and a resulting transient temperature fluctuation in the microchannel system. Future work is leveraging and optimizing this experimental structure for imaging bubble formation and for studying the critical heat flux.

INTRODUCTION

The rapid development of high-speed, high-power, high-density integrated circuits requires compact devices for cooling. Almost 20 years ago, Tuckerman and Pease [1] predicted that single-phase forced convective cooling should be feasible for circuit power densities of more than 1000 W/cm^2, under large flow rates (above 600 ml/min). Since then, microchannel heat exchangers have been under constant research interest. Modeling and optimization of single-phase heat exchangers were well-studied [2-3], while theoretical and experimental studies on two-phase flow in microchannels were very limited. Peng et al. [4] investigated flow transition and heat transfer in V-shaped microchannels with a hydraulic diameter ranging from 200-600 μm. Jiang et al. [5] studied phase change in diamond-shaped microchannels with hydraulic diameters less than 100 μm. However, data interpretation in the previous research was impaired by a lack of detailed information about the rate of heat transfer into the fluid and the two-phase flow conditions inside the microchannels. Whether saturated nucleate boiling can or cannot occur in microchannels under 100 μm hydraulic diameter remains unknown.

The present work develops parallel-flow microchannel heat exchangers with integrated heaters and thermometers. This work is part of an ongoing effort to experimentally study both temperature distribution and heat transfer rates in microchannels. In our system, glass channel cover slips make it possible to directly observe the flow conditions inside the microchannels. Channels of 50-100 μm width and 50 μm depth are micromachined into a 2 cm by 6.5 cm silicon chip with an effective cooling area of 4 cm^2 or 2 cm^2. Heaters and thermometers are formed by ion implantation on the backside of the same chip to provide precise heat flux control as well as temperature distribution measurements.

DESIGN

For viscous internal fluid flow, fluid friction causes a pressure drop along the flow direction. For aspect ratios near unity, the pressure drop ΔP associated with this single-phase flow can be calculated from [6]

$$\Delta P = \rho \cdot \frac{64}{Re} \cdot \frac{LV^2}{2D_H} \qquad (1)$$

where ρ is the density of the fluid, Re is the Reynolds number, L is the length of the channel, V is the flow velocity, and D_H is the hydraulic diameter of the channel. The Reynolds number can be interpreted as the ratio of the flux of momentum to the viscous force within the channels. For channels in which the flow is predominantly fully developed, which is the case in our work, the pressure drop required to provide a certain mass flow rate is inversely proportional to Re. This relationship causes the pressure gradient along microchannels to be particularly large and, for a given set of dimensions, places an important lower bound on the pressure delivered by the system pump.

The advantage of this large pressure drop is a very high heat removal rate by forced convection. The approximate relationship between the single-phase fluid temperature rise along the microchannels ΔT_W and the heat rate input q is given by

$$q = \dot{m} c_p \Delta T_W \qquad (2)$$

where $\dot{m}$ is the fluid mass flow rate, and c_p is the fluid specific heat. Because of the higher fluid temperature near the exit, the local chip temperature also increases along the length of the channel. The increase in chip temperature along the channel length is reduced by conduction in the silicon, which augments the heat flux into the fluid near the inlet. However, the total heat rate is related to the average chip and fluid temperatures in

$$q = hA(T_{C,avg} - T_{W,avg}) \qquad (3)$$

where h is a length-averaged convection coefficient, A is the total wetted area, $T_{C,avg}$ and $T_{W,avg}$ are average chip and fluid temperatures respectively. The convection coefficient h can be increased by a factor of 100 to 1000 with forced convection, and under the assumption of uniform heating, it is independent of flow velocity in fully developed laminar flow, which has a Re number less than about 2300 [6].

Using implanted resistors as heaters and thermometers, the heat input q is determined by

$$q = UI \qquad (4)$$

where U and I are input voltage and current. Since semiconductor resistors have a temperature coefficient that can be calibrated, the local temperature can be determined from the

resistance variation of the resistors. Hence we are able to control the heat input to the fluid, as well as to monitor the local and average chip temperature.

Table 1. Design parameters of three types of microchannel heat exchangers.

	Type 1	Type 2	Type 3
Width (μm)	75	50	100
Depth (μm)	50	50	50
Length (cm)	2	2	2
Number of channels	50	20	30
Effective cooling area (cm^2)	4	2	4
Maximum flow rate (ml/min)	10	5	10
Hydraulic diameter (μm)	60	50	67
Reynolds number[*]	56	88	79
Pressure drop[*] (psi)	19.9	53.7	20.1

[*]*The Reynolds number and the pressure drop are calculated under the maximum water flow rate of the design type. Water properties are taken at ambient conditions.*
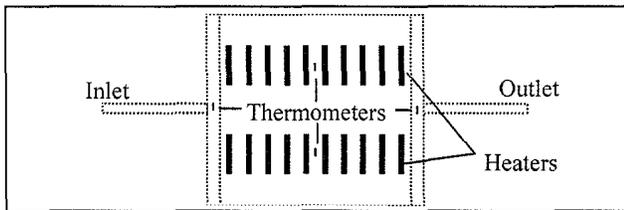
Figure 1. Schematic of the heaters and thermometers. The dashed lines outline the microchannels (the middle square) and the inlet/outlet on the channel side. On the back side, 20 heater strips are symmetrically distributed over the cooling area, connected in parallel; and 4 separate thermometers are located right at the entrance, exit and two middle spots.
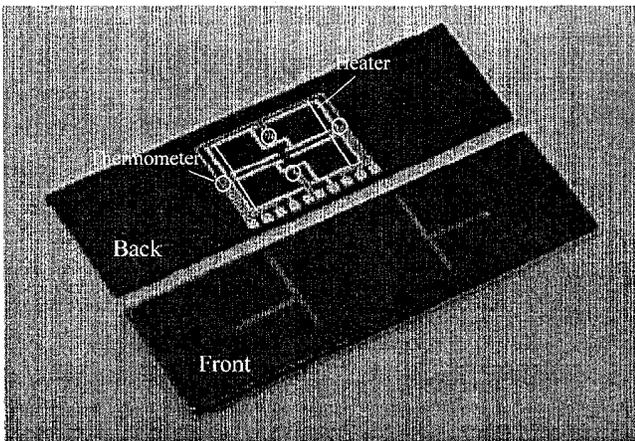
Figure 2. The micro heat exchanger chip. Both the front and back side of the heat exchanger are shown. On the front side, the two "T" shaped channels are 1mm wide inlet and outlet, the square between is the effective cooling area, which is 2 cm by 2 cm in this device. Locations of heaters and thermometers are marked on the back side.

The design considerations for the system include the pressure drop under the maximum flow rate of 5-10 ml/min, and the resistance and distribution of the heaters and thermometers. We use Equation (1) to calculate the maximum pressure in the

system; then estimate the water and chip temperature under a certain heat input rate from Equations (2) and (3); and finally decide the heater resistance based on the desired heat power. On the basis of trial calculations, we have selected three design types, whose parameters are listed in Table 1. Figures 1 and 2 show the layout of the heaters and the thermometers. The resistance of the parallel heaters is 100 Ω, while each of the four thermometers has a resistance of 2000 Ω.

FABRICATION

The fabrication process begins the fusion bonding of our own SOI (Silicon-On-Insulator) wafers. First, a pair of SiO_2 coated N-type <100> wafers are placed face to face and self-attached by Van der Waals force. An additional 1-hour annealing at 950 C gives very good fusion bonding. Then we grind the top Si wafer to a desired thickness of 50 μm. Second, heaters and thermometers are implanted from the support wafer side (backside). Because the parallel design of the heaters greatly lowers the total heater resistance, the heaters and the thermometers are implanted together. The implant dose is 1e14 Boron at 40keV. Since the heaters are operated under high power conditions, a long annealing time is necessary (we recommend 6 hours at 1150 C). The target surface resistivity after annealing is 200 Ω per square. Third, aluminum tracks and pads are deposited to make electrical connections. Fourth, DRIE is used to etch the channels from the SOI side, and then to etch the two inlet/outlet through holes from the backside (shown as dashed rectangle in Figure 3(C)--these are not in the same plane). And finally, a Pyrex7740 glass piece is anodically bonded to the front microchannel side to seal the channels.

(a) Fusion bonding SOI wafers

(b) Implanting heaters and thermometers
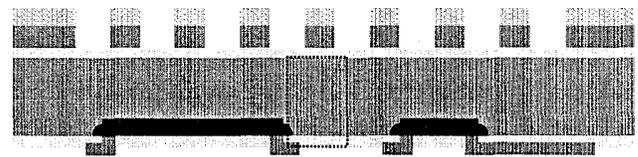
(c) DRIE etching channels and inlet/outlet holes (dashed rectangle)
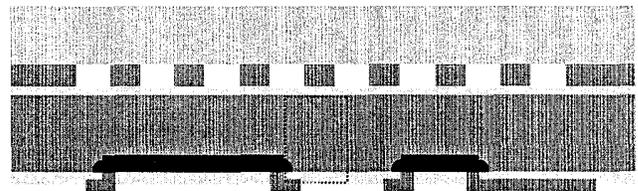
(d) Si-glass anodic bonding
Figure 3. Schematic of fabrication process.
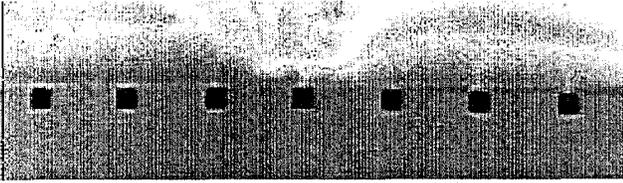
Figure 4. An SEM picture of the Si-glass bonding cross section.

EXPERIMENTAL SETUP

Fig. 5 shows the experimental setup. A fixture was developed to simplify the exchange of test devices. The fixture has internal flow directors and O-rings to securely seal the inlet and outlet holes in the heat exchanger. It also allows us to etch the inlet and outlet holes from silicon side, which is much easier than from glass side.
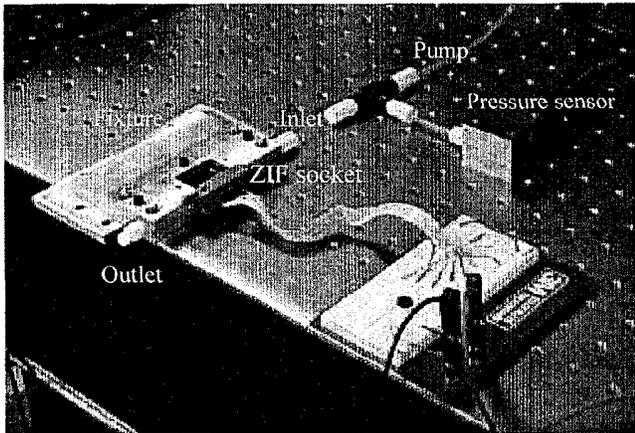
Figure 5. Experimental setup. The test chip is supported and sealed by the fixture with internal flow directors. The ZIF socket connects the heaters and thermometers to the breadboard, where input power is supplied and the thermometer resistances are measured. A pressure transducer is attached at the inlet. The outlet fluid temperature is measured by a thermocouple.

As shown in Figure 6, we attach a surfboard (the breadboarding medium for surface mount) to the resistor side of the heat exchanger, and wire-bond the aluminum pads to the surfboard, which has 10 standard single-in-line pins. We insert
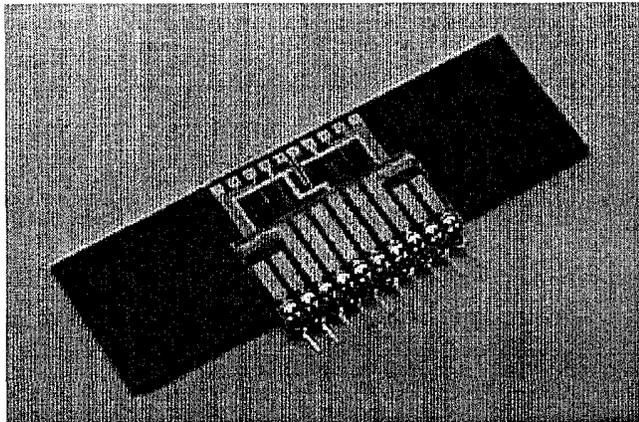
Figure 6 The wire-bonded surfboard and the heat exchanger.

the heat exchanger to the fixture, then clamp the ZIF (Zero-Insertion-Force) socket to the surfboard, and thermometer resistances can be measured from the bread board. A pressure sensor located right at the connection of the pump monitors the pressure drop. Since the exit of the fixture is open to the air, the pressure readout at the inlet can be considered the pressure drop along the microchannels. We also have a thermocouple at the outlet to measure the fluid temperature.

EXPERIMENTAL RESULTS

Six heat exchangers of three types listed in Table 1 were fabricated and tested in our experiments. DI water was used as cooling fluid. Because of the fabrication tolerance, the micromachined channels may not have the exact values as designed in Table 1. The measured actual dimensions of the testing devices are listed in Table 2.

Table 2. Actual dimensions of testing heat exchangers.

Device number	1	3	4	5	6	7
Type	1	3	1	3	2	2
Width (um)	75	100	80	100	55	55
Depth (um)	50	50	42	42	42	45
Length (cm)	2	2	2	2	2	2
Number of channels	50	30	50	30	20	20
Hydraulic diameter (um)	60	67	55	59	48	49

Thermometers on the same chip usually have similar but not identical resistance and thermal coefficients. We calibrated each of the thermometers on all the devices. The calibrations showed a very good repeatability before and after heat transfer experiments. The calibration curves of each thermometer were plotted and fitting equations were obtained.

$$y = 0.0000029669x^3 - 0.0216735619x^2 + 53.1292930645x - 43552.2477599307$$
$$R^2 = 0.9990074198$$
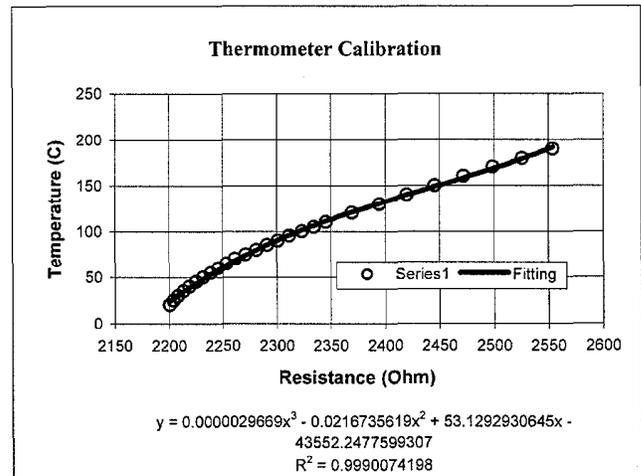
Figure 7. Calibration of a typical thermometer. We have seen resistance variations of 5-10% in the same wafer, so all thermometers were individually calibrated. Repeat calibrations on resistors after use indicated variations less than 0.2%.

The first phase of the experiments was single-phase measurements. Figure 8 shows the measured and simulated pressure drops in the microchannels for various flow rates. The simulation model in these data is a simple laminar flow pressure

drop relation with a constant friction factor (see Equation (1)). The pressure drop data agree reasonably well with the fully-developed laminar flow model, supporting the use of the analogous fully-developed laminar heat transfer model, in which the Nusselt number is constant and the convection coefficient is independent of the flow velocity. Because the heat flux into the fluid varies along the length of the channel, due to conduction in the silicon, the use of a constant convection coefficient is not precisely correct. However, the error due to this approximation is small compared to the uncertainty in the measured temperature data.
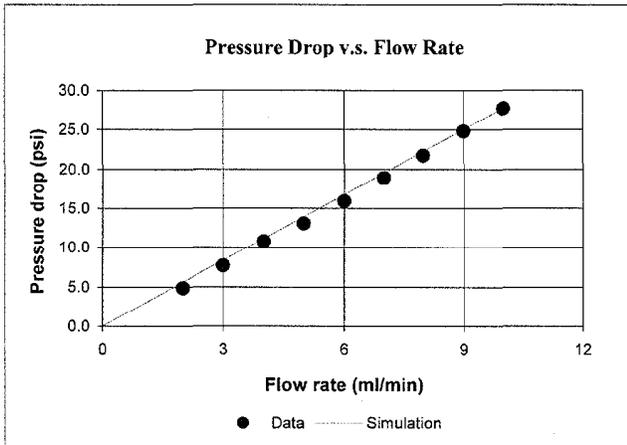
Figure 8. Pressure drop along the microchannels of device No. 4, a Type 1 device. This figure shows the measured and simulated pressure drop along the microchannels as a function of water flow rate at room temperature. A simple laminar flow model described in Equation (1) with constant friction factor accurately describes this data. Although pressure drop seems proportional to flow velocity squared in Equation (1), it is actually linearly proportional to flow velocity (or flow rate) because Reynolds number is also proportional to flow velocity.
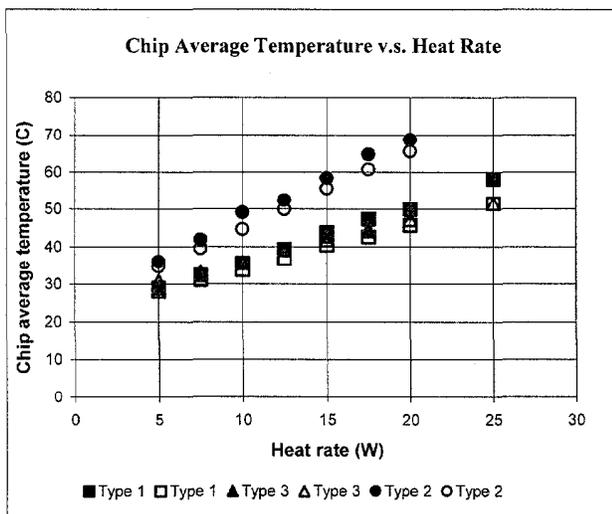
Figure 9. Chip average temperature as a function of heat rate for six devices of three types. As shown in Table 1, Type 1 and Type 3 have an effective cooling area of 4cm², with different size and number of microchannels, but similar entire convection area; Type 2 has an effective cooling area of 2cm². Type 1 and Type 3

devices were tested under 10 ml/min flow rate, while 5 ml/min for Type 2. Since Type 2 has a lower flow rate and less cooling area, the chip temperature rise is higher than Type 1 and Type 3 at the same heat rate. All trend lines extrapolate to the room temperature 22 C for zero heat input.

Next, we measured chip temperature as a function of heating rate for the three devices for the single-phase flow, as shown in Figure 9. In accordance with Equation (2) and (3), under a fixed mass flow rate, the average chip temperature linearly increases with the heat rate. 25 W of heat removal rate at 5-10 ml/min flow rate yields up to 50 C average chip temperature rise.

Figure 10 shows the average chip temperature as a function of water flow rate for a fixed heat rate of 25 W. Under single-phase flow, the chip temperature varies linearly with the inverse of the flow rate, according to Equations (2) and (3). This simple statement of the dependence neglects losses due to convection and radiation from the chip surface and into the fixture.
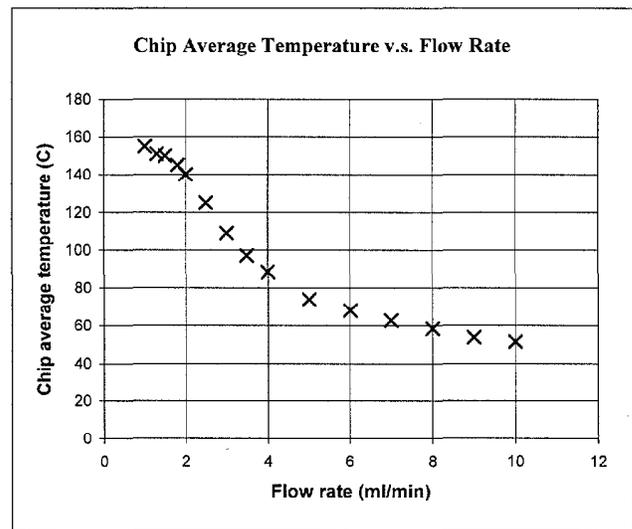
Figure 10. The Chip average temperature as a function of flow rate below 10 ml/min for a fixed heating power of 25 W. Device No. 4 of Type 1 was tested as flow rate was reduced from 10 ml/min to lower values, and boiling was observed at flow rates below 3 ml/min (as in Figure 10). After the onset of boiling, the chip temperature increases faster.

Figure 11 is an image of water boiling in microchannels with a hydraulic diameter of 55 μm. An interesting fact is, while the resistance readout of thermometers was quite stable in single-phase experiments, we observed resistance variations corresponding to 1-3 C after the onset of boiling. The formation of bubbles in these microchannels dramatically alters the fluid flow and heat transfer. The chip temperature change after the onset of boiling is plotted in Figure 10 together with single-phase flows. For this case the dependence of the chip temperature on the flow rate is increased, which indicates that the onset of boiling has substantially decreased the heat transfer coefficient in portions of the channels. Dry-out in the downstream portions of the channels is believed to be responsible for this decrease, which causes the temperature rise to increase substantially compared that for fully single-phase convection. This experiment did not resolve a diminished sensitivity of the temperature to the flow rate, which was expected at the onset of boiling before dry-out.
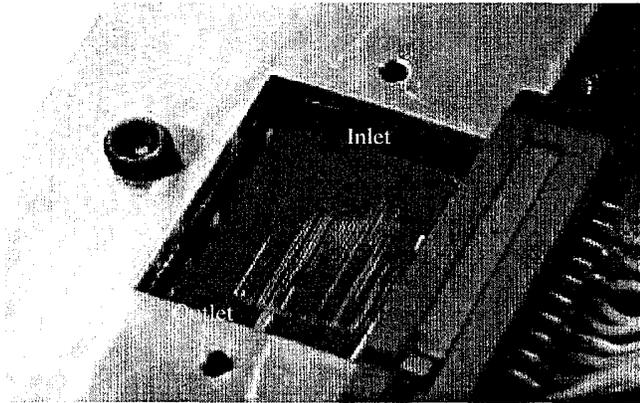
Figure 11. A close-up of water boiling in the microchannels with a hydraulic diameter of 55 μm. The shiny lines in the cooling area are where the bubbles are formed and move through the chip from the inlet to the outlet. Since there is a temperature distribution on the silicon substrate, the entire cooling area does not boil at the same time. Bubbles generate at the exit manifold of channels first, then the origin of the bubble line gradually moves to the entrance manifold in accordance with the local substrate temperature rise.

However, we still believe that the rate of chip temperature rise should be slowed down at the moment of boiling starts but before the dry-out occurs. Using the same device, we fixed the flow rate at 1 ml/min, carefully adjusted the heat input rate and monitored the four temperature readouts. The fixture was placed under a microscope in order to detect the bubble formation.
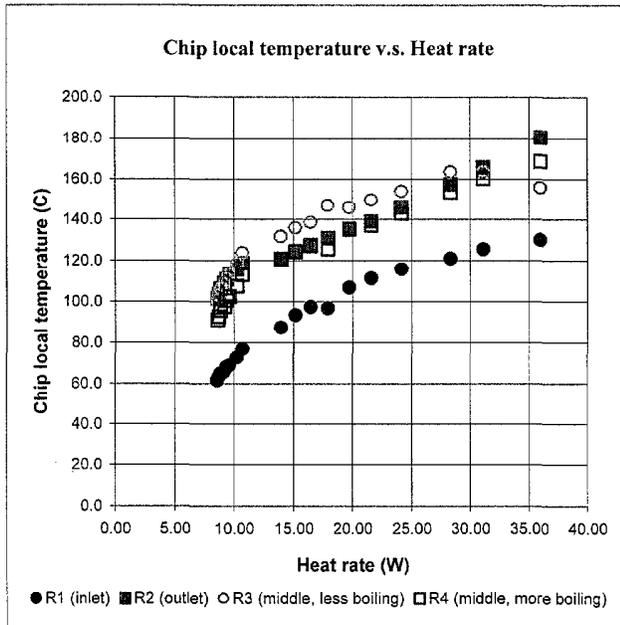
Figure 12. Boiling curve under a constant flow rate of 1ml/min. Device No. 4 was tested for phase change in the microchannels. The temperature readouts of the thermometers are no longer as stable as in the single-phase flow. The local temperature fluctuation is closely related to bubble generation.

Figure 12 plots the four temperature readouts as a function of heat removal rate. Local temperatures are plotted individually

because spatially averaged temperature may dampen any significant fluctuation caused by bubble generation. Each of the four curves shows a clear slope change after the onset of boiling. The increase rate of chip temperature is slowed down by a factor of 5-10 under two-phase flow regime. We also noticed a local temperature drop of about 5 C at the outlet when the first bubble formed in one of the microchannels.

Because of the density difference between liquid and gas phase and the associated bulk accelerations of the bulk flow, a significant change in pressure drop after boiling begins can be expected. Figure 13 is the pressure readout at the inlet when the flow rate is fixed at 1ml/min and the heat rate increases from 0 to 35 W. In the single-phase flow regime, the pressure drop decreases corresponding to the decrease of water viscosity at higher temperatures. However, after the onset of boiling, the pressure drop increases dramatically due to rapid vaporization.
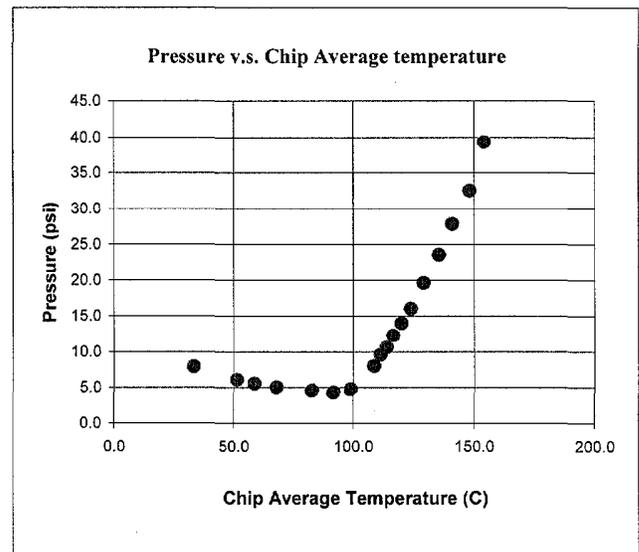
Figure 13. Inlet pressure change from single-phase to two-phase flow. Testing condition is the same as in Figure 12. The pressure drop as well as the flow conditions along the microchannels can vary dramatically because of the property difference between water and vapor. At a fixed 1 ml/min flow rate, the Reynolds number in the liquid phase is as low as 5.8, while the flow velocity is 0.1m/s. If the fully developed boiling state is reached (with the quality equal to 1), the vapor velocity in the microchannels can be as high as 156m/s.

DISCUSSION

As shown in Figure 11, the boiling is not uniform across the channels. This may be caused by a non-uniform heating or non-uniform flow distribution. The non-uniform heating may be caused by the breaking down of PN junctions and by conduction in the silicon. Our experiments show that the heaters require a long annealing and/or a low background doping level in order to work reliably under high power input. The reason is that the resistors are isolated by PN junctions that can break down at a certain voltage threshold. Both a long annealing and a lightly doped substrate help to increase this break-down threshold. On the other hand, however, this also indicates that more work is

279

needed to establish stable two-phase, sub-critical heat flux convection.

A large group of microchannels with a large cooling area is probably best for an actual microchannel heat exchanger application, but not for certain heat transfer experiments. As a general guide, the total convective thermal resistance can be significantly reduced by increasing the number of microchannels, while the large cooling area with distributed heaters makes it more difficult to maintain a uniform heating condition. In addition, it is also very difficult to evenly distribute the fluid into the individual microchannels. We have observed a roughly parabolic flow profile across the microchannels with our current inlet manifold design, which gives a higher flow rate to the middle channels. As an example, a single channel or a small group of microchannels is a better design for critical heat flux study.

The manifold design can also influence the boiling condition in the microchannels. In this design, the two "T" shaped channels are used as fluid inlet and outlet. Since the manifold has a much larger surface area than the microchannels in the effective cooling area, the bubbles tend to form in the manifold. However, boiling in the manifold can be reduced by increasing the thermal resistance of the manifold.

We also notice the temperature increase of the fixture during the experiments, indicating a heat loss from the fixture. The current fixture is made of aluminum, which has a very high thermal conductivity. To reduce the heat loss from the fixture, we can choose materials with low thermal conductivity, reduce the length of the internal flow channels, or place thermal isolations in the device itself—for example, cutting trenches in the silicon chip.

CONCLUSIONS

This work shows that micromachined thermometers and heaters can aid with the study of single and two-phase convection in microchannels. With our current microchannel test systems, we have successfully observed phase change in microchannels as well as conducted detailed flow and heat transfer experiments. With DRIE technique, we are able to control the aspect ratio of the microchannels. These implanted resistors provide a reliable method to control the heat input rate as well as to measure the local temperature change. We measured up to 50 C average chip temperature rise, under DI water single-phase convection at 5-10 ml/min flow rate at 25 W heat removal rate. We have also observed variations in the resistances of the thermometers as boiling begins in these microchannels. Temperature variations of up to 3 C at frequencies of 1 Hz and faster have been observed, indicating that dynamic thermal measurements are possible in these structures.

Ongoing research is studying nucleation in microchannels and the relationship with channel wall roughness and the critical heat flux. Through these experiments, we aim to construct a more accurate model for the design of stable two-phase microchannel heat exchangers for integrated circuits.

ACKNOWLEGMENTS

REFERENCES

1. D. B. Tuckerman, and R. F. W. Pease, "High-Performance Heat Sinking for VLSI", *IEEE Electron Device Letters*, Vol. EDL2, No. 5, pp126 (1981).

2. V.K. Samalam, "Convective Heat Transfer in Microchannels", *Journal of Electronic Materials*, Vol. 18, No. 5, pp611 (1989).

3. R. W. Knight, D. J. Hall, J. S. Goodling, and R. C. Jaeger, "Heat Sink Optimization with Application to Microchannels", *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, Vol. 15, No. 5, pp832 (1992).

4. X. F. Peng, H. Y. Hu, and B. X. Wang, "Flow Boiling through V-Shaped Microchannels", *Experimental Heat Transfer*, Vol. 11, pp87 (1998).

5. L. Jiang, and M. Wong, "Phase Change in Microchannel Heat Sinks with Integrated Temperature Sensors", *Journal of MEMS*, Vol. 8, No. 4, pp358 (1999).

6. F. M. White, "Fluid Mechanics (4th Edition)", McGraw-Hill Companies, Inc., ISBN 0-07-069716-7, (1999).

HIGH SENSITIVITY MEMS ULTRASOUND ARRAYS BY LATERAL FERROELECTRIC POLARIZATION

J.J. Bernstein , J. Bottari, K. Houston, G. Kirkos, and R. Miller
The Charles Stark Draper Laboratory,
Cambridge MA 02139-3563

B. Xu, Y. Ye and L.E. Cross
Materials Research Lab, Penn State University,
University Park, PA 16802

ABSTRACT

This paper discusses the design of advanced micromachined ferroelectric ultrasound transducers for use at 3 MHz. 16 X 16 arrays of resonant monomorph sensors have been constructed, with sol-gel PZT as the active ferroelectric layer deposited on insulating layers of ZrO_2 and SiO_2. A novel in-plane polarization of the PZT is used to maximize sensitivity, while trading off reduced output capacitance to match the CMOS buffer electronics. This results in about 30 dB improved sensitivity compared to conventional polarizing across the thickness of the PZT layer. Fluid-filled through wafer holes are used as an acoustic matching network to achieve resonance at both 1 and 3 MHz. A lumped element equivalent circuit model is presented, as well as Finite Element Analysis results. Performance predictions for projector efficiency and receive response are given. Test results are presented including transmit response, receive sensitivity, and frequency response.

INTRODUCTION

There is great interest in two-dimensional arrays of acoustic transducers for 3-D acoustic imaging. Previous work reported the use of micromachining technology combined with sol-gel PZT (Lead Zirconate Titanate) to create 2-dimensional arrays of MHz acoustic transducers on silicon wafers [1]. Micromachined Ultrasonic Transducers (MUT's) based on electrostatic transducers have been developed for both air and underwater applications [2]. Miniature ultrasonic transducers have been fabricated on silicon using films of PVDF (Poly-vinylidene-difluoride) as the transduction material [3,4].

In this paper we describe a novel in-plane polarization technique for increasing the sensitivity of ferroelectric monomorphs by 30 dB or more. Potential applications of these transducer arrays include a hand-held diver's sonar, medical ultrasound imaging, and nondestructive testing (NDT).

THICKNESS MODE VS. IN-PLANE POLARIZATION

Ferroelectric films on silicon are usually grown on an inert conducting layer such as Pt which forms one of the device electrodes, with the other electrode on the upper surface. Most work with PZT on silicon has been confined to thin films less than 1 μm thick of PZT, although films as thick as 12 μm can be deposited [5]. For a uniaxial stress σ_{xx} in a PZT film of thickness t_{PZT} (Fig. 2) the output voltage is

$$V = \sigma_{xx} \cdot t_{PZT} \cdot g_{31} \qquad (1)$$

where g_{31} is the receive coefficient of the PZT for a stress applied orthogonal to the poling direction.

Fig. 2 shows the normal thickness-mode polarization of a piezoelectric or ferroelectric film. For a ferroelectric film, the possibility exists to pole the film in the plane of the substrate (Fig. 3), so that much larger inter-electrode l spacing is possible. The output voltage for this case is

$$V = \sigma_{xx} \cdot l \cdot g_{33} \qquad (2)$$
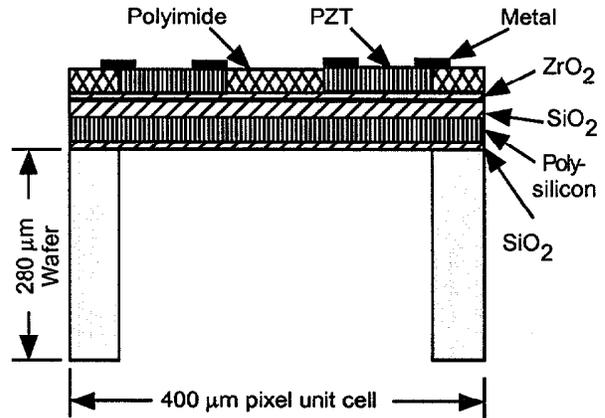
Fig. 1. Cross-section of one sensing pixel showing fluid-filled hole terminating on sensing diaphragm.
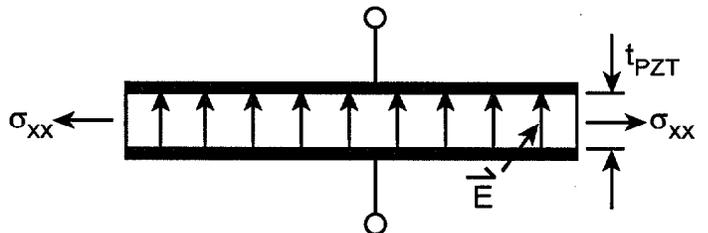
Fig. 2. Thickness polarization of ferroelectric film.
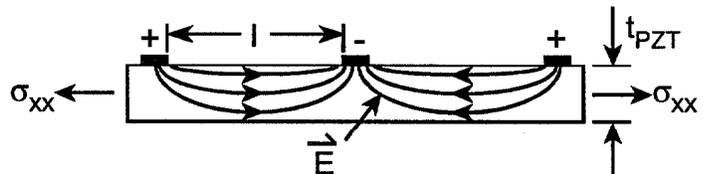
Fig. 3. Lateral polarization of ferroelectric film.

Since g_{33} is normally twice g_{31}, a 6 dB increase in coupling coefficient results from using in-plane polarization. If we assume a 1 μm thick PZT film with 30 micron electrode spacing, then one obtains a factor of 30 increase in voltage sensitivity from electrode spacing, and a factor of 2 from higher ferroelectric coupling constant, for a total of 60X higher sensitivity (+35.5 dB). An additional benefit is that only a single conductive layer is required, which greatly simplifies the fabrication process. This increase in sensitivity is

obtained at the cost of reduced output capacitance, but an overall increase in signal/noise is achieved.

Poling Procedure: The PZT films are poled using the transducer's electrodes at 10 volts/μm of inter-electrode gap [6]. Over-voltage of 20 V/μm causes the PZT film to crack. 16 X 16 arrays were poled one row at a time using a probe card with 16 high voltage electrodes and 1 ground electrode. The poling voltage was applied slowly using a ramp up, soak, and ramp down of 2 minutes each.

Sensitivity-Capacitance Tradeoff

For simple ferroelectric pressure transducers, the quantity $M_0^2 C$ (where M_0 is the open-circuit sensitivity (V/Pa) and C is device capacitance) is characteristic of the transducer and is not affected by partitioning the transducer into series or parallel connected pieces, attaching an ideal transformer, or reconfiguring the electrodes. Hence capacitance can be traded off for sensitivity. The optimal signal to noise ratio for a capacitive sensor is achieved when the sensor capacitance is approximately equal to the input capacitance of the buffer amplifier.

The sonar pixels were designed to have an output capacitance of between 0.2 and 0.6 pF, to match the input capacitance of the CMOS readout circuit, which was 0.25 pF.

ANALYTICAL MODEL

A lumped element acoustic model (Fig. 4) of the transducer was analyzed to provide insight into the design and optimization of the transducer. The model contains an acoustic source in the far-field, acoustic radiation elements L_R and R_R representing radiation mass and resistance [7], a transmission line *TL* which represents wave propagation down the fluid-filled holes, and lumped impedances C_{dia} and L_{dia} which represent the compliance and mass of the air-backed diaphragm.

This lumped element model can be solved analytically, allowing optimization to be carried out. In addition the model provides insights which are absent from the finite element results.

Maximum sensitivity occurs at the system resonances, when the reactances cancel leaving only the radiation resistance.

The thickness of the wafer was chosen to provide resonance at both 1 MHz and 3 MHz, the desired operating frequencies of an underwater sonar. This calculated to a wafer thickness of approximately 295 μm which was used for fabrication.
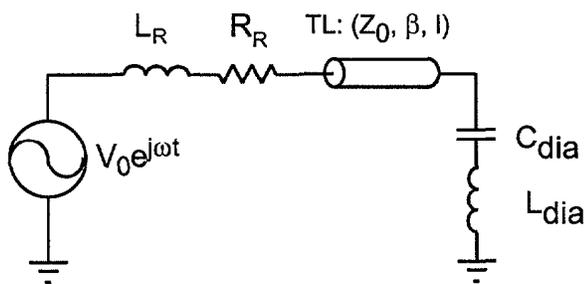
Figure 4. *Lumped element equivalent circuit model of monomorph pixel on a fluid-filled hole.*

Fig. 5 shows the calculated average diaphragm deflection (per Pa input) vs. frequency for 3 wafer thicknesses which give the first, second or third resonance at 3 MHz. To achieve the first resonance at 3 MHz required a wafer thickness of only 43 μm, which is too thin to handle. To achieve the third resonance at 3

MHz required a wafer 540 μm thick, but would not yield a peak near 1 MHz as desired. A 4th (diagonal) line is added for reference which shows the free particle displacement of a 1 Pa acoustic wave in water.
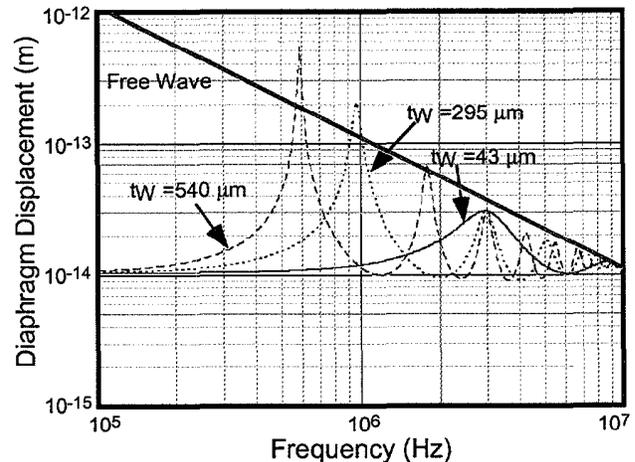
Figure 5. *Average diaphragm deflection vs. frequency for sensor diaphragm mounted on hole with three different lengths to achieve a resonance at 3 MHz.*

FINITE ELEMENT ANALYSIS

Finite element analysis was carried out on 13 different final designs. All pixels measure 0.4 mm X 0.4 mm. The designs are grouped into 3 general types: Quad designs (Fig. 6) with 4 diaphragms per unit cell, "racetrack" designs with 2 or 3 diaphragms per unit cell (Fig. 7), and Nona designs with 9 diaphragms / unit cell.

Fig. 6 shows a quad cell, with 4 diaphragms wired in series. The center metal bond pad is for bump bonding to a CMOS "Transmit-Receive Integrated Circuit" developed by Lockheed Martin IRIS.

ANSYS was used for all FEA computations. Fig. 8 shows the device as simulated by ANSYS. The diaphragm model contains the details of all layers, and was electrically driven by 1 V.

The PZT response to the driving voltage deforms the diaphragm and generates an outgoing acoustic wave. The semi-infinite fluid filled space is represented by a hemisphere of fluid elements (the "mushroom cap" of Fig. 8). The outer boundary of this hemisphere is made up of "infinite absorbing elements" which prevent reflections back towards the transducer. .
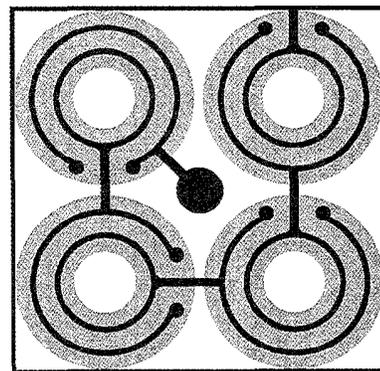
Fig. 6. *Quad transducer has 4 diaphragms wired in series per unit cell (0.4 mm X 0.4 mm). Concentric electrodes and PZT layer are shown.*
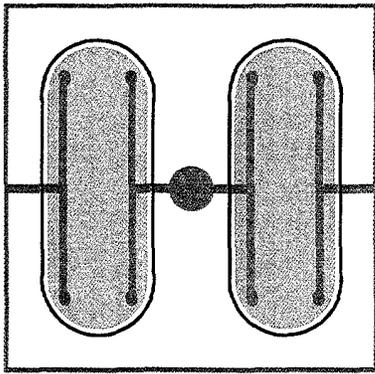
Fig. 7. *"Race-track" transducer has two elongated diaphragms wired in parallel per unit cell.*
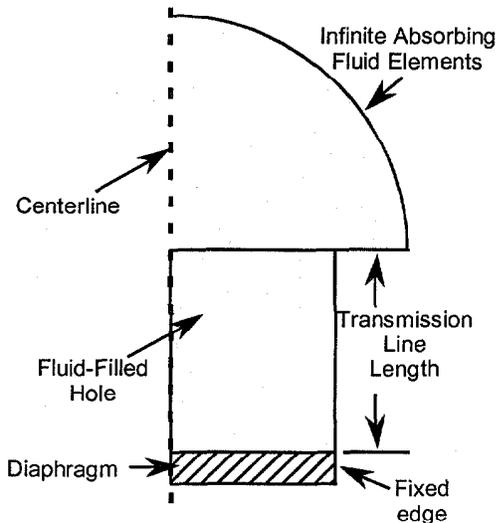
Fig. 8. *FEA problem as solved with ANSYS for radially symmetric diaphragm (centerline shown) with diaphragm fixed at outer edge.*

The transmit response for the transducer at 1 meter was calculated assuming conservation of energy in the acoustic wave and expressed in standard sonar units as dB ref. 1μPa/V at one meter.

Two approaches were used to predict receive sensitivity: static calculations augmented by the Q of the transducer, and reciprocity relations which allow the receive sensitivity to be directly calculated from the transmit response for any reciprocal transducer.

Predicted Transmit and Receive Sensitivity

The FEA results were translated into standard sonar units for transmit and receive. These are listed for several transducer designs in Table 1 below. Receive sensitivity was calculated from transmit response using reciprocity relations.

DEVICE FABRICATION

Sonar transducer arrays were fabricated on 100 mm diameter silicon wafers 290 μm thick. After initial oxidation, a 3 μm thick layer of undoped polysilicon was deposited by Low Pressure Chemical Vapor Deposition at 570°C from silane, followed by deposition of 2 μm of SiO_2 . The wafers were then coated with 0.5 μm of ZrO_2 and then 3 μm of PZT using sol-gel techniques

[6]. The PZT layer was patterned using conventional photolithography and wet etching.

Table 1. Predicted Transmit and Receive Responses

Transducer	Unit Cell	Transmit resp.	Receive
Name	Cap.	1 Unit Cell	Sensitivity
Units -->	(F)	dB//μPa/V @ 1m	dB//V/μPa
quad80	2.21E-13	114.1	-209.0
nona50	6.64E-13	120.4	-213.5
RT120	3.42E-13	114.1	-211.2

A 3 μm polyimide dielectric layer is then spun on and cured. This low dielectric constant layer reduces the stray capacitance to the substrate of all metal lines which are not in contact with the PZT. A metal layer consisting of Ti/Pt/Au is then sputtered and patterned by a standard liftoff technique. The wafers are then coated with a top-coat of polyimide and patterned to expose the bondpads only.

Holes are then etched from the back of the chip using an STS Inductively Coupled Plasma etcher, masked with a thick photoresist layer. This etcher allows vertical holes to be etched through the wafer.

"Mono" devices with a single diaphragm per pixel are shown in Fig. 9. This device had an output capacitance of 0.7 pF and sensitivity (-210 dB //1V/μPa).

Fig. 10 shows a cross-section of an array which has been cracked to reveal the hole side-walls. The etch walls are close to vertical.
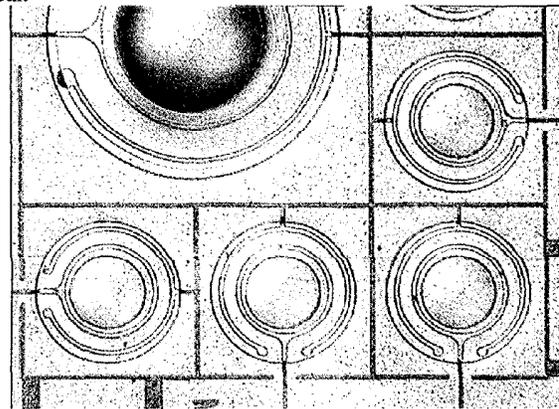
Fig. 9. *Circular "Mono" pixels are poled radially with concentric ring electrodes.*

TEST RESULTS

16 X 16 arrays were tested at Lockheed Martin by bump-bonding an electronics readout chip to the array. In-house testing was carried out by mounting small 3 X 3 arrays into Kovar flat-packs with holes cut in the package to allow water and acoustic signals to reach the chip. The chips were wire-bonded to the flat-pack, which was soldered to a PC board, to which coax cables were attached. The entire assembly was then coated with silicone rubber for water-proofing, with a hole to allow sound to reach the sensor array.

Fig. 11 shows receive voltage sensitivity for a Mono-150 device (single diaphragm, radius = 150 μm). Peak response of –210 dB ref. V/μPa is at 3.8 MHz, (slightly higher than the desired system frequency), with an additional peak at 1.5 MHz. Noise floor of the system is also shown. This data was compensated for 190 pF of stray cable capacitance and pre-amp input impedance on each pixel.

Fig. 10. *Through-wafer hole terminates on PZT monomorph diaphragm.*

Fig. 12 shows transmit response from a Mono-150 design pixel. The peak transmit response (at 3.8 MHz) is about 122 dB// µPa/V at 1 meter. A peak at 1.5 MHz is also visible.

Fig. 11. *Receive data for Mono 150 design pixel. Notice broad peak at 3.8 MHz.*

Fig. 12. *Transmit Voltage Response (TVR) and noise floor, MONO 150 pixel on a test chip.*

SUMMARY AND CONCLUSIONS

This paper has described a new type of MEMS ferroelectric transducer, the monomorph with in-plane polarization. This device allows sensitivity to be traded off against capacitance to optimize the system signal/noise ratio, particularly with low input-capacitance CMOS amplifiers. A novel dielectric buffer layer (ZrO_2) compatible with PZT sol-gel deposition was developed. 16 X 16 arrays of MEMS sonar pixels were successfully built and tested. Sensitivity of –210 dB ref. 1V/µPa at 3 MHz was attained, which is an improvement of about 30 dB over previous MEMS ferroelectric sensors at a lower frequency (0.8 MHz) [1]. The devices are reciprocal and can be used for both transmit and receive.

ACKNOWLEDGEMENTS

REFERENCES

1. J. Bernstein, S.L. Finberg, K. Houston, L.C. Niles, H.D. Chen, L.E. Cross, K.K. Li, and K. Udayakumar., "Micromachined High Frequency Ferroelectric Sonar Transducers", *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, Vol. 44, No. 5, September 1997, pp. 960-969.
2. X. Jin, I. Ladabaum, and B.T. Khuri-Yakub, "The Microfabrication of Capacitive Ultrasonic Transducers", *IEEE Journal of Microelectromechanical Systems*, Vol. 7, # 3, September 1998, pp. 295-302.
3. R.G. Swartz, Application of Polyvinylidene Fluoride to Monolithic Silicon PVF2 Transducer Arrays, Ph.D. Dissertation, Stanford University, (May 1979).
4. J.H. Mo, J.B. Fowlkes, A.L. Robinson, and P.L. Carson, "Crosstalk Reduction with a Micromachined Diaphragm Structure for Integrated Ultrasound Transducer Arrays," *IEEE Trans. Ultrasonics, Ferroelectrics and Frequency Control*, Vol. 38, (January 1992), pp. 48-53.
5. H.D. Chen, K.R. Udayakumar, C.J. Gaskey, L.E. Cross, J.J. Bernstein, and L.C. Niles, "Fabrication and electrical properties of lead zirconate titanate thick films", *Journal of the American Ceramic Society*, vol. 79, no.8, pp. 2189-92, 1996.
6. B. Xu, Y. Ye, L.E. Cross, J. Bernstein and R. Miller, "Dielectric hysteresis from transverse electric fields in lead zirconate titanate thin films", *Applied Physics Letters*, vol. 74, # 23, pp. 3549-3551, (7 June 1999).
7. Kinsler, Frey, Coppens and Sanders, Fundamentals of Acoustics, 3rd edition, (John Wiley and Sons, 1982), pp. 191-193.
8. Leo L. Beranek, Acoustical Measurements, 1988 Edition, (Acoustical Society of America / American Institute of Physics, 1988), pp. 113-122.
9. R. J. Bobber, Underwater Electro Acoustic Measurements, (Peninsula Press, 1988) pp. 27-45.
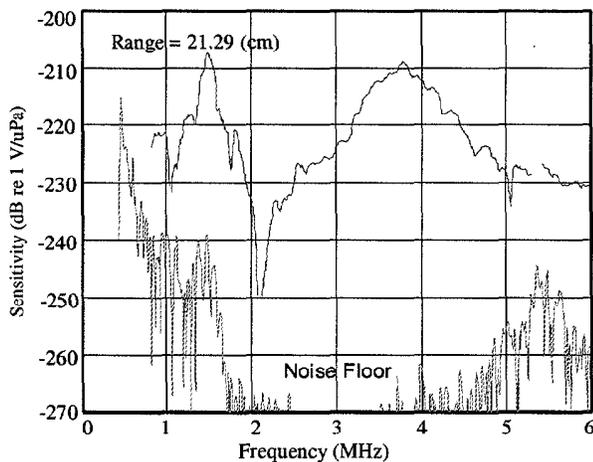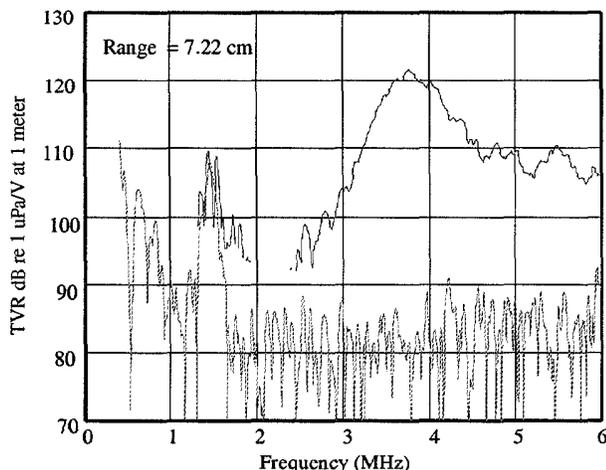
A High-Performance Hybrid CMOS Microaccelerometer

Arvind Salian, Haluk Kulah, Navid Yazdi[†], Guohong He and Khalil Najafi

Center for Integrated Microsystems
1301 Beal Avenue, University of Michigan, Ann Arbor, MI 48109-2122
email: najafi@umich.edu: Tel: (734) 763 6650

[†]Now with: *Center for Solid-State Research*, Arizona State University, Tempe, AZ 85287-5706

ABSTRACT

A hybrid microaccelerometer subsystem consisting of an all-silicon μg capacitive microaccelerometer and a CMOS capacitive interface circuit is presented. The measured sensitivity for a device with 2mm x 1mm proof mass and 1.4μm air gap is 1.4pF/g and the calculated mechanical noise floor for the device is 0.39μg/√Hz in atmosphere. The circuit has a 95dB dynamic range, a low offset of 370μV and can resolve better than 75aF. The complete module has a measured sensitivity of 160mV/g and noise floor of 3.6μV/√Hz (-110dBV/√Hz), indicating that the current system is capable of resolving about 20μg/√Hz.

Keywords: Inertial sensors, μg accelerometer, interface circuit.

INTRODUCTION

High-precision microaccelerometer subsystems are needed in navigation/guidance, microgravity measurements, tilt control, platform stabilization, and position sensing [1,2]. The sensor should have high sensitivity, low temperature sensitivity, and good long-term stability. The interface circuit should have good dc response, low offset, and high offset and gain stability. Surface micromachined accelerometers attain high capacitive readout sensitivity by integrating the sensor and its interface circuitry [2]. However, these sensors typically have a noise floor of several 100μg's because of their small proof mass and small full-scale capacitance change. Bulk micromachined accelerometers have higher sensitivity and hence can attain sub-10μg/√Hz performance [3]. In this paper, we report the implementation and testing of a hybrid capacitive silicon microaccelerometer subsystem with a CMOS readout interface for high-precision applications mentioned above.

μg ACCELEROMETER STRUCTURE

A precision μg microaccelerometer is typically operated in closed-loop to obtain higher bandwidth, full-scale range, and linearity. For a high performance accelerometer, the main design consideration for the device structure is meeting the resolution specification which is determined by its mechanical and readout electronics noise. Mechanical noise can be reduced by increasing the size of the proofmass and reducing damping. Electrical signal-noise ratio can be improved by reducing electronic noise, and increasing proofmass size and device capacitance variation. The latter two provide a larger signal pick-up and result in less stringent readout circuit requirements. Mechanical noise is mainly due to Brownian motion and can be expressed in terms of mass size and total damping [4]. The key element in obtaining low mechanical noise floor without vacuum packaging is increasing mass size and controlling damping [3,4].
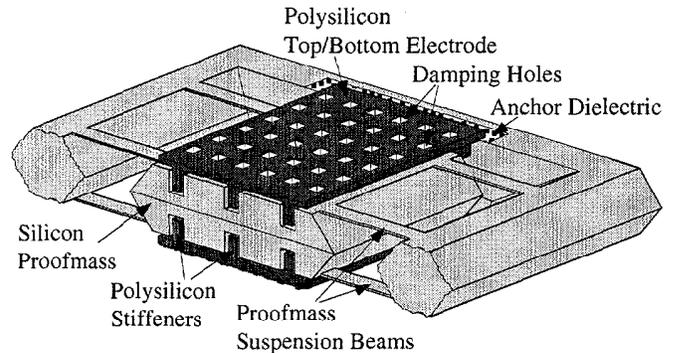
Figure 1. Microaccelerometer structure [3].

Our approach, as shown in Figure 1, addresses all these design issues by combining surface and bulk micromachining in order to achieve high device sensitivity, low noise floor, and controllable damping - all by using a single silicon wafer [3]. The central idea behind the process is to use the whole wafer thickness to attain a large proofmass, to utilize a sacrificial thin film to form a uniform and conformal gap over a large area, and to create electrodes by depositing polysilicon on the wafer. These electrodes, while thin, are made very stiff by embedding thick vertical stiffeners in them so that force rebalancing of the proofmass becomes possible [5]. The technology utilizes a trench refill technique to form thick stiffeners by depositing thin polysilicon films. Any damping hole configuration and geometry can be easily formed on the poly electrodes to optimize damping coefficient and capacitance, while there is no concern with large hole density and its effect on the electrode softening as the plate stiffness is mainly provided by the embedded stiffeners. There are eight boron-doped suspension beams which are symmetric with respect to the proofmass centerline and result in low cross-axis sensitivity. The electrodes are polysilicon plates created on both sides of the proofmass and anchored on an isolation dielectric at the frame.

FABRICATION PROCESS / RESULTS

The fabrication process, as shown in Figure 2, requires six masks and provides a yield of >80%. The process starts with a shallow (3μm) p++ boron diffusion on <100> double-polished p-type Si wafers. Both sides of the wafer are patterned and the patterns are aligned to each other. The shallow boron diffusion is performed at 1175°C for 30min and defines the beams, the

proofmass and the supporting rim. Then, 5000Å of LPCVD oxide is deposited and patterned to form dimples. These dimples reduce the contact area and help reduce stiction. They also limit the range of travel and help against shock. The oxide also serves as a mask for the trench etch step. Next 40-60µm deep, 6µm wide trenches are etched in the silicon to be used later to form the vertical electrode stiffeners.

(a) Shallow boron diffusion

(b) Deposit oxide (sets the dimple height), etch trench for vertical electrode stiffner

(c) Trench refill using sacrificial oxide, nitride and polysilicon

(d) Dope poly, etch poly, deposit capping oxide, etch oxide for EDP etch, open contacts, deposit metal

(e) Release in EDP, etch sacrificial oxide in HF
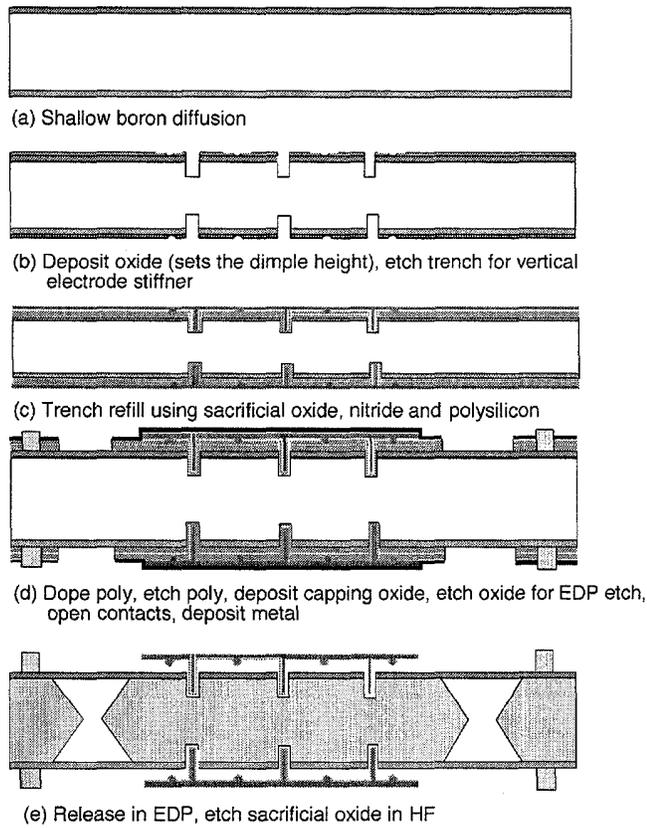
Figure 2. *Fabrication process sequence of a µg accelerometer.*
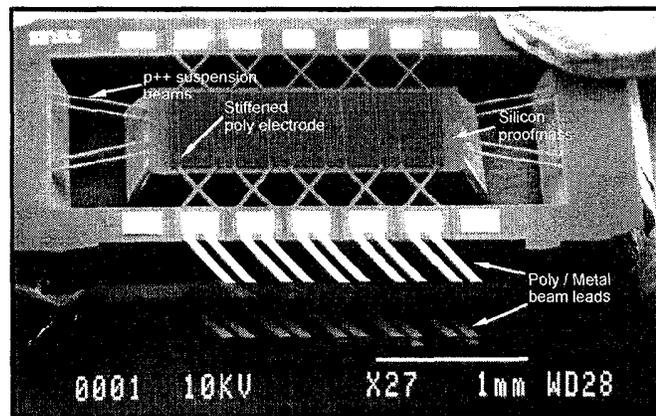
Figure 3. *SEM view of a high-precision µg accelerometer with 2mm x 1mm proofmass.*

The trenches are then refilled completely with sacrificial LPCVD oxide, LPCVD nitride and low-stress LPCVD polysilicon. The polysilicon is doped with boron at 1050°C and then annealed at 1050°C. The polysilicon is patterned to form the electrode

plates with damping holes. Next 4000-5000Å LPCVD (capping) oxide is deposited. The oxide is patterned to form metal contacts and openings to bulk Si for the subsequent proofmass release. Then Cr/Au (300Å/5000Å) is evaporated and lifted-off. The proofmass is released in EDP at 110°C. The sacrificial oxide is etched in straight HF. The released structures are soaked in methanol overnight and then oven dried at 110°C.

Figure 3 shows a SEM view of the device with 2mm x 1mm proofmass. The device has 5 electrically isolated electrodes on each side, which are anchored at the rim using stiffened poly supports. Taking advantage of the selectivity of the planes in EDP, the proof mass corners have been well compensated as seen in Figure 3. Table 1 summarizes the device specifications.

Table 1. *Silicon Sensor Characteristics*

Sensor Parameters	Value
Proof Mass Size	2000 x 1000 x 450 $(µm)^3$
Rest Capacitance (measured)	38pF (5 Electrodes)
Sensitivity (calculated)	1.25pF/g (Bridge)
Sensitivity (measured)	1.4pF/g (Bridge)
Resonant Frequency	770Hz
Mechanical Noise Floor (calculated)	0.39µg/√Hz

INTERFACE CIRCUIT

The microaccelerometer is interfaced with capacitive readout circuitry to detect capacitance change and to operate the sensor in open-loop or force-rebalance the proof mass in closed-loop. Figure 4 shows the block diagram of the interface circuit [6]. Two fixed reference capacitors are used to form a full-bridge with the sensor capacitive half-bridge, and the sensor top and bottom electrodes are used as the input nodes to the chip front-end.
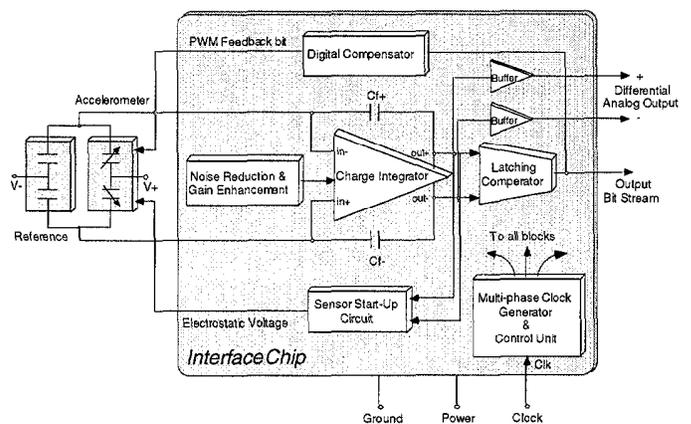
Figure 4. *Block diagram of the CMOS capacitive interface chip. The circuit is capable of both open-loop and closed-loop operation and utilizes a switched-capacitor readout technique to reduce sensitivity to parasitic capacitances [6].*

Fig. 5 shows the photograph of the circuit in 1.2µm CMOS. It has a fully-differential charge integrator with correlated double sampling to cancel 1/f noise, amplifier offset and compensate finite

286

amplifier gain. It utilizes digital compensation, chopping, and safe start-up to ensure high performance readout operation. It has 95dB dynamic range, a low offset of 370μV and can resolve better than 75aF/√Hz. The interface circuit has an adjustable gain in the 0.3-1.1 V/pF range. Table 2 summarizes the circuit specifications.
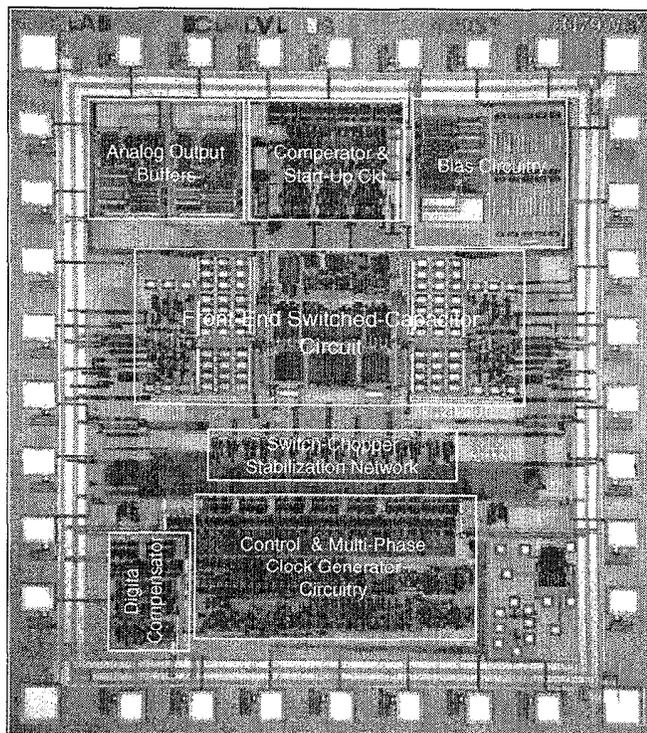
Figure. 5. CMOS capacitive readout interface chip [6].

Reference Capacitors   *Interface VLSI Chip measuring 2.6mm x 2.4mm*

Printed circuit board measuring 12mm x 16mm   *Silicon Accelerometer with 4mm x 1mm proof mass*
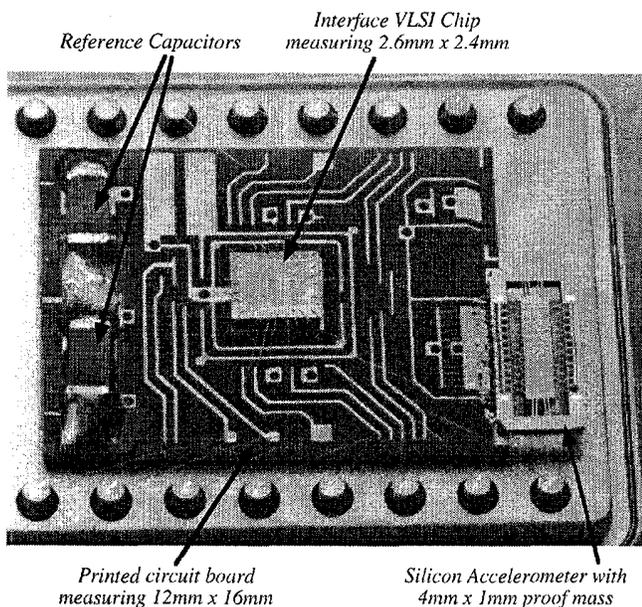
Figure 6. *Hybrid packaged accelerometer and the interface chip in a standard 24-pin IC package.*

HYBRID SUBSYSTEM-TEST & MEASUREMENTS

Figure 6 shows the hybrid subsystem with the sensor and the circuit assembled onto a PC board and mounted inside a standard DIP package. The sensor and the interface circuit are packaged close to each other to minimize parasitics. The device is mounted on the edge of the PCB to reduce any packaging induced stress.

The device (2mmx1mm) has a measured open loop sensitivity of 1.4pF/g. Figures 7 through 9 show the open-loop test results for this system. The open-loop tests were done on a dividing head, in a 1g gravitational field, by changing the acceleration on the sensor from -1g to +1g. While changing the applied acceleration, both analog output signals were measured. Figure 7 shows the results obtained from this measurement. As evident, both signals are sinusoidal and vary from 0 to 0.16 volts. The deviation from an ideal sinusoid is an artifact of the measurement setup and can be eliminated by taking a differential measurement, as shown in Figure 8.
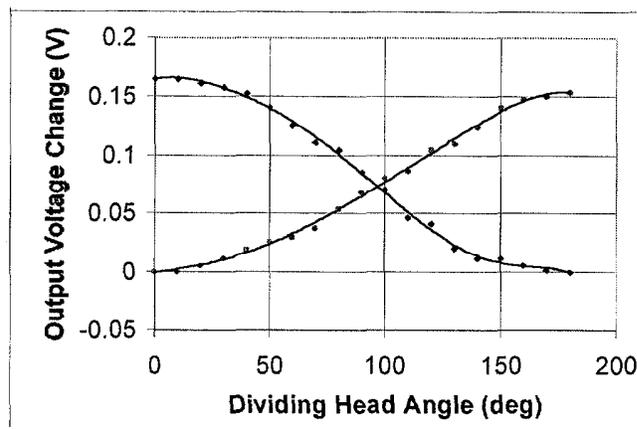
Figure 7. Output Voltages of the sensor system for ±1g range.
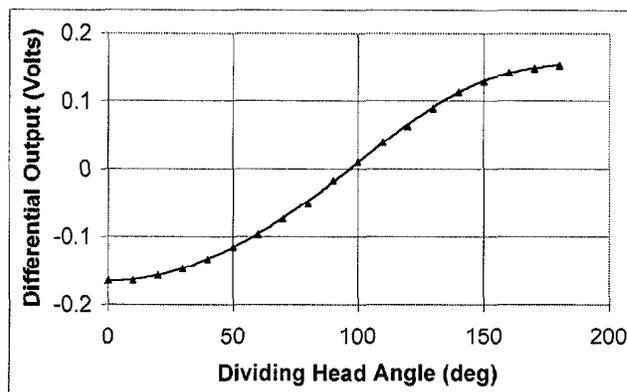
Figure 8. *Differential output voltage of the sensor system for ±1g range.*

Figure 9 shows the open-loop test results for this system with respect to acceleration indicating a sensitivity of 160mV/g. The expected sensitivity of the system (untrimmed) is ~420mV/g. The noise floor is measured using a dynamic signal analyzer HP3561. Figure 10 shows the measured output noise spectrum of the complete module with a dc input of 1g, which shows a noise floor of 3.6μV/√Hz. This indicates that the current system is capable of resolving about 20μg/√Hz. The sensitivity is lower than expected

(<3x). The reasons for this are not clear yet and are being investigated.

However, it should be noted that the circuit operates at its lower gain mode, i.e. ~300mV/pF, and hence the overall sensitivity can be increased by laser trimming the integration capacitances in the circuit. Moreover, the overall resolution can be further reduced down to sub-10μg/√Hz by using an accelerometer with higher sensitivity. For example, it can be decreased lower than 1.9μg/√Hz by using a sensor with 39pF/g sensitivity. Therefore the results obtained through the test of this hybrid-system are promising for very high resolution and higher sensitivity accelerometer systems. Temperature and shock tests have not yet been performed on these devices.
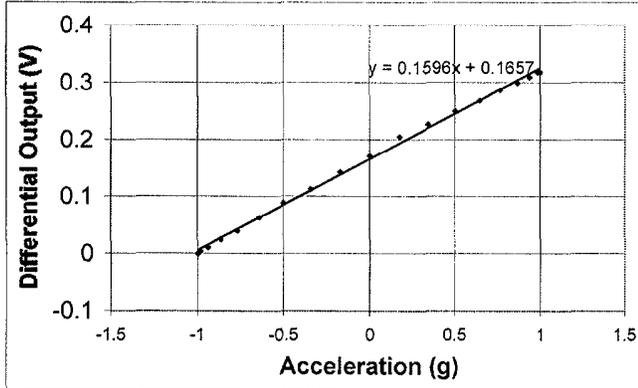
Figure 9. *Open loop test results for the sensor sub system for a device with 2mm x 1mm proof mass. The measured sensitivity is 160mV/g.*
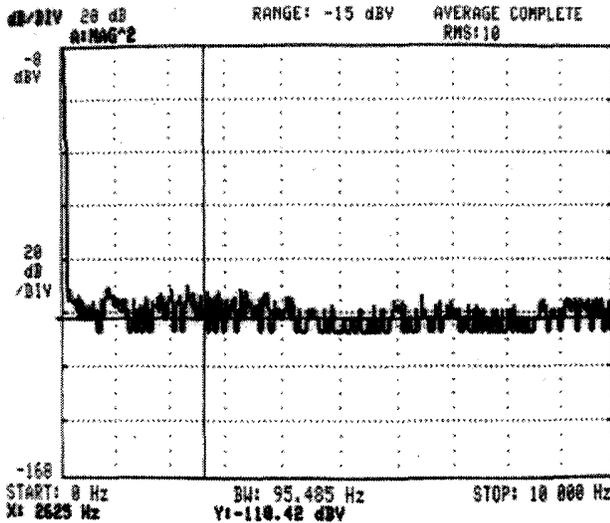
Figure 10 *Measured output spectrum of sensor/circuit module. Measured resolution of the system is ≈20μg/√Hz.*

SUMMARY

A high performance μg accelerometer with high sensitivity, low noise, and controllable damping is presented. This is achieved by combining surface and bulk micromachining and utilizes the whole wafer thickness to attain a large proof mass with integrated polysilicon electrodes. The measured sensitivity for a device with 2mm x 1mm proof mass and 1.4μm air gap is 1.4pF/g. The calculated mechanical noise floor for the device is 0.39μg/√Hz at

atmosphere. The circuit has a 95dB dynamic range, a low offset of 370μV and can resolve better than 75aF. The hybrid subsystem has a noise floor of 3.6μV/√Hz and a sensitivity of 160mV/g, indicating that the current system is capable of resolving about 20μg/√Hz. Table 2 summarizes the circuit and the circuit/sensor specifications.

Table 2. Summary of the circuit performance specifications and circuit/sensor performance

Interface IC Parameters	Value
Size	2.6 x 2.4 mm^2
Sampling clock	200kHz
Power dissipation	<6.6mW @ 5V
Capacitance sensitivity	0.3-1.1V/pF (adjustable)
Noise floor	85 μV/√Hz @ 1.1V/pF
Capacitance resolution	<75 aF
Dynamic range	95dB
Offset W/o switch chopper stabilization with chopper stabilization	2.7mV 370μV (82 ppm w.r.t. full range)
Sensor and Interface Module	**Value**
Equivalent acceleration resolution	~20μg/√Hz (measured)
Full scale range	±1.2g with 5V supply

REFERENCES

[1] N. Yazdi, F. Ayazi, and K. Najafi, "Micromachined Inertial Sensors," *Proc. IEEE,* vol. 86, pp. 1640-1659, Aug. 98
[2] C. Lu, M. Lemkin, and B. Boser, "A monolithic surface micromachined accelerometer with digital output," *IEEE J. Solid-State Circuits,* vol. 30, no. 12, pp. 1367-1373, Dec. 1995.
[3] N. Yazdi and K. Najafi, "An All-Silicon Single-Wafer Fabrication Technology for Precision Microaccelerometers," *Transducers '97,* Chicago, IL, USA, pp. 1181-1184, June 1997.
[4] T. Gabrielson, "Mechanical-Thermal Noise in Micro-Mechanical Acoustic and Vibration Sensors," *IEEE Trans. on Electron Devices,* vol. 40, pp 903-909, May 1993.
[5] A. Selvakumar and K. Najafi, "High-Density Comb Array Microactuators Fabricated Using A Novel Bulk/Poly-Silicon Trench Refill Technology," *Digest of Hilton-Head Solid-State Sensor and Actuator Workshop,* pp. 138-141, June 1994.
[6] N. Yazdi, and K. Najafi, "An Interface IC for A Capacitive Silicon μg Accelerometer,"*Tech. Digest, 1999 IEEE Int. Solid-State Circuits Conference* (ISSCC99), San Francisco, CA, USA, February 1999.

288

A High Aspect-Ratio Polysilicon Vibrating Ring Gyroscope

Farrokh Ayazi*, Hsiao H. Chen, Fatih Kocer, Guohong He, and Khalil Najafi

Center for Integrated MicroSystems
University of Michigan, 1301 Beal Avenue, Ann Arbor, MI 48109-2122
Email: najafi@umich.edu, Tel: (734) 763-6650, Fax: (734) 763-9324

*Currently with: Georgia Institute of Technology, Atlanta, GA 30332-0250

ABSTRACT

This paper presents the fabrication and testing of a high aspect-ratio 80μm tall polysilicon ring gyroscope (PRG) fabricated using a new *dry-release* poly-silicon MEMS technology. This single-wafer technology is capable of producing electrically isolated vertical electrodes as tall as the main body polysilicon structure (50 to 100's μm tall) with various size air-gaps ranging from sub-micron to tens of microns. An open-loop sensitivity of 200μV/deg/sec in a dynamic range of ±250 deg/sec was measured under low vacuum level for a prototype device tested in hybrid format. The resolution for a prototype sensor with a quality factor (Q) of 1200, drive amplitude of 0.15μm, and sense node parasitic capacitance of 2pF was measured to be less than 1 deg/sec in 1Hz bandwidth. Elimination of the parasitic capacitance and improvement in the quality factor of the ring structure will improve the resolution to 0.01 deg/sec in 1Hz bandwidth. Three hundred microns long clamped-clamped beam resonators fabricated in this technology have shown measured quality factors as high as 85,000 in 1mTorr vacuum.

Keywords: Gyroscope, high aspect-ratio, silicon micromachining, vibratory gyroscope, yaw rate sensor.

INTRODUCTION

High-performance microgyroscopes [1] are needed in many applications, including inertial navigation, control, and defense/avionics/space. The first nickel *vibrating ring* gyroscope was presented at the 1994 Hilton Head Workshop [2]. The vibrating ring gyroscope, shown in Fig. 1, consists of a ring, eight semicircular support springs, and drive, sense and control electrodes that are located around the structure. The ring is electrostatically vibrated into its first flexural mode with fixed amplitude. When the device is subjected to rotation around its normal axis, Coriolis force causes energy to be transferred from the first mode to the second flexural mode, which is located 45° apart from the first mode, causing amplitude to build up proportionally in the latter mode; this build-up is capacitively monitored. This device provides a number of advantages, including excellent mode matching, high-resolution, low zero-rate output, and long-term stability. The technological requirements for improving gyroscope performance and the scaling limits of ring gyroscopes were previously presented at MEMS 98 conference [3]. This paper reports the fabrication and testing of a high aspect-ratio (20:1), 80μm tall, 1.1mm in diameter polysilicon ring gyroscope (PRG) fabricated using a new *dry-release* poly-silicon MEMS technology [4]. This all silicon single-wafer technology is capable of producing electrically isolated vertical electrodes as tall as the main body structure (50 to 100's μm tall) with various size air gaps ranging from sub-micron to tens of microns. It provides all the features required for eventual realization of "inertial-grade" micromachined gyroscopes and other MEMS devices.
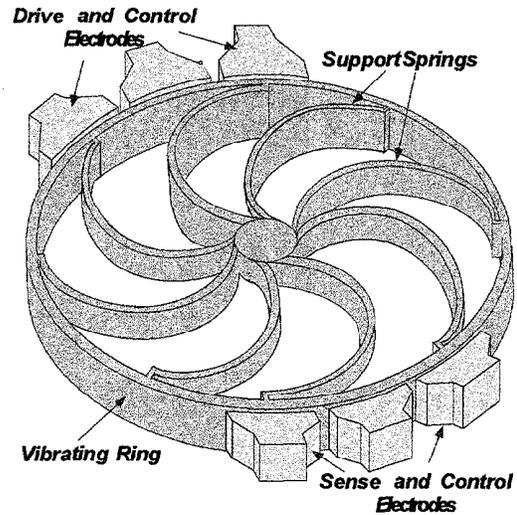
Figure 1: Structure of a vibrating ring gyroscope.

FABRICATION TECHNOLOGY

Figure 2 shows the fabrication process flow for the high aspect-ratio dry-release poly-silicon technology. First, a thin layer of LPCVD silicon nitride is deposited and patterned to serve as an isolation dielectric layer underneath the electrode bonding pads. Deep trenches with straight sidewalls (90°±1°) are then dry etched into a low-resistivity silicon substrate using the STS Silicon Deep Reactive Ion Etcher (DRIE). Trenches with smooth and straight sidewalls are needed to eliminate void formation in the polysilicon beams which will be formed by back-filling these trenches. Mechanical structures are created by refilling trenches with polysilicon deposited over a sacrificial oxide layer. The structural polysilicon layer has to be doped to lower its electrical resistance. Doping with boron is preferred over phosphorous due to the fact that the etch rate of boron-doped poly is much less than the etch rate of phosphorous-doped poly in the HF:H₂O (1:1) solution, leaving the poly beams intact after relatively long HF release step.

Silicon sense electrodes *as tall as the ring structure* are released from the substrate using a *dry directional/isotropic SF₆ silicon etch* performed in an STS etcher. This dry release step consists of a deep, directional etch (depth of the trenches + 20μm) followed by an isotropic SF₆ silicon etch. The sacrificial oxide layers are then etched away in HF:H₂O (1:1) solution to create submicron capacitive air gaps between the sense-electrodes and the ring structure.

This technology has been employed to fabricate a number of thick polysilicon vibrating shell gyroscopes [4] with silicon electrodes as tall as the shell structures. Figure 3 shows the SEM view of an 80μm tall, polysilicon ring gyroscope. The ring is

1.1mm in diameter and the diameter of the support post is 120μm. The width of the ring and support springs is 4μm. Sixteen electrodes are evenly located around the structure; they are approximately 60μm tall and 150μm long and are separated from the ring by a 1.4μm capacitive air-gap. Figure 4 shows close-up of a silicon electrode as tall as the main body structure separated from 80μm thick polysilicon ring by 1.4μm air-gap created through sacrificial oxide etching. The silicon is anchored on top to the supporting polysilicon layer, which is in turn anchored through the isolating nitride layer to the substrate. Vertical polysilicon stiffeners (trench-refilled) are also incorporated in the structure of the electrodes to improve rigidity.
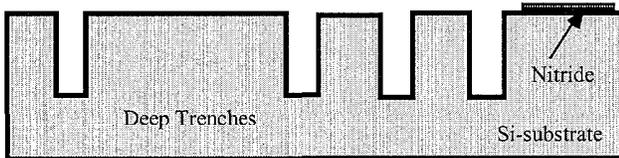
a) Deposit and pattern the isolating nitride layer;
b) dry etch deep trenches to define the main body structure.
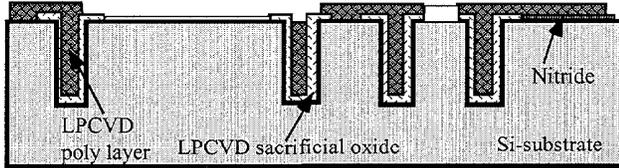
c) Deposit LPCVD sacrificial oxide and dope surface of the oxide; d) refill trenches with LPCVD polysilicon; e) etch back poly; f) pattern oxide; g) deposit, dope and pattern poly.
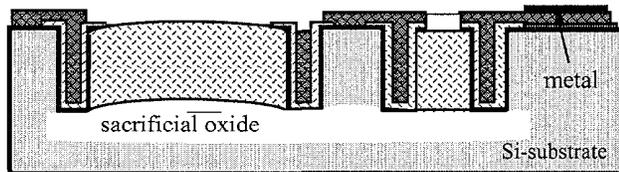
h) Deposit and pattern Cr/Au;
i) SF$_6$ dry directional etch + undercut and release electrodes;
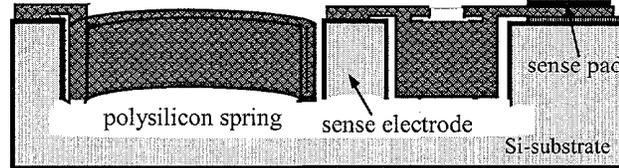
j) Etch the sacrificial oxide and completely release structures.

Figure 2: Fabrication process flow for the six-mask, high aspect-ratio dry-release poly-silicon MEMS technology.

Size of the air-gaps in this technology is not limited to sub-micron levels. Larger air-gaps can be realized with polysilicon trench-refilled electrodes. Larger air-gaps are defined by the distance between two adjacent trenches that form the parallel plates of a capacitor. The silicon between these trenches is then directionally etched in the STS machine during the deep release step. If the spacing is so small that etch openings cannot be placed between adjacent trenches, then the silicon between the two trenches is protected with resist on the top but undercut at the bottom and the sides during the deep release step. This narrow piece of silicon which is held by the oxide on the sidewall of trenches will then fall into the HF:H$_2$O solution during the sacrificial oxide etch. Figure 5 shows the SEM view of wide

capacitive gaps (10-20μm) with polysilicon electrodes fabricated using this technology.
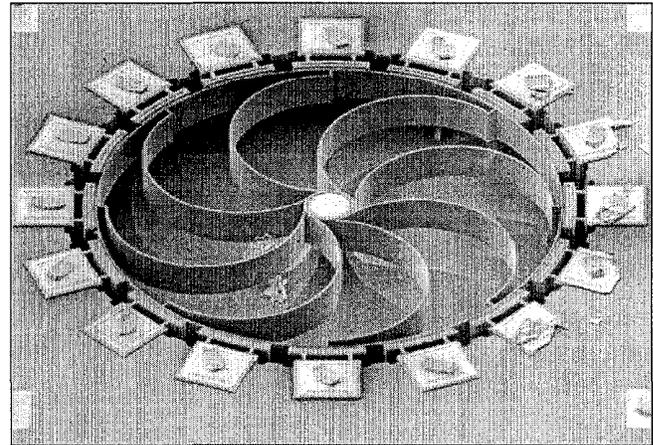
Figure 3: SEM view of a prototype 1.7mm×1.7mm polysilicon ring gyroscope fabricated through this technology. The dry-release process was entirely carried in the STS machine.
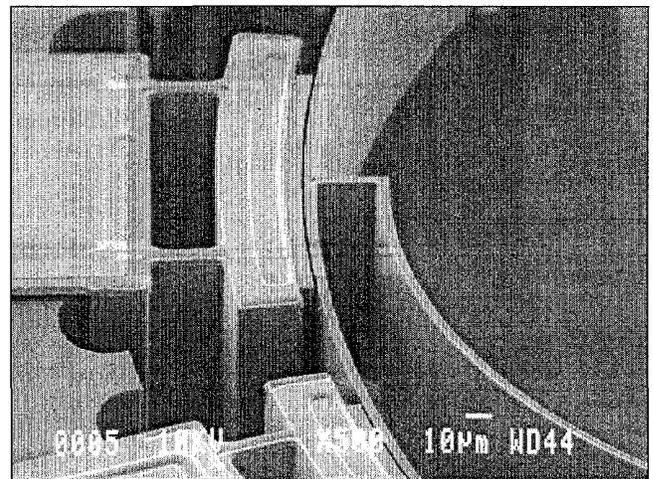
Figure 4: Close-up of a released sense electrode separated by 1.4μm air-gap from the vibrating ring. Each sense electrode is 60μm tall. Ring and springs are 4μm wide and 80μm tall.
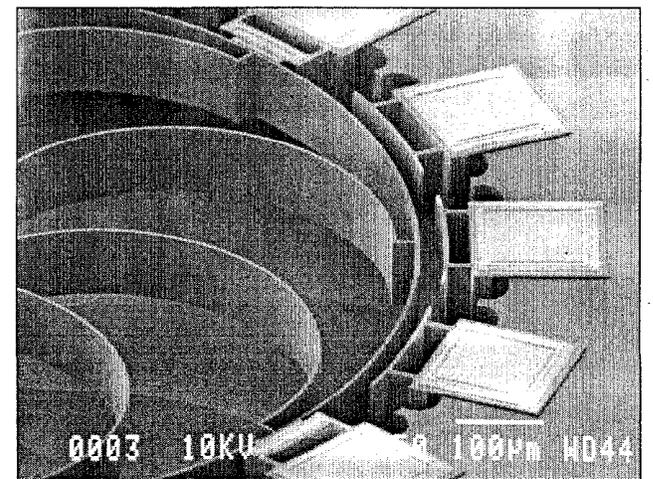
Figure 5: Various size air gaps from sub-micron to tens of micron.

This technology provides several important features required for high-performance microgyroscopes. First, the height of the polysilicon main body structure *"and the electrodes"* can be increased to 100's μm using DRIE. Second, since the body to sense-electrode gap is defined by the thickness of the sacrificial layer, it can be reduced to sub-micron level; *these two factors together will significantly increase the sense capacitance and hence the sensitivity*. Third, the ring to drive-electrode spacing can be made arbitrarily larger than the ring to sense-electrode spacing to achieve larger drive amplitude and in turn, reduce the noise floor. Fourth, improved long-term stability and temperature sensitivity is obtained due to the all-silicon feature of this technology.

TEST RESULTS

An 80μm tall prototype polysilicon ring gyroscope was tested open loop under vacuum. The sensor chip was connected in hybrid format to an NMOS source-follower buffer chip, as shown in Fig. 6, to measure ring vibrations. The low-impedance output signal of this buffer chip was then amplified off-chip using discrete amplifiers. Drive and control circuitry needed for tuning and open or closed loop operation of the ring gyroscope were implemented off chip using discrete components.

Electronic tuning is an attractive feature of the ring gyroscope. Any frequency mismatch between the sense and drive modes is electronically compensated using tuning electrodes around the structure. By matching the drive and sense mode frequencies, sensitivity is amplified by the Q of the structure. The frequency of the flexural mode of a prototype device is measured to be ~28.3kHz (with 7V bias applied to the ring) which agrees well with calculation and ANSYS simulation results. As shown in Fig. 7, drive and sense mode frequencies of the ring gyro were 63Hz apart due to imperfections. The balancing voltages required to null this 63Hz of frequency split between the modes were only ±0.9V. After balancing, the two peaks merge together and the sense and drive mode frequencies become equal. We have also been able to balance ring structures with up to 1kHz of frequency split (f_0=28kHz) by applying less than 6 Volts tuning voltages to the balancing electrodes. Small capacitive gaps (1.4μm) have indeed shifted the tuning voltages down to CMOS acceptable levels, increasing the tuning capability.

The *measured* Q of 1.1mm, 80μm-thick ring structure was ≈1000-2000. This is much smaller than anticipated due to excessive
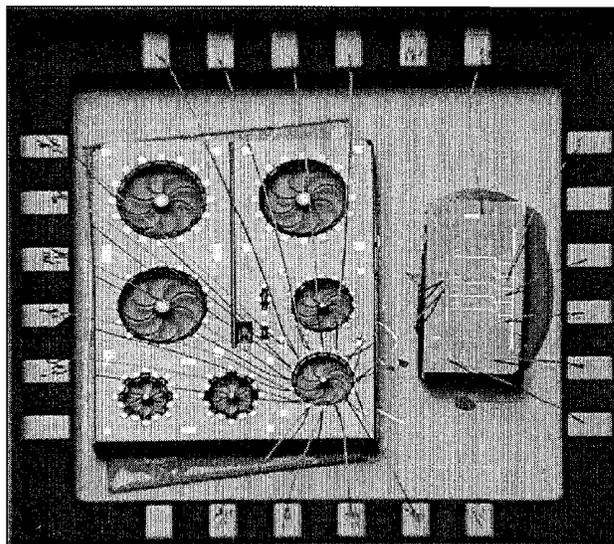
Figure 6: Hybrid attachment of the sensor chip to the buffer chip.

anchor and bulk losses (voids inside the poly beams). Modified devices with solid post design are expected to have higher Q in the range of 10,000-20,000 and hence higher sensitivity. *80μm tall, 300μm long clamped-clamped beam micromechanical resonators fabricated in this technology have* <u>shown quality factors as high as 85,000 in 1 mTorr vacuum</u> [4].

Figure 8 shows the response of a prototype ring gyroscope to a 5Hz sinusoidal input rotation rate with peak-to-peak amplitude of 120 deg/sec. Figure 9(a) shows the response of a prototype device to input rotation rates up to ±250 deg/sec in a 5Hz bandwidth. It should be noted that this measurement was carried out under poor vacuum level and the quality factor was about 250 which is much less than the actual Q of the ring structure under high vacuum. This drop in the quality factor reduces the open loop output signal by a factor of 5-8. Also, the output signal has been further reduced due to large parasitic capacitance in the output sense node. The parasitic capacitance associated with the hybrid connection (C_{inp} of the buffer + capacitance of bonding pads and wirebonds) have been measured to be around 2pF total. This large capacitance has degraded the output signal by an additional factor of 5. The sensitivity of a prototype device under poor vacuum (Q≈250) and large parasitics (C≈2pF) was measured to be 200μV/deg/sec.

(a)

(b)
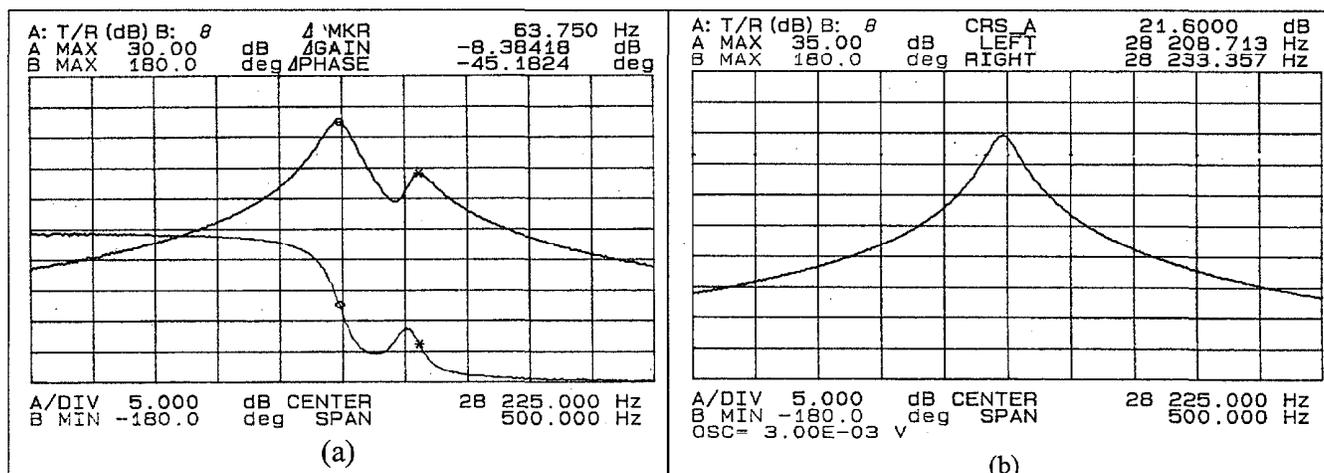
Figure 7: Electronic balancing of the ring gyroscope. (a) *Before balancing*: Two flexural resonant peaks of a prototype ring gyroscope have different frequencies (Δf=63Hz for f_0=28kHz). (b) *After balancing*: two peaks merge together and the frequencies become equal.
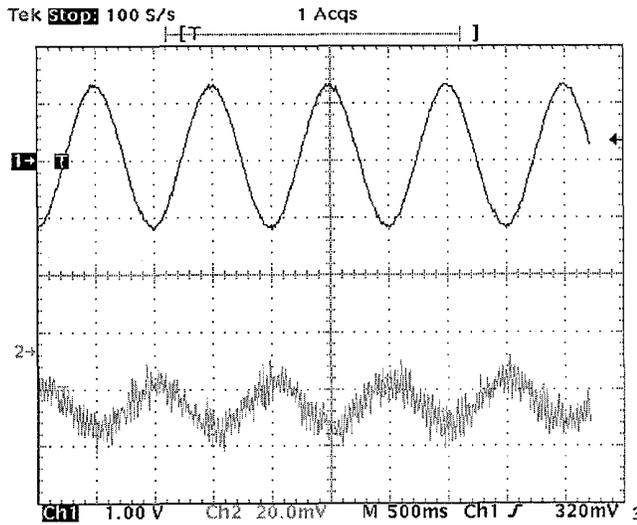
Figure 8: Response of a prototype single ring gyroscope to a 5Hz sinusoidal rate with a peak-to-peak amplitude of 120 deg/sec.

Another set of measurements was taken under improved vacuum levels. The Q was measured to be ≈1200 after balancing. Figure 9(b) shows the response of the sensor to rotation rates in the range of –20 deg/sec to 20 deg/sec. The resolution was measured to be ≈1 deg/sec. This resolution for a Q of 1200 and parasitic capacitances of 2pF is in agreement with theory. The resolution of the sensor is currently limited by the noise from the circuitry. Elimination of parasitic capacitances and improvement in Q will reduce the resolution to 0.01 deg/sec in 1Hz bandwidth, which is the Brownian noise floor for the single ring polysilicon gyroscope with a drive amplitude of 0.15μm (Table 1). By increasing the drive amplitude to 1μm in an asymmetric polysilicon ring gyroscope design, a minimum detectable signal of 5×10^{-3} deg/sec (18 deg/h) in a 10Hz bandwidth can be achieved.

CONCLUSIONS

This paper presents the fabrication and testing of a high aspect-ratio 80μm tall polysilicon ring gyroscope fabricated using a new *dry-release* poly-silicon MEMS technology. An open-loop sensitivity of 200μV/deg/sec in a dynamic range of ±250 deg/sec was measured under low vacuum level for a prototype device tested in hybrid format. The resolution for a prototype sensor with a quality factor (Q) of 1200, drive amplitude of 0.15μm, and sense-

node parasitic capacitance of 2pF was measured to be less than 1 deg/sec in 1Hz bandwidth. Elimination of the parasitic capacitance and improvement in the quality factor of the ring structure will improve the resolution to 0.01 deg/sec in 1Hz bandwidth.

Table 1: Specifications and calculated performance of the polysilicon vibrating ring gyroscope.

Design Parameter	Target Value
Material Quality Factor	20,000
Ring Diameter	1.1 mm
Sense Electrode Gap Spacing	1.4 μm
Ring and Springs Width	4 μm
Height of the Ring Structure	80 μm
Support Post Diameter	120 μm
Height of Each Electrode	60 μm
Resonant Frequency	27.3 kHz
Interface Circuit Input Noise (V_n)	$1\ \mu V/(Hz)^{0.5}$
Parasitic Capacitance ($C_{device}+C_{pad}$)	500 fF
Polarization Voltage (V_p)	7V
Drive amplitude	0.15μm
Brownian Noise Floor	$0.01\ deg/sec/(Hz)^{0.5}$
Min. Detectable Signal (10Hz BW)	**0.04 deg/sec**

ACKNOWLEDGMENTS

REFERENCES

[1] N. Yazdi, F. Ayazi, and K. Najafi, "Micromachined Inertial Sensors," *Invited paper, Proceedings of the IEEE,* pp.1640-1659, Aug. 1998.

[2] M. W. Putty and K. Najafi, "A Micromachined Vibrating Ring Gyroscope", *Digest, Solid-State Sensors and Actuators Workshop,* Hilton Head, SC, June 1994, pp. 213-220.

[3] F. Ayazi and K. Najafi, "Design and Fabrication of a High-Performance Polysilicon Vibrating Ring Gyroscope," in *Proc. 1998 IEEE/ASME Micro Electro Mechanical Systems Workshop,* Heidelberg, Germany, 1998, p. 621.

[4] F. Ayazi and K. Najafi, "High Aspect-Ratio Dry-Release Poly-Silicon MEMS Technology for Inertial Grade Microgyroscopes," in *Proc. IEEE Position Location And Navigation Symposium (PLANS 2000),* San Diego, CA, March 2000, pp. 304-308.

(a)

(b)
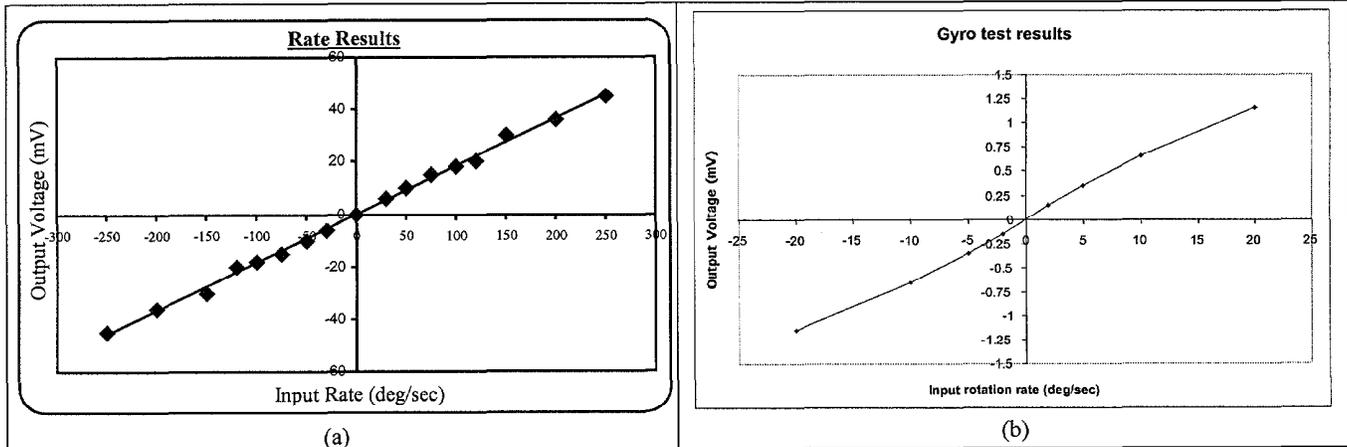
Figure 9: (a) Measured rate results obtained from a prototype ring gyroscope in a 5Hz BW, range=±250°/sec; (b) Measured rate results obtained from a prototype ring gyroscope in a 5Hz BW, range=±25 °/sec.

YAW SENSOR DESIGN OPTIMIZATION USING FINITE ELEMENT METHOD

George Q. Jiang
Delphi Delco Electronics Systems, MS Fab A, P.O. Box 9005,
Kokomo, IN 46904-9005

ABSTRACT

Delphi Delco Electronics Systems, in collaboration with the Delphi R&D Center, has developed a surface-micromachined angular rate (yaw) sensor for automotive applications. This sensor is based on the same principle used for the vibrating shell gyroscope. The sensing element consists of a combined ring and comb structure, which is supported by eight springs attached to a center hub. Modal analysis has been performed on the sensing element to study the mode shapes of different designs using the finite element method (FEM). It was discovered that mode shapes were distorted due to the asymmetry of the supporting spring geometry in previous designs. The distorted mode shape limited the sensor's performance. The finite element method was also used to optimize a novel spring design to restore the sensing element's mode shape. The optimized design has eliminated the sensor stiction at hight temperature, reduced the total harmonic distortion (THD), simplified application specific integrated circuit (ASIC) design and sensor calibration, and improved sensor performance and yield.

INTRODUCTION

The surface-micromachined, vibrating ring angular rate sensor is fabricated with a electroforming process that allows for the inclusion of large amounts of on-chip circuitry with the sensing element on the same die. Wafer-level vacuum packaging has also been developed to further reduce the cost [1]. It is the low cost that enables wider automotive applications including yaw control, rollover protection, and turn-by-turn navigation presently implemented only in the luxury-vehicle market.

The sensing element is based on the principle of the shell vibrating in its fundamental flexural mode analyzed by G. H. Bryan in 1890 [2]. In this mode of vibration, the lip of the shell (or the rim of a ring) vibrates in an elliptical shape that has two nodal diameters. The four points on the ring that have no radial deflection are called nodes, and the four points with maximum radial deflection, located 45° from the nodes, are called antinodes (Figure 1). When the device is rotated, the Coriolis force causes the vibrating pattern to rotate and the amplitude of vibration at the defined nodal points to increase. To measure angular rate of rotation, the vibration signal at the nodes is fed back to reduce the amplitude, thereby maintaining a stationary pattern with nodes at the defined locations. The signal is proportional to the angular rate and is used as the output [3].

This principle has been successfully applied to a navigation-quality gyroscope developed by Delco Systems Operations for many years [4]. To simplify the manufacturing process and to reduce the cost, a surface-micromachined, vibrating-ring angular rate sensor was developed at the Delphi R&D Center (formerly the GM Research Center). Delphi Delco Electronics Systems has improved the design further and made it a manufacturable, low-cost product. The sensing element consists of three structures: the ring itself; the eight combs attached to the ring at the antinodes and

nodes; and eight support springs with one end attached to the ring and the other end attached to the center hub anchored on the substrate (Photo 1). The interdigitated combs form the parallel plate capacitors that are used for either driving the ring to resonance or detecting the rate signal. Different spring designs, such as a semicircle (C-shaped) or alternating, conjoined semicircles (S-shaped) [not shown], were used to support the ring and comb.
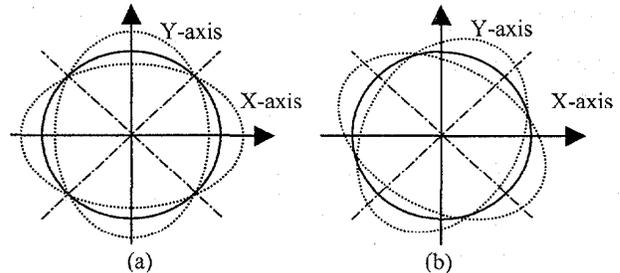
Figure 1. Mode shape of an ideal ring at flexural mode. (a) vibration pattern at station (b) vibration pattern at rotation
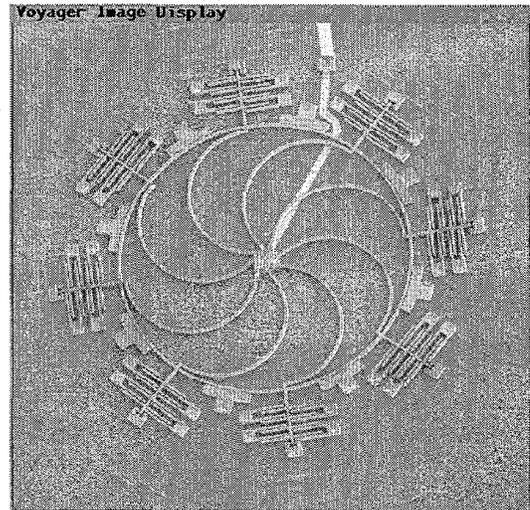
Photo 1. Resonating ring-comb angular rate sensor

ANALYSIS AND RESULTS

It was very difficult to calibrate the prior-generation yaw sensor to a scale factor variation of less than 3% over the automotive operating temperature range (-40 to 85°C). The sensor also showed large, asymmetrical THD over the operating temperature range. At high temperature, the combs touched and shorted to the fixed electrodes, resulting in failure. To investigate the source of the problem, analyses were performed on the mode

shape of the spring-supported, vibrating-ring yaw sensors and the results were compared with the mode shape of the ideal ring.

The Mode Shape of an Ideal Ring

The mode shape of an ideal ring can be expressed as

$$U_r = A_m \times cos(2\theta) \qquad (1)$$
$$U_\theta = 0.5A_m \times sin(2\theta) \qquad (2)$$

where U_r is the radial deflection of the ring
U_θ is the tangential motion of the ring
A_m is the maximum amplitude of the deflection
θ is the angle of the ring with respect to the x-axis

The mode shape of an ideal ring can also be plotted as in Figure 2. It can be seen that at antinodes, which are located at 0 (360), 90, 180 and 270 degrees, the ring has maximum radial deflection while the tangential motion is zero. On the other hand, at the nodes, which are located at 45, 135, 225 and 315 degrees, the ring has maximum tangential motion but no radial deflection.

The FEM is also used to obtain the mode shape of an ideal vibrating-ring. The results are the same as shown in Figure 2.

$$(\underline{\quad} U_r \, ; \, \underline{\cdots} \, U_\theta)$$
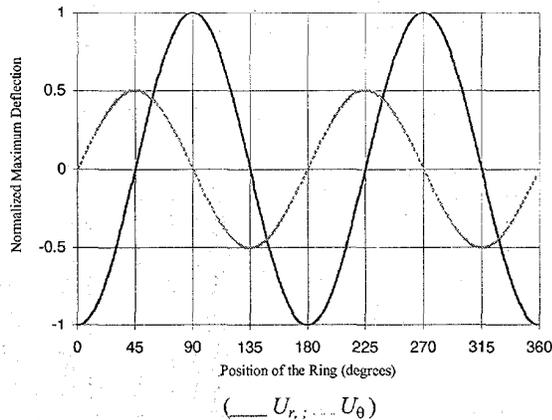
Figure 2. Mode shape of an ideal ring

The Mode Shape of a Ring Supported by Springs

Due to the complexity of the structure, no analytical mode shape solution is available for the springs-supported ring-comb sensing element. The FEM program, ANSYS®, is used to perform modal analysis on these structures. The element SHELL43 is used to build the model. Nodal displacement is extracted from the simulation result file and analyzed. The mode shape of the ring-comb structure supported by C-shaped springs is plotted in Figure 3.

It is clear from Figure 3 that the antinodes, where no tangential motion exists, and the nodes, where no radial motion exists, are no longer 45 degrees apart. Furthermore, the mode shape is not symmetrical about either antinodes or nodes as it is for an ideal ring. In other words, if the nodes are not aligned with the sensing electrodes at 45, 135, 225, and 315 degrees, the radial motion will cause capacitance change or offset even when the yaw sensor is not subjected to angular rotation. If the balance mechanism is used to align the nodes with sensing electrodes located at 45, 135, 225, and 315 degrees, the antinodes will be shifted away. The electrodes located at 0, 90, 180, and 270 degrees will see tangential motion in addition to radial motion. This

additional motion distorts the signal at the pickup electrodes, increases the 2× frequency signal and therefore the total harmonic distortion (THD), which makes the task of calibrating the gyroscope's performance over a temperature range difficult. The tangential motion of the ring also causes comb rocking and even failure due to combs touching and becoming stuck to the sensing electrodes.

$$(\underline{\quad} U_r \, ; \, \underline{\cdots} \, U_\theta)$$
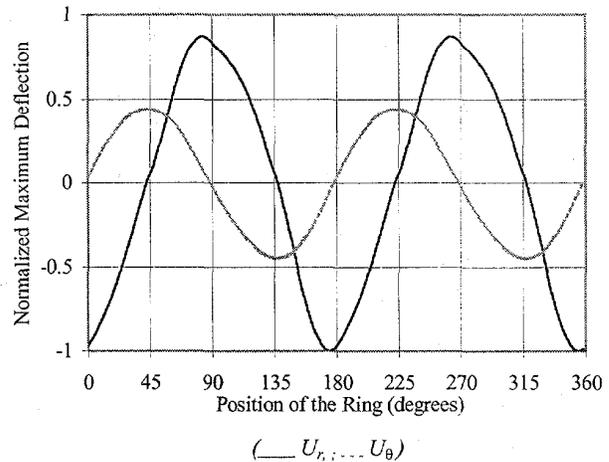
Figure 3. Mode shape of a ring supported by C-shaped springs

The mode shape of the resonance ring yaw sensors, supported by alternating, conjoined semicircle, S-shaped springs, has also been analyzed using ANSYS®. The simulation shows similar mode-shape distortion.

DESIGN OPTIMIZATION

The cause of mode-shape distortion aforementioned is the asymmetry of the previous spring designs, such as the C-shaped or the S-shaped spring configurations. To miniaturize the device and maintain low resonance frequency, however, a simple and asymmetrical shape of springs is often the choice. To minimize the mode-shape distortion, a novel-shaped spring system was invented which contains eight double-semicircle, S-shaped springs with two different radii (Photo 2). The ratio of the radius of the large semicircle to the radius of the small semicircle is optimized so that the mode-shape distortion will be minimized. With eight combs attached to the ring, the ratio of the radius of the large semicircle to the radius of the small semicircle is chosen as 3.28 : 1. Figure 4 shows the simulation results of the mode shape of a ring supported by the optimized springs.

It can be seen that in the optimized supporting spring design, the tangential motion at antinodes and the radial deflection at nodes are minimized. The mode shape of the resonating yaw sensor has been restored.

SIGNIFICANT IMPACT

Using the optimized supporting-spring design, the high temperature stiction problem has been eliminated. The THD over temperature has been reduced by 50% (Figure5). The most significant impact is that the combination of the new spring design and a simplified and matched ASIC design has dramatically increased the yield of the high-performance yaw sensor. The optimized design meets the requirement of a scale factor variation less than 3% over automotive operating temperatures with a greatly improved yield.
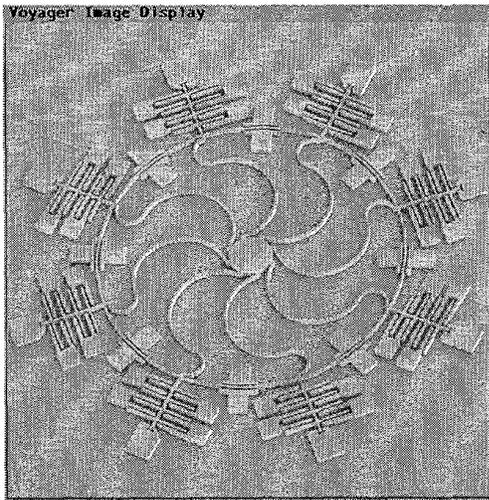
Photo 2. *Resonance ring yaw sensor supported by the optimized springs*

$$(\underline{\quad\quad} U_r\,;\,\underline{\cdot\,\cdot\,\cdot}\,U_\theta)$$
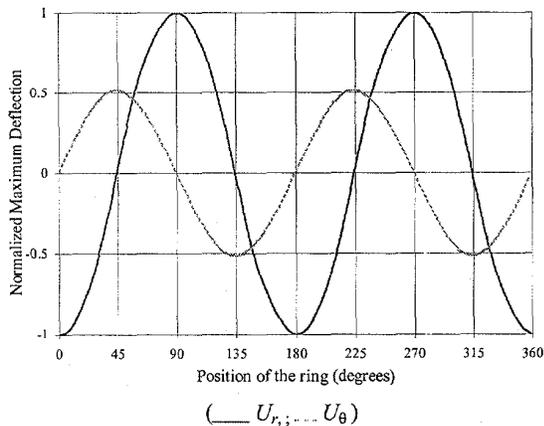
Figure 4. *Mode shape of a ring supported by optimized springs*
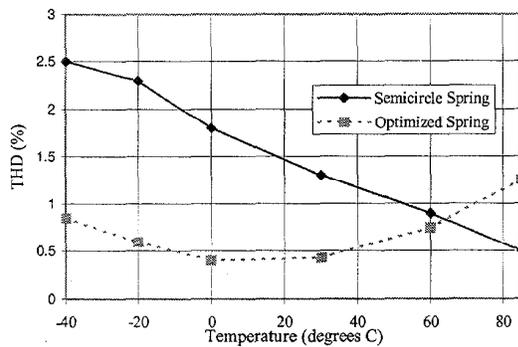
Figure 5. *THD of different spring designs over operating temperature*

CONCLUSIONS

The finite element method has been used to analyze the mode shape of the resonating-ring angular rate sensor. It is found that the mode shape of prior-generation sensors was distorted by the asymmetry of C- and S-shaped supporting springs. Mode-shape distortion caused undesired radial deflection at the nodes

and tangential motion at the antinodes. This, in turn, caused the following: large asymmetry THD over the operating temperature range, sensing structures shorted at high temperature, difficulty in calibration, and low yield.

The finite element method has also been used to develop a novel spring design in which each spring is composed of two semicircle-shaped springs with different radii. The ratio of the two radii is optimized so that the distortion of the mode shape of the vibrating ring is minimized. The new design has eliminated the sensor stiction at high temperature, minimized the THD, simplified the calibration process, and greatly improved the yield of the high-performance yaw sensor.

ACKNOWLEDGMENTS

REFERENCES

1. S. Zarabadi, T.Vas, D. Sparks, J. Johnson, Q. Jiang, M. Chia, E. Borzabadi, A Resonance Comb/Ring Angular Rate Sensor Vacuum Packaged Via Wafer Bonding, Spring SAE Tech. Proc., No. 1999-01-1043.
2. G. H. Bryan, On a Revolving Cylinder or Bell, *Proceedings of the Royal Society (London)*, p. 47, 1890.
3. M. Putty, K. Najafi, A Micromachined Vibrating Ring Gyroscope, *Solid-State Sensors and Actuators Workshop*, Hilton Head, South Carolina, USA, pp. 213-320, June, 1994.
4. E. Loper, D. Lynch, Projected System Performance Based on Recent HRG Test Results, *Proc. IEEE 5th Digital Avionics System Conf.* Seattle, Washington, USA, pp. 18.1.1-18.1.6, 1983.

SENSE FINGER DYNAMICS IN A ΣΔ FORCE-FEEDBACK GYROSCOPE

Joseph I. Seeger, Xuesong Jiang, Michael Kraft*, and Bernhard E. Boser

Berkeley Sensor & Actuator Center
University of California, Berkeley, CA 94720-1774

*Department of Electronics and Computer Science, Southampton University
Southampton SO17 1BJ, United Kingdom

ABSTRACT

This paper presents a gyroscope that uses ΣΔ force-feedback for closed-loop Coriolis-acceleration sensing and 1-bit digital output. The noise floor is approximately $3°/s/\sqrt{Hz}$ at atmospheric pressure. The effect of capacitive sense-finger resonance on the ΣΔ feedback loop is analyzed. In order to minimize quantization noise the resonant frequency of the sense fingers should be designed above $f_s/4$, if the sense and feedback electrodes are separate. If the same electrodes are used for sensing and feedback, the finger resonance should be designed near $f_s/6$, where f_s is the sampling frequency.

INTRODUCTION

A vibratory-rate gyroscope determines the angular velocity of an object by vibrating a proof-mass attached to the object and measuring its Coriolis acceleration. Previous micromachined gyroscopes have demonstrated both open-loop [1] and closed-loop [2] Coriolis-acceleration sensing with analog outputs. Closed-loop sensing has the potential to increase the sensor bandwidth, improve linearity, and reduce sensitivity to process and temperature variations. This is especially important if the frequency of the Coriolis acceleration matches the proof-mass resonant frequency. Sigma-delta (ΣΔ) modulation is a technique that provides high-resolution analog-to-digital conversion [3]. An approach demonstrated in accelerometers [4], combines ΣΔ modulation and force-feedback to make a sensor in which acceleration signals are converted directly to digital output. Recently, a ΣΔ force-feedback accelerometer has reported 18-bit resolution [5].

Closed-loop sensing can also cause instability, and ΣΔ modulation tends to increase the sensor noise due to quantization error. In ΣΔ force-feedback, the mechanical design of the proof-mass affects both the stability and the quantization noise. With proper design, ΣΔ force-feedback can provide a digital output without degrading the sensor performance. This paper presents a gyroscope that uses ΣΔ force-feedback for closed-loop Coriolis acceleration sensing and digital output. Additionally, this paper investigates the effect of capacitive sense finger resonance on the ΣΔ force-feedback performance.

DESIGN SUMMARY

Figure 1 shows a monolithic, surface-micromachined, vibratory gyroscope that is sensitive to rotations about the Z-axis, the axis normal to the plane of the chip. The gyroscope was fabricated by Sandia National Labs in an integrated MEMS process with a 2.25μm thick mechanical polysilicon layer and 5V, 2μm minimum gate length CMOS transistors. The gyroscope operates at atmospheric pressure.
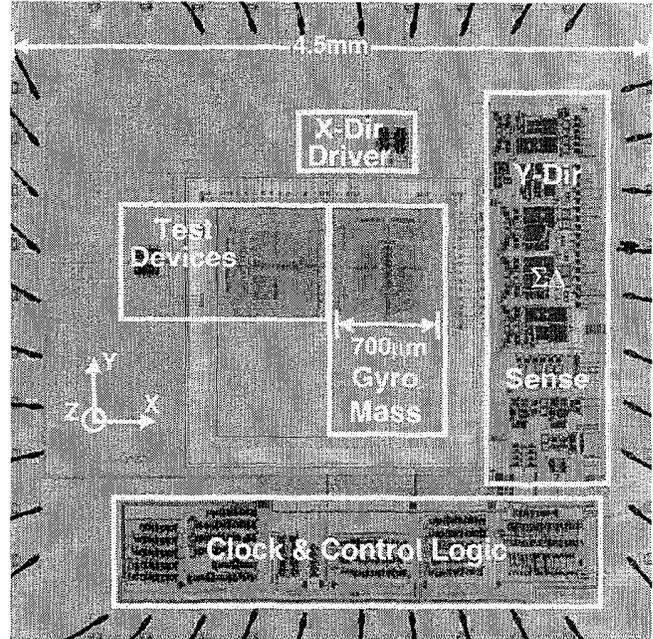
***Figure 1.** Photograph of the gyroscope die*

Figure 2 shows a block diagram of the gyroscope. It can be divided roughly into three parts: the proof-mass, the electronics for vibrating the proof-mass, and the electronics for sensing Coriolis acceleration. The proof-mass is vibrated along the X-axis (the drive mode), and its acceleration is measured along the Y-axis (the sense mode). The proof-mass is represented by the drive and sense dynamics, which ideally are coupled only by the Coriolis acceleration. The resonant frequency and quality factor in air, before electrostatic spring tuning, are approximately 10.5kHz and 4 in the sense mode and 9.5kHz and 3 in the drive mode. The proof-mass is 0.3μg. It is a single structure (Figure 3) that consists of three sets of gap-closing electrodes: electrostatic actuation in the X-direction (70fF, 2μm gap), capacitive sensing in the Y-direction (70fF, 1.2μm gap), and force feedback in the Y-direction (2fF, 1.2μm gap). A fourth set of electrodes cancel quadrature coupling between the drive and sense axes [1].

The drive electronics consist of an off-chip oscillator and a 1-bit quantizer on chip that switches a fixed voltage between the two fixed drive electrodes. The structure is driven approximately ±1μm at its resonant frequency using 3.2V. In this prototype design, the drive mode is open-loop, without any electronic position sensing.

The sense electronics implement a ΣΔ feedback loop that is similar to [4] and is explained in detail in [6]. The loop consists of 1) a capacitance sensing amplifier for position measurement; 2) correlated double sampling (CDS) to cancel offset and flicker
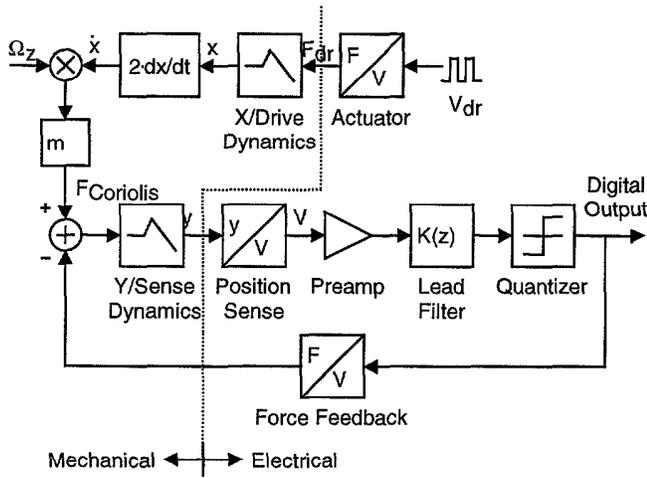
Figure 2. $\Sigma\Delta$ *Gyroscope block diagram*
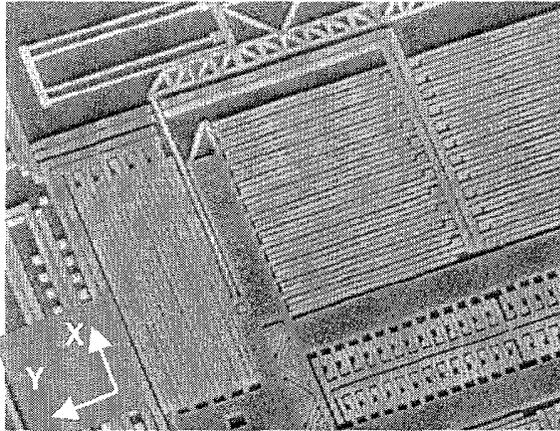
Figure 3. *SEM of the proof-mass*

noise; 3) a gain stage; 4) a discrete-time lead compensator; 5) a one-bit quantizer, and 6) a feedback pulse that drives one of the two fixed feedback electrodes. The sense dynamics act as a noise-shaping filter for the $\Sigma\Delta$ loop. The sampling frequency f_s is 1MHz.

SENSE DYNAMICS

Ideally, the proof-mass dynamics in the sense-direction can be modeled with one resonant mode by a lumped mass, spring, and damper. In reality, the proof-mass is a distributed element with many resonant modes. Figure 4 illustrates a model of the sense dynamics that includes a flexible beam, representing the electrodes used for position sensing, and a lumped mass, representing the rest of the structure. The flexible beam has length L and thickness Nh, where N is the number of sense electrodes, and h is the thickness of the proof-mass. The spring represents the suspension compliance in the sense direction. The model also includes damping on the lumped mass, distributed damping on the flexible beam, and a distributed spring to account for electrostatic tuning of the sense fingers. All elements are constrained to move in the Y-direction. The deflection of the beam $y(x,t)$ is a function of position and time. The deflection of the lumped mass is $y(0,t)$, and the deflection of
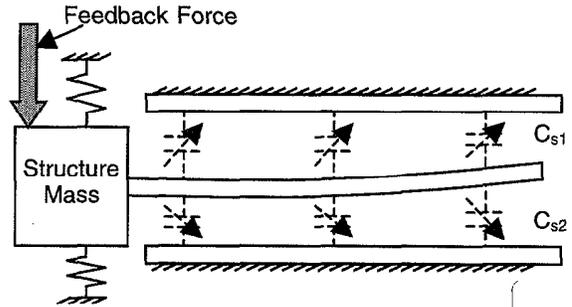
Figure 4a. *Model of the non-collocated sense dynamics*
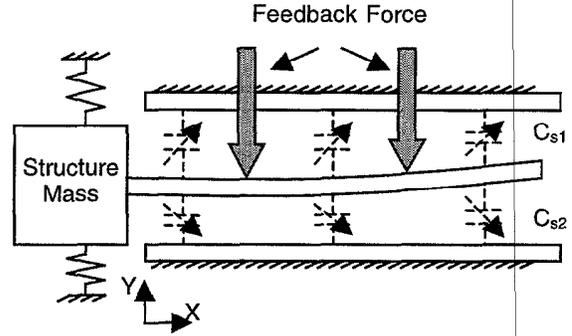
Figure 4b. *Model of the collocated sense dynamics*
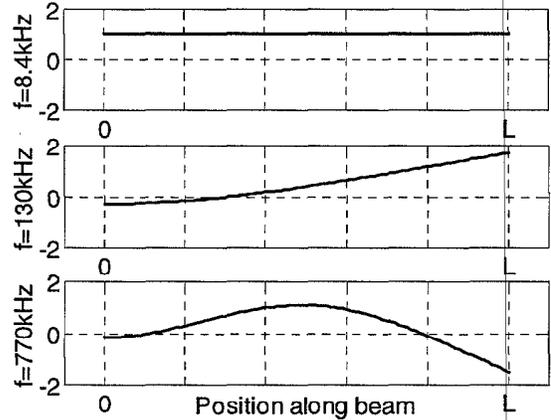
Figure 5. *First three computed sense-finger mode shapes*

the beam tip is $y(L,t)$. Neglecting fringing fields, the total capacitance from the moving, flexible electrode to the upper, fixed electrode is approximately:

$$C_{s1}(t) = \varepsilon N h \int_0^L \frac{1}{d - y(x,t)} dx,$$

where ε is the permittivity of air, and d is the nominal gap. The position sense electronics measure the differential capacitance, C_{s1}-C_{s2}, which can be simplified, assuming small deflections, to

$$\Delta C_s(t) \approx \frac{2\varepsilon h}{d} \int_0^L \frac{y(x,t)}{d} dx.$$
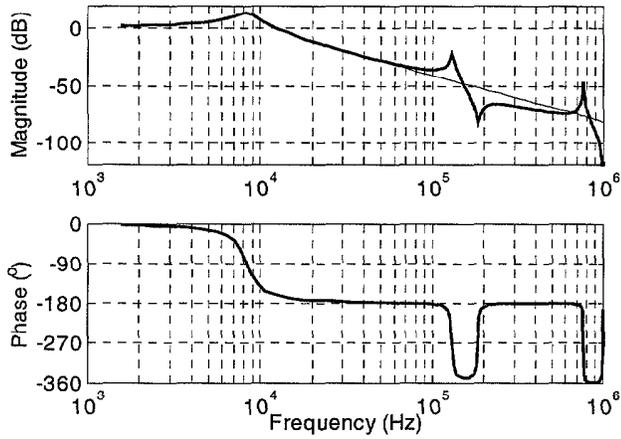
Figure 6. Non-collocated (thick) and second-order (thin) frequency response of the average beam position to the feedback force.
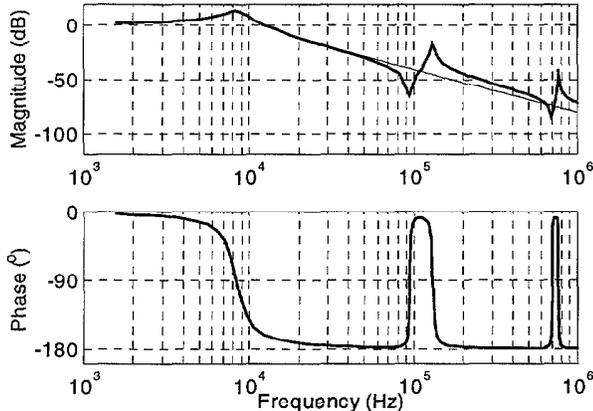
Figure 7. Collocated (thick) and second-order (thin) frequency response of the average beam position to the feedback force.

The feedback force can be applied to the lumped mass (Figure 4a) or distributed across the sense finger (Figure 4b). The former case, the non-collocated design, models an implementation in which the feedback electrodes are separate from the sense electrodes. The latter case, the collocated design, models an implementation, such as [4], in which the fixed electrodes are time multiplexed for sensing and force-feedback.

The resonant frequencies and mode shapes are independent of where the force is applied. Figure 5 shows the first three mode shapes and their associated resonant frequencies computed using the Rayleigh-Ritz method [7]. The lowest resonant mode corresponds to the lumped mass and the sense electrodes moving together. The resonant frequency is approximately $\omega_0 = \sqrt{(k - k_{el})/m}$, where k is the suspension spring constant, k_{el} is the negative spring constant due to electrostatic tuning, and m is the total mass of the structure. The higher resonant modes correspond to small motion of the lumped mass and significant bending of the capacitive comb finger. The finger resonant frequencies can be approximated by the resonant frequencies of a cantilevered beam with a rectangular cross-section

$$\omega_i = \alpha_i^2 \frac{W}{L^2} \sqrt{\frac{E}{12\rho}} \quad \alpha_1 = 1.875, \alpha_2 = 4.694$$
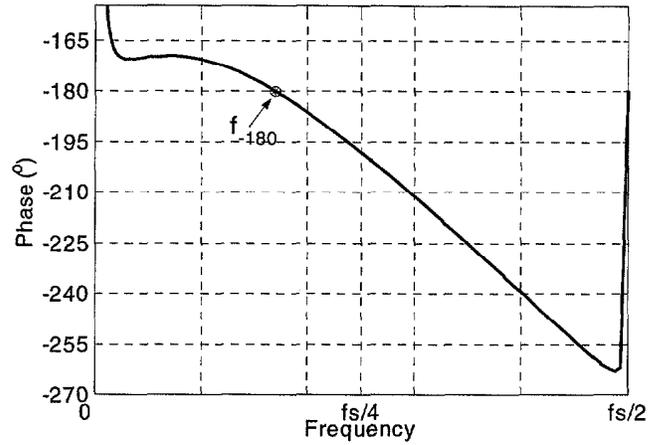
Figure 8. Phase versus frequency of second-order dynamics, discrete-time compensator, and delayed feedback pulse.

where i is the mode number, W is the finger width, E is Young's modulus, and ρ is the material density [7].

The Bode plot of differential capacitance response to a sinusoidal feedback force is shown in Figure 6 for the non-collocated design. As is typical of non-collocated dynamics, there is an additional 180° of phase lag at frequencies above the resonant mode of the finger. Figure 7 shows the frequency response for the collocated design. Collocated dynamics provide 180° of phase lead below the finger resonance.

ΣΔ LOOP ANALYSIS

Even though the 1-bit quantizer is a nonlinear element, the ΣΔ loop can be analyzed using linear system techniques. The quantizer can be modeled by a linear gain plus a quantization error source. Reference [8] calculates the effective gain and the quantization error assuming that the signal at the input of the quantizer is white noise. The effective gain of the quantizer varies as a function of the signal at the input of the quantizer.

The output of the ΣΔ modulator includes quantization noise that is shaped by the noise transfer function, $1/(1+T(j\omega))$, where $T(j\omega)$ represents the open-loop gain. In order to reduce the quantization noise contribution in the signal band, the open-loop gain should be maximized in the signal band. However, the open-loop gain must be limited to prevent instability.

With the ideal second-order sense dynamics, the open-loop gain and hence the ΣΔ modulator performance is limited by the phase lag due to correlated double sampling and due to the time delay between the position sense and the feedback pulse. To compensate for this lag, a lead network is included in the ΣΔ loop. Figure 8 shows a plot of the phase versus frequency for an open-loop system consisting of second order dynamics, a lead compensator, $C(z)=2-z^{-1}$, and a feedback pulse centered about $T_s/2$, one-half of the clock period. Below $f_s/6$, the net phase is greater than −180° due to the lead compensator. At higher frequencies, the phase drops below −180° due to the time delay of the feedback pulse. The −180° crossover frequency f_{-180} represents the maximum unity-gain frequency for stable closed loop operation. This frequency sets a limit for the low-frequency gain. If f_s is increased, f_{-180} is also increased and therefore the loop gain can be increased. Thus, in the ideal case, increasing the sampling frequency tends to reduce the quantization noise. The −180° crossover frequency varies depending on the feedback delay, CDS,
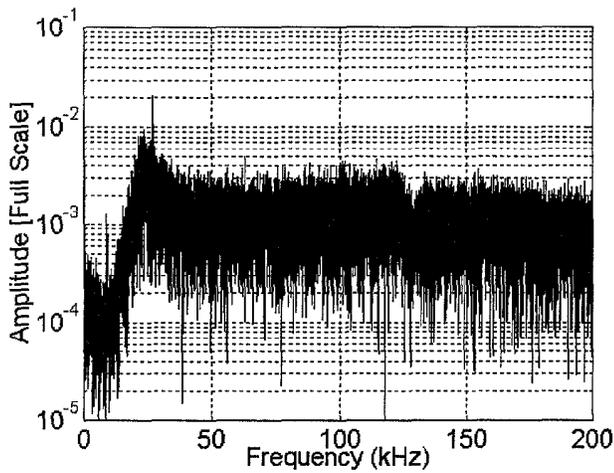
Figure 9. *Measured bitstream output spectrum*

Figure 10. *Measured response to a 50°/s, 4Hz sinusoidal rotation rate (0.25Hz/bin)*

and aliasing. For reasonable values, f_{-180} varies between $f_s/6$ and $f_s/4$.

The non-collocated dynamics introduce an additional 180° of phase lag at the first resonant frequency of the flexible finger (ω_1). The phase lag due to the finger resonance sets an upper limit on the −180° crossover frequency. If the resonant frequency of the finger is below $f_s/4$, then f_{-180} approximately equals the finger resonant frequency and the maximum loop gain is reduced from the ideal. In this case, the $\Sigma\Delta$ loop might oscillate at the finger resonant frequency. If the finger resonance is greater than $f_s/4$, the phase lag due to the finger resonance does not degrade the $\Sigma\Delta$ performance. In the case of non-collocated dynamics, increasing f_s does not reduce quantization as much as in the ideal case.

The collocated dynamics introduce 180° of phase lead below the finger resonance. Collocated dynamics are stable in feedback provided that there is sufficient phase at the resonant frequencies of the flexible beam. If the resonant frequency is above $f_s/6$, the compensator does not provide sufficient phase lead. The increased amplitude response near the resonant frequency of the beam significantly reduces the maximum low-frequency loop gain. In this case, the $\Sigma\Delta$ loop might oscillate at the beam resonant frequency. If the resonant frequency is below $f_s/6$, the compensator provides sufficient phase lead to maintain stability. If the resonant frequency is too low, a higher order resonant mode of the beam that is above $f_s/6$ can degrade the $\Sigma\Delta$ performance.

EXPERIMENTAL RESULTS

The sense mode was designed with non-collocated feedback. The first resonant mode of the sense electrodes is 130kHz. A 150,000 point FFT of the 1-bit $\Sigma\Delta$ modulator output is shown in Figure 9. The spectrum matches simulations run in SIMULINK. The peak at 130kHz is due to the resonant mode of the sense fingers. The peak at 22kHz indicates the unity-gain frequency. The unity gain frequency is significantly less than the finger resonance. This results from increased electronic noise at the quantizer input which reduces the effective quantizer gain. Thus, the finger dynamics do not affect this implementation, but they would become significant if the unity gain frequency were increased either by increasing the feedback force or by decreasing the electronic noise.

Simulations indicate that the residual motion of the proof-mass is 2.2Å_{rms} and the sense fingers flex 1.3Å_{rms}. Figure 10 shows the narrow-band response to a 50°/s, 4Hz sinusoidal
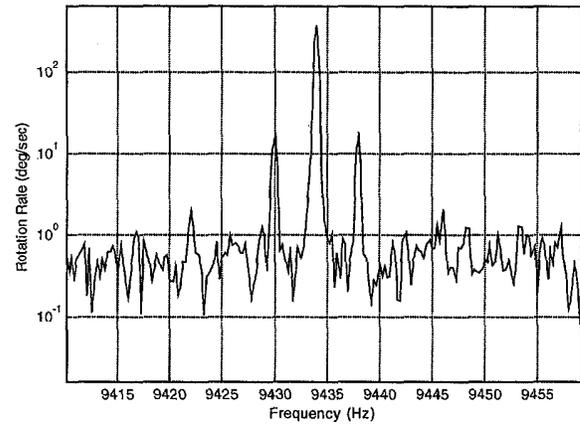
rotation. The noise floor is approximately $3°/s/\sqrt{\text{Hz}}$. The performance is limited by quantization noise.

CONCLUSIONS

The $\Sigma\Delta$ force-feedback gyroscope is functional. The resonant mode of the long capacitive fingers used for position sensing is excited by feedback, and it has been measured. In order not to degrade the $\Sigma\Delta$ performance, if the feedback and sense electrodes are separate, the sense fingers should be short enough so that the resonant frequency is above $f_s/4$. If the sense electrodes are also used for feedback, the fingers should be made long enough so that the resonant frequency is about $f_s/6$. In theory, the finger dynamics can limit the $\Sigma\Delta$ performance, but in practice, they are not the only factor that limits the $\Sigma\Delta$ force-feedback performance.

ACKNOWLEDGEMENT

REFERENCES

1. W.A. Clark, R.T. Howe, and R. Horowitz, "Surface Micromachined Z-Axis Vibratory Rate Gyroscope," *Solid-State Sensor and Actuator Workshop*, Hilton Head Island, SC, June (1996), p.283-287.
2. S. An, Y.S. Oh, K.Y. Park, S.S. Lee, and C.M. Song, "Dual-Axis Microgyroscope with Closed-Loop Detection," *Sensors and Actuators, 73*, (1999) pp. 1-6.
3. S.R. Norsworthy, R. Schreier, and G.C. Temes, *Delta-Sigma Data Converters*. Piscataway, NJ: IEEE Press, (1997).
4. M. Lemkin and B.E. Boser, "A Three-Axis Micromachined Accelerometer with a CMOS Position-Sense Interface and Digital Offset-Trim Electronics," *IEEE Journal of Solid-State Circuits*, April (1999). p.456-68.
5. M.A. Lemkin, T.N. Juneau, W.A. Clark, T.A Toessig, and T.J. Brosnihan, "A Low-Noise Digital Accelerometer Using Integrated SOI-MEMS Technology," *Transducers '99*, Sendai Japan, June (1999) pp. 1294-1297
6. X. Jiang, J.I. Seeger, M. Kraft and B.E. Boser, "A Monolithic Surface Micromachined Z-Axis Gyroscope with Digital Output," *Symposium on VLSI Circuits*, Honolulu, HI, June (2000).
7. L. Meirovitch, *Principles and Techniques of Vibrations*, Upper Saddle River, NJ: Prentice Hall, (1997). pp. 403, 518-535
8. S. H. Ardalan and J. J. Paulos, "An Analysis of Nonlinear Behavior in Delta-Sigma Modulators," *IEEE Transactions on Circuits and Systems*, June (1987), p.593-603.

Development of a Wafer-Bonded, Silicon-Nitride Membrane Thermal Shear-Stress Sensor with Platinum Sensing Element

Anthony Cain[§], Venkataraman Chandrasekaran, Toshikazu Nishida and Mark Sheplak
Interdisciplinary Microsystems Group
University of Florida
Gainesville, Florida 32611-6250

ABSTRACT

This paper presents the development of a thermal shear-stress sensor for quantitative turbulence measurements. The shear-stress sensor was fabricated by an aggressive wafer bonding process that produces a 1500 $\mathring{A}$-thick silicon nitride membrane over a vacuum cavity which supports a platinum sensing element. Static sensitivity data at multiple overheat ratios (11 *mV/Pa* at an overheat of 1.0), pressure sensitivity spectra (< 1 $\mu V/Pa$), noise floor spectra (100 $nV/\sqrt{Hz}$) and direct, *in-situ* dynamic calibration data are presented. These calibrations were performed using a four-point probe configuration to exclude the effects of external bias circuitry and to isolate the sensor performance.

INTRODUCTION

Wall shear stress measurement within a turbulent boundary layer provides insight into complex flow phenomena, including skin friction, flow separation and reattachment. The mean component indicates the average state of the flow over a surface, and the dynamic component is a signature of the processes responsible for the unsteady momentum transfer to the wall [1]. The fluid structures of interest are dimensionally on the order of 100 micrometers, requiring a sensor element of the same size to minimize spatial averaging. Also, a shear stress sensor with a usable bandwidth of over 10 *kHz* is needed to accurately capture the complete spectrum of turbulent fluctuations [2].

Conventional shear stress measurement techniques can be broadly classified into two types: direct (*e.g.,* floating element devices) and indirect (*e.g.,* thermal sensors). Micromachined varieties of both of these types of sensors have been developed for turbulence measurements [3-6]. Thermal sensors [3,4] are generally simpler to fabricate and are more robust than floating element sensors [5,6]. However, these require an empirical correlation, valid for very specific conditions, to relate the Joulean heating rate to the induced shear stress [7]. Since the device operates on thermal transfer principles, the ratio of the convection of heat into the flow (used for shear stress transduction) and the various storage and dissipation mechanisms ultimately limit the dynamic response and the operational bandwidth. In particular, the unsteady heat conduction into the substrate results in a roll-off in the gain factor of the frequency response function (FRF), as well as a phase lag at low frequencies [8]. Therefore, these sensors require extensive *in-situ* calibration before they can be utilized for quantitative measurements in a flow field. Liu *et al.* [3] presented a novel polycrystalline silicon-based sensor on a membrane over a vacuum cavity to reduce substrate conduction effects. This device showed improved sensitivity over traditional devices, but the sensor operation was limited to moderate thermal overheats (0.65), and an experimental verification of the FRF using a known shear-stress input was not reported.

In this work, a shear stress sensor is presented that employs a thin-film platinum sensing element on top of a silicon-nitride

[§] Anthony Cain is presently affiliated with the Air Force Research Laboratory, AFRL/MLPO, WPAFB, Dayton, OH 45433-7707.

membrane. This membrane is stretched over a circular vacuum cavity with the goal of providing an extended thermal operating range, a lower noise floor, and reduced pressure sensitivity. Advantages of platinum-based sensors over polycrystalline silicon-based sensors include higher thermal coefficient of resistance (TCR), higher thermal operating range, reduced 1/*f* noise [9], and no piezoresistive-induced pressure sensitivity.

The details of the fabrication process for the shear-stress sensor are provided and experimental results presented. Specifically, rigorous static and dynamic characterizations of the sensor are reported for a constant current (CC) mode of excitation. In addition, the experimental verification of potential error sources is provided, including dynamic pressure sensitivity and noise floor spectra.

DESIGN CONSIDERATIONS

Typical metals used for conventional hot-film and hot-wire sensors include tungsten, platinum, and platinum-iridium [10]. It is desirable to use a material with a high temperature coefficient of resistance (TCR), high resistivity, and low thermal conductivity. The first two parameters will determine the sensitivity to velocity or shear stress and the latter parameter will affect the amount of heat that is conducted into the electrical contacts. Of the three metals, bulk platinum offers the second highest TCR, the highest resistivity, the highest operating temperature and the second lowest thermal conductivity. It is also favorable because it does not form an oxide. Due to these factors, platinum is widely considered to be the optimal metal for use as a hot-film sensor [10].

During operation, the sensing element is resistively heated to a temperature greater than the gas temperature defined by the non-dimensional thermal overheat ratio,

$$a_T = \frac{T_s - T_g}{T_g}$$

where T_s and T_g are the absolute sensor and gas temperatures, respectively. The convection of heat from the sensor is related to the wall shear stress by the Reynolds analogy [11] and is measured by monitoring changes in the temperature-dependent resistance of the sensing element:

$$R_s = R_r \left[1 + \alpha (T_s - T_r) \right],$$

where R_s is the sensor resistance, T_r is a reference temperature corresponding to the sensor reference resistance, R_r, and α is the TCR. The sensor can be operated in one of three biasing modes: constant current (CC), constant voltage (CV), or constant temperature (CT). Constant current and constant voltage modes are open-loop systems that use constant current and voltage sources, respectively, to provide heating power to the sensor. These modes do not regulate sensor temperature, and thus the sensor resistance varies with changes in the fluid temperature and velocity. Consequently, when thermal overheat is specified for CC and CV, it is the zero mean flow overheat relative to the initial fluid temperature. In comparison, the CT mode is a

closed-loop system, providing heating power to the sensor to maintain a constant sensor temperature and thus resistance [10].

To accurately capture the complete spectrum of shear stress fluctuations in a turbulent flow, a sensor should possess a flat, zero-gain, minimum-phase frequency response. This dictates that the sensor have equal static and dynamic sensitivities over the frequency range of interest. In general, the inherent compliance, inertance and dissipation of the system limit the dynamic response. The limiting factors are thermal inertia of the substrate and the sensing element, which restrict the usable bandwidth of the sensor [12] and the unsteady heat conduction into the substrate. This further complicates the dynamic response. In order to minimize uncertainties in correlation and spectral analysis of the measured shear stress, an *in-situ* dynamic calibration technique is required to identify the frequency response function (FRF).

SENSOR FABRICATION

The device structure consists of a 1500 $\overset{o}{A}$-thick x 4 μm-wide x 200 μm-long platinum sensing element on top of a 1500 $\overset{o}{A}$-thick silicon nitride membrane which seals a 200 μm-diameter and 10 μm-deep vacuum cavity. Figure 1 is a plan-view SEM of the active area of the shear-stress sensor. Two gold leads at each end of the sensing element permit 4-point probe characterization exclusive of the effects of the biasing circuitry. Titanium serves as an adhesion layer for the leads and sensing element. The shear-stress sensors were fabricated using a silicon-rich silicon-nitride/silicon fusion-bond and thin-back process [13] outlined in Figure 2. The process begins with two p-type <100> single-polished silicon wafers. Cavities are etched via DRIE in the handle wafer followed by growth of 7000 $\overset{o}{A}$ of thermal oxide as a bonding layer. The membrane material, 1500 $\overset{o}{A}$ of silicon-rich silicon nitride, is deposited on the device wafer and polished via CMP in preparation for bonding (A). The wafers are bonded in 100% O_2 ambient and annealed at 1100 $^{\circ}C$ for one hour. A vacuum is formed within the cavities during the anneal as O_2 is consumed by the growth of SiO_2 [14]. The handle wafer is protected and the device wafer is thinned back to the silicon nitride using 20% KOH at 60 $^{\circ}C$ (B). The sensing element and leads are patterned via image reversal lithography and titanium, platinum, and gold are realized via E-beam deposition. The sensing element and leads are realized via lift-off and the gold is etched from the sensing element using a KI wet etch (C). This compact, three-mask process yielded strict diaphragm geometry control (±3%) for the vacuum cavity. Several die configurations were designed including 9-sensor arrays for wind tunnel testing and single sensor dies with and without vacuum cavities for sensor characterization. Packaging for characterization experiments was accomplished by flush mounting the sensor dies within a Lucite plug with front to back electrical contacts. The sensor leads are ball-bonded to the package contacts using 0.001*in.* gold wire and are oriented downstream of the sensor to minimize flow disruption during testing.
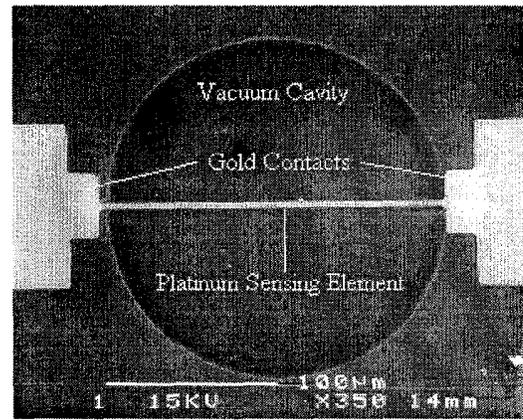
Figure 1. Plan-view SEM of the active area of the shear stress sensor.
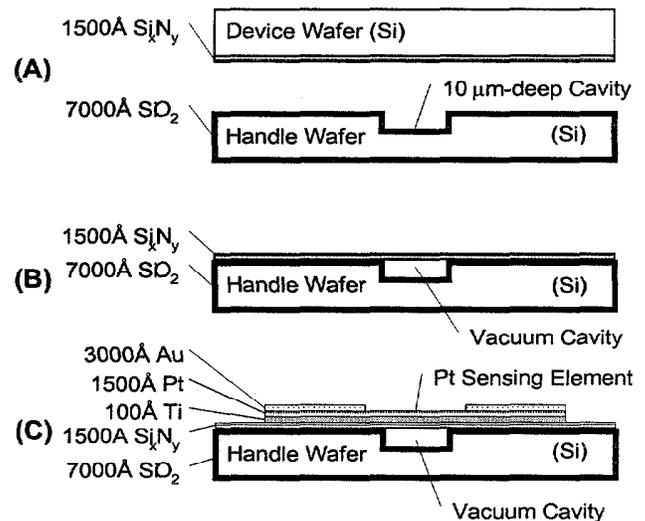
Figure 2. Cross-sectional schematic of the sensor fabrication sequence.

RESULTS AND DISCUSSION

All sensor measurements were performed using a 4-point probe configuration to exclude the effects of the external circuitry and the parasitic lead impedance. This entails maintaining a constant current through the sensing element and monitoring the varying voltage across the sensor using a Keithley 2400 source meter.

The temperature coefficient of resistance of the thin-film platinum resistor was experimentally determined using a high-temperature alumina substrate jig that permitted a 4-point resistance measurement of the sensor within a furnace. A TCR of 0.0029 $^{\circ}C^{-1}$ possessing a maximum non-linearity of 2.7% over a temperature range of 20 $^{\circ}C$ to 400 $^{\circ}C$ was observed (Figure 3). The effectiveness of the vacuum cavity in thermally isolating the sensing element from the substrate was investigated by comparing the sensor overheat as a function of input power for sensors with and without a cavity (Figure 4). The sensors with cavities exhibited much greater overheat than those without cavities. The sensors without cavities could only achieve a maximum overheat of 0.01 before exceeding a destructive current density limit of 1×10^7 A/cm^2.
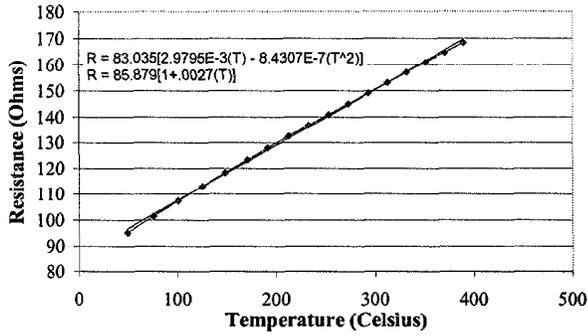
Figure 3. *Resistance vs. temperature relationship for the thin-film platinum sensing element.*
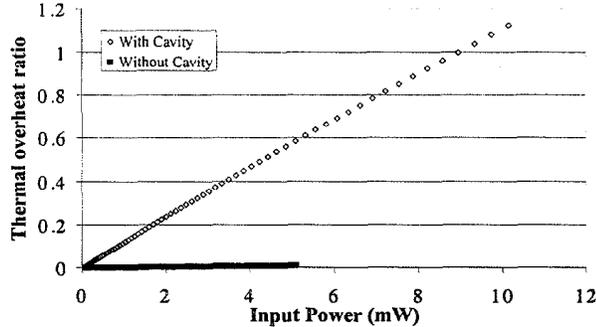
Figure 4. *Thermal overheat ratio vs. input power for sensors with and without cavities.*

Static wall shear stress sensitivity experiments were performed in CC mode of excitation using a laminar flow cell for thermal overheats of 0.2 - 1.0 and wall shear stresses from 0 *Pa* - 1.7 *Pa* (Figure 5). The static sensitivity increased with higher thermal overheat with a maximum sensitivity of 11 *mV/Pa* at an overheat of 1.0. A lumped, two-dimensional power balance relationship for hot film sensors [15] is commonly used as the fitting equation for these calibrations. This relationship suggests a $1/3^{rd}$-power dependence of the sensor heating power on shear stress based on several major assumptions, most notably, that the thermal boundary layer lies within the linear region of the velocity boundary layer and the sensing element lies on a semi-infinite substrate. This device did not exhibit the $1/3^{rd}$-power dependence which is physically plausible since the sensor is located on a diaphragm over a vacuum cavity and not on a semi-infinite substrate.

Dynamic wall shear stress sensitivities were obtained at multiple mean shear stress levels and overheats, using a novel, *in-situ* dynamic calibration technique. This technique provides known sinusoidal shear-stress perturbations generated via acoustic plane-wave excitation [16,17]. These calibrations were performed in a constant-current mode of excitation for thermal overheats of 0.6 - 1.0 and for mean shear stress levels of 0.03 *Pa* to 0.06 *Pa*. A constant amplitude acoustic excitation of 105 dB SPL (ref. 20 *μPa*) was used for the calibration. The dynamic shear stress FRF is a function of the static sensitivity as shown in Figure 6 but is independent of the overheat ratio as seen by the overlapping curves in Figure 7. An increase in dynamic sensitivity with higher thermal overheats was observed and in general, the sensor exhibited ≈ 40 *dB/decade* roll-off with a corner frequency of ≈ 600 *Hz*, indicative of a highly damped 2nd order system. Previous work [18] uses a 1/2-order system to

model the response of a sensor on a semi-infinite medium. The difference in response is explained by the presence of a sealed vacuum cavity that drastically reduces the unsteady heat conduction into the substrate and limits the dissipation due to conduction losses into the thin membrane.

The dynamic range of the sensor is ultimately limited by its cross sensitivities to non-shear-stress inputs and the device noise floor. Experiments were performed to quantitatively measure both the pressure sensitivity and the noise floor. The dynamic pressure sensitivity spectrum was determined within a plane-wave tube by orienting the sensor normal to the propagation of the acoustic waves under zero mean flow conditions. Frequency sweeps at multiple sound pressure levels ranging from 120 *dB* to 155 *dB* (ref. 20 *μPa*) at a constant overheat of 1.0 were used for the measurements. The results indicate negligible pressure sensitivity (< 1 *μV/Pa*, Figure 8) up to 10 *kHz*. The noise floor spectra at zero mean flow and multiple overheats were also measured (< 100 *nV/√Hz*) and is shown in Figure 9. Electromagnetic interference appears to be the major contributor at low frequencies with all harmonics decaying within 1 *kHz*. Any observed response of the sensor to pressure fluctuations may also be due to buoyant convection from the sensor.
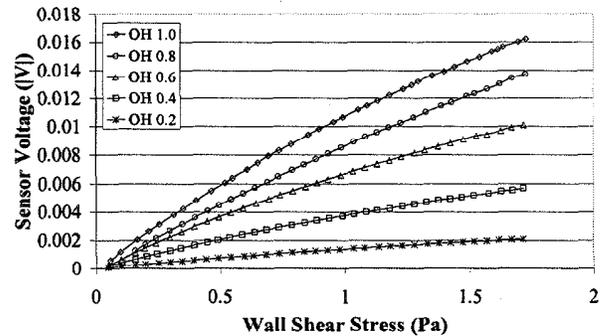
Figure 5. *Change in sensor voltage as a function of wall shear stress at multiple thermal overheats for the sensor with vacuum cavity (4th order polynomial curve fits included).*
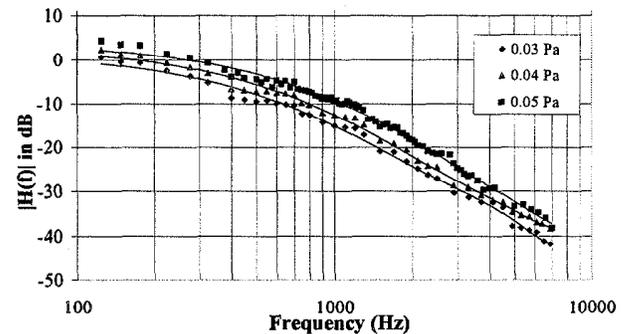
Figure 6. *Frequency response function of the shear stress sensor at an overheat of 0.81 for multiple mean shear stress levels (4th order polynomial curve fits included).*
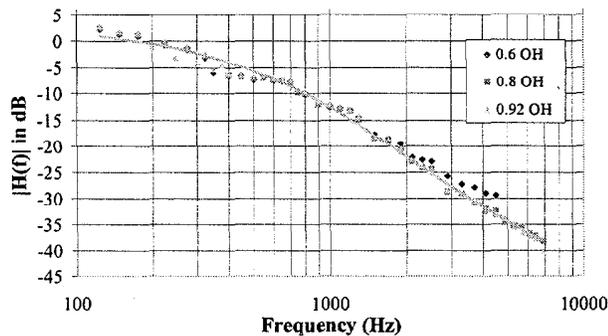
Figure 7. Frequency response function of the shear stress sensor at a mean shear stress of 0.04 Pa for multiple overheat ratios (4th order polynomial curve fits included).
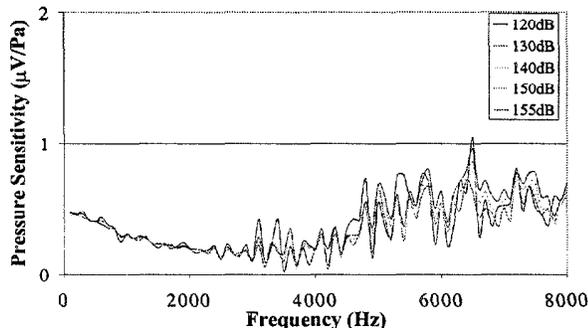
Figure 8. Dynamic pressure sensitivity spectrum of the shear stress sensor with cavity at multiple sound pressure levels (dB ref 20 μPa).
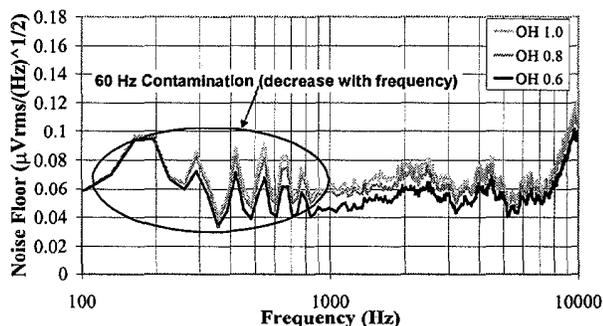
Figure 9. Measured noise spectrum of the shear stress sensor at zero mean shear stress and multiple overheats.

CONCLUSIONS AND FUTURE WORK

A thermal shear stress sensor for wind-tunnel measurements has been developed. The devices were fabricated using a novel wafer-bonding process that permits the fabrication of a sealed cavity thermal isolation structure. Thorough characterization of the sensor in a four-point probe configuration to exclude the effects of the external circuitry has been performed, including an *in-situ* dynamic calibration. Noise floor measurements in the operational range of the sensor have been obtained. The dynamic range of operation of the sensor has been experimentally verified (9 μPa-1.7 Pa) and the FRF obtained from 100 Hz to 8 kHz. The sensitivity to pressure and its effect on the dynamic response has been shown to be negligible. Future work will focus on characterizing the sensor

in a constant temperature mode of operation and investigating the effects of sensor orientation in a flow field.

REFERENCES

1. J. H. Haritonidis, "The Measurement of Wall Shear Stress," in Advances in Fluid Mechanics Measurements, Springer-Verlag, (1989), pp. 229-261.
2. A. Padmanabhan, *et al.*, "Micromachined Sensors for Static and Dynamic Shear Stress Measurements in Aerodynamic Flows," *Technical Digest*, Transducers '97, Chicago, IL, (1997), 137-140.
3. C. Liu, *et al.*, "A Micromachined Flow Shear-Stress Sensor Based on Thermal Transfer Principles," *JMEMS*, Vol. 8, No. 1, Mar. (1999), pp. 90-99.
4. E. Kalvesten, "Pressure and Wall Shear Stress Sensors for Turbulence Measurements," Ph.D. Dissertation, Royal Institute of Technology, Stockholm, Sweden, (1996).
5. A. Padmanabhan, *et al.*, "Wafer-bonded floating-element shear stress microsensor with optical position sensing by photodiodes," *JMEMS*, Vol. 5. No. 4, Dec. (1996), pp. 307-315.
6. T. Pan, *et al.*, "Microfabricated Shear Stress Sensors, Part 1: Design and Fabrication," *AIAA Journal*, Vol. 37, No. 1, Jan. (1999).
7. V. A. Sandborn, Resistance Temperature Transducers, Metrology Press, Fort Collins, Co., (1972).
8. B. J. Bellhouse, *et al.*, "The Determination of Fluctuating Velocity in Air with Thin Film Gauges," *J. Fluid Mech.*, Vol. 29, No.2, (1967), pp. 289-295.
9. S-L Jang, "A Model of 1/f Noise in Polysilicon Resistors," *Solid-State Electronics*, Vol. 33, (1990), pp. 1155-1162.
10. L. M. Fingerson, *et al.*, "Thermal Anemometers," in Fluid Mechanics Measurements, R. J. Goldstein, Ed., Hemisphere, New York, (1983) pp. 99-154.
11. F. M. White, Viscous Fluid Flow, McGraw-Hill, NY, (1974), pp. 123-124.
12. B. J. Bellhouse, *et al.*, "The Measurement of Fluctuating Skin Friction in Air with Heated Thin-Film Gauges," *J. Fluid Mech*, Vol. 32, part 4, (1968), pp. 675-680.
13. M. Sheplak, *et al.*, " Dielectrically-Isolated, Single-Crystal Silicon, Piezoresistive Microphone," Technical Digest, Solid-State Sensor and Actuator Workshop, Hilton Head, SC, 23-26, June 1998.
14. M. A. Huff, *et al.*, "Design of Sealed Cavity Microstructures Formed by Silicon Wafer Bonding," *JMEMS, Vol. 2, No. 2, June (1993), pp. 74-81.*
15. T. J. Hanratty, *et al.*, "Measurement of Wall Shear Stress," in Fluid Mechanics Measurements, R. J. Goldstein, Ed., Hemisphere, New York, (1983) pp. 99-154.
16. M. Sheplak, *et al.*, "Dynamic Calibration of a Shear Stress Sensor using Stokes Layer Excitation," AIAA Paper 98-0585, (1998).
17. V. Chandrasekaran, *et al.*, "Dynamic Calibration Technique for Thermal Shear Stress Sensors with Variable Mean Duct Flows," AIAA Paper 2000-0508.
18. S. C. Ling, *et al.*, "The Hot Film Anemometer: A New Device for Fluid Mechanics Research," *J. Aero. Sci.*, 23, (1956), pp. 890-891.

MEMS Thermal Shear-Stress Sensors: Experiments, Theory and Modeling

Qiao Lin, Fukang Jiang, Xuan-Qi Wang, Zhigang Han, Yu-Chong Tai
James Lew* and Chih-Ming Ho*
Electrical Engineering 136-93, Caltech, Pasadena, CA 91125
* MAE Department, UCLA, Los Angeles, CA 90024

ABSTRACT

We have made many versions of MEMS thermal shear-stress sensors which have received successful applications. However, it has been found that the classical theory for conventional thermal shear-stress sensors is inapplicable to MEMS sensors. This paper then presents a systematic study of this issue and for the first time, an adequate theoretical analysis of MEMS sensors is developed and examined by experiments. The resulting 2D MEMS shear-stress sensor theory, which includes heat transfer effects ignored by the classical theory, is verified by experimental data. We also perform 3D heat transfer simulation and the results agree with the testing data and support the proposed new theory.

INTRODUCTION

When a fluid flows over a solid surface, viscous effects generate shear stress at the surface. The measurement of the surface shear stress, which has important applications in fluid dynamics and control [1], can be achieved by thermal sensing methods. This has been demonstrated with several different designs of thermal shear-stress sensors [2-5]. When compared to their conventional counterparts, MEMS sensors are extremely small in size (~ 200 μm) and offer superior spatial resolution and minimal flow interference. Due to improved thermal isolation using a vacuum or air cavity underneath the sensing element (Fig. 1), these sensors are also highly sensitive and consume very little power. While the effort in developing micromachined shear-stress sensors has been successful, a thorough theoretical understanding of the MEMS sensor operation has been lacking. In fact, it has been experimentally observed that our MEMS devices often disagree with the classical hot-wire or hot-film theory, which states that the heat removed by the flow is proportional to the 1/3-power of the shear stress [5]. This suggests that there may be phenomena that the classical theory does not consider, and that a new theory should be developed for the operation of MEMS thermal shear-stress sensors.

This work presents a systematic study including both experimental and theoretical investigations to address this issue. By wind-tunnel testing of various designs of MEMS shear-stress sensors with different membrane thickness, size and material, we first obtain experimental data that confirms the inadequacy of the classical hot-wire/hot-film theory. A more thorough theoretical analysis follows, identifying that this invalidity is due to the lack of a thin thermal boundary layer in the flow. Then, a 2-D MEMS shear-stress sensor

theory is developed. We show that by incorporating important heat transfer effects that are ignored by the classical theory, the new model provides a closed-form approximate solution and consistently describes all the MEMS sensors. Moreover, we also present 3D heat transfer simulation that agrees with experiments and supports our new 2-D model. This work demonstrates that many classical assumptions made for conventional thermal devices need to be carefully re-examined for miniature MEMS devices.

MEMS THERMAL SHEAR-STRESS SENSORS

Fig. 1 illustrates one half of a MEMS thermal shear-stress sensor cut along the flow. During operation, a boron-doped polysilicon thermistor (used as a hot wire) is heated electrically and maintained at a constant temperature. Since heat transfer rate from the hot wire to the fluid flow is related to the wall shear stress, the power consumption of the hot wire can then be used to determine the shear stress. The hot wire lies on a membrane (typically made of silicon nitride or Parylene), which is in turn suspended over the substrate by a vacuum or air cavity. By improving thermal isolation, the cavity increases sensor sensitivity while lowering power consumption. The detailed fabrication process of these sensors is available in [2,5].
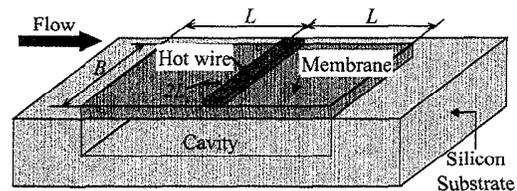
Fig. 1. MEMS shear-stress sensor structure.

EXPERIMENTS

We have conducted wind-tunnel experiments with four different MEMS shear-stress sensors. Three of the sensors (Fig. 2(a)) have a polysilicon hot wire on a silicon-nitride membrane. These sensors have membrane thickness and hot-wire width of 1.5 μm and 7 μm (the "thin-nitride" sensor), 3 μm and 7 μm ("thick-nitride"), and 1.5 μm and 3 μm ("thin-wire"), respectively [2-4]. In the fourth sensor (the "Parylene" sensor), the hot wire (width = 10 μm) is on a Parylene membrane (Fig. 2(b)) [5]. The polysilicon hot
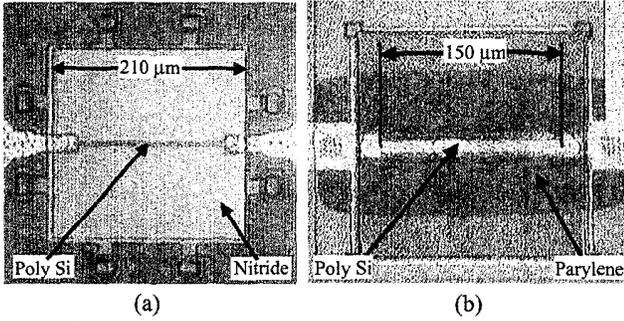
Fig. 2. MEMS shear stress sensor using (a) nitride and (b) Parylene membranes.
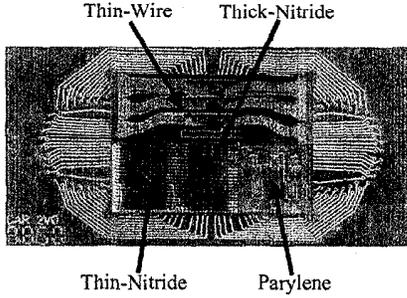
Fig. 3. Four MEMS sensors on the same PCB.

wires in all four sensors have a thickness of 0.5 μm. These MEMS sensors are carefully placed on a single PC Board (Fig. 3) to ensure uniform testing conditions and consistent testing results. The PCB is then flush-mounted on the inner wall of a 2-D wind-tunnel [3].

The wind-tunnel testing results are shown in Fig. 4. During sensor operation, the power (denoted P) needed to maintain the hot wire at a constant temperature depends on the wall shear stress (denoted τ). The classical theory states $P-P_0 \propto \tau^{1/3}$, where $P_0 = P|_{\tau=0}$. However, the significantly curved lines in Fig. 4 clearly indicate that this is not true.
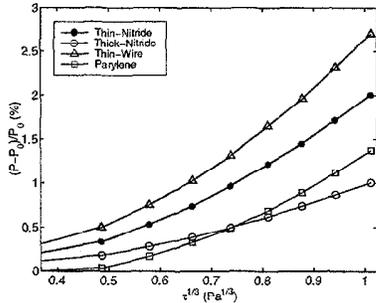
Fig. 4. Measured shear-induced power change vs. $\tau^{1/3}$.

INADEQUACY OF THE CLASSICAL THEORY

To explain the inadequacy of the classical theory, we now carefully examine the classical shear-stress sensor model. This model (Fig. 5) considers a fluid flowing past a plane surface. The temperature at the surface is ambient
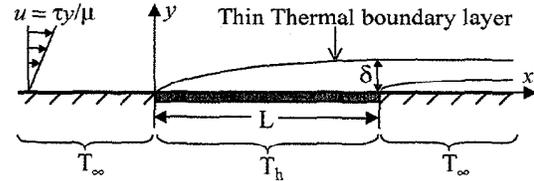
Fig. 5. Classical shear-stress sensor model.

(T_∞) except over a heated length L, where the temperature is T_h. Assume $\delta \ll L$, i.e., the thermal boundary layer (the heated region in the fluid) is very thin compared to the heated length. The heat transfer in the flow is then governed by

$$u\frac{\partial T}{\partial x} = \alpha \frac{\partial^2 T}{\partial y^2}, \quad (1)$$

where T is the fluid temperature and α the fluid's thermal diffusivity. Note that constant material properties are assumed throughout this paper. Under the assumption that the thermal boundary layer is also thin compared with the velocity boundary layer [7], the flow velocity u is linear, i.e.,

$$u = \frac{\tau y}{\mu}, \quad (2)$$

where μ is the fluid's dynamic viscosity. Solving this problem yields the classical result that the heat transferred to the fluid over the heated length is proportional to $\tau^{1/3}$. To study the validity of the classical theory, we note that with $\bar{\tau} = \tau L^2 / \mu\alpha$, the thermal boundary layer thickness is given by $\delta/L = 0.34\bar{\tau}^{-1/3}$. Thus, for the assumption $\delta \ll L$, and hence the classical theory, to hold, one must have

$$\bar{\tau}^{1/3} \gg 1. \quad (3)$$

For the MEMS sensors used in our experiment, we can readily calculate that $\bar{\tau}^{1/3} \leq 2.8$ for the three nitride-based sensors, and $\bar{\tau}^{1/3} \leq 1.6$ for the Parylene-based sensor. Thus, for our MEMS sensors, the condition (3) is readily violated and the classical theory is hence invalid.

A NEW THEORY FOR MEMS SENSORS

Given the failure of the classical theory, a new theory needs to be developed to consistently describe MEMS sensor operation. Clearly, the new theory must allow a thick thermal boundary layer ($\delta \sim L$). Heat conduction in the membrane is also important, since it can be shown that the abrupt change in the prescribed surface temperature, as assumed in the classical model, is ill-posed in the absence of a thin thermal boundary layer. We use the same notation for fluid properties as used in the classical model.

Our MEMS sensor model is shown in Fig. 6. In this 2D problem, we consider forced convection in the fluid coupled with heat conduction in the membrane (length $= 2L$). The membrane's mid-point is maintained at constant temperature T_h, representing the hot wire. This essentially ignores the hot wire's streamwise width, which is much

smaller than L. Forced convection in the fluid is then governed by [7]

$$u\frac{\partial T}{\partial x} = \alpha\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2}\right). \quad (4)$$
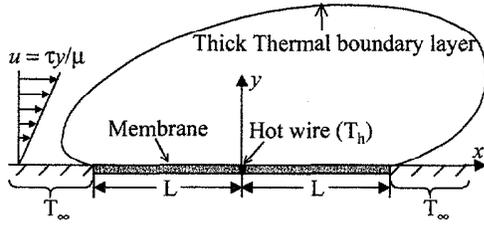
Fig. 6. MEMS shear-stress sensor model.

Comparing Equations (1) and (4), we see that the classical theory ignores the contribution from streamwise heat conduction, which is however very important for MEMS sensors due to the lack of a thin thermal boundary layer. On the other hand, the flow velocity profile (2) is still valid due to the small sensor size.

Heat conduction in the membrane is governed by

$$k_m t \frac{\partial^2 T_m}{\partial x^2} + k\frac{\partial T}{\partial y}\Big|_{y=0} = 0, \quad (5)$$

where T_m is the membrane temperature averaged over the thickness t ($t \ll L$), and k_m is the membrane's thermal conductivity. The coupled heat transfer problem is closed by the following boundary conditions: $T = T_h$ for $x = 0$ and $y = 0$; $T = T_\infty$ for $|x| > L$ and $y = 0$ as well as for $x \to \pm\infty$ or $y \to +\infty$; and $T_m = T$ as $|x| \le L$ and $y = 0$.

While the solution to this model is generally not in closed form, its functional form can be identified. Define three dimensionless parameters by

$$\bar{\tau} = \frac{\tau L^2}{\mu\alpha}, \quad \bar{P} = \frac{P}{2k(T_0 - T_\infty)B}, \quad \text{and} \quad \lambda = \frac{k_m t}{kL}, \quad (6)$$

where B is the sensor's transverse length (Fig. 1). It can be shown that the solution to the MEMS sensor model is given by the dimensionless power $\bar{P}$ as a function of the thermal conductivity ratio λ and the dimensionless shear stress $\bar{\tau}$. That is, the solution has the functional form

$$\bar{P} = \bar{P}(\lambda, \bar{\tau}) = \bar{P}_0(\lambda) + \Delta\bar{P}(\lambda, \bar{\tau}),$$

where we have decomposed the dimensionless power into two components: the power in still fluid $\bar{P}_0 = \bar{P}(\lambda, 0)$, and

the shear-induced power change $\Delta\bar{P} = \bar{P} - \bar{P}_0$.

Thus, the heat transfer problem can be solved numerically in terms of these dimensionless parameters, as shown in Fig. 7. Furthermore, this numerical solution can be approximated by the following practically useful closed-form expressions. For $0.04 \le \lambda \le 5$ and $1 \le \bar{\tau} \le 50$, within 5% error we have

$$\bar{P}_0 = 1.96\lambda(1 + 0.435/\lambda^{0.9})$$

and

$$\Delta\bar{P} = \frac{1.31\lambda^{0.4}\bar{\tau}^{1.47}}{(1 + 2.07\lambda^{0.6})(1 + 0.56\bar{\tau}^{0.3}/\lambda^{0.037})^5}.$$

To study the sensors used in our experiments, we can alternatively approximate the numerical solution by

$$\Delta\bar{P} = [0.129\lambda^{0.45}/(1 + 1.81\lambda^{0.55})]\bar{\tau}^{0.85}$$

for $0.6 \le \bar{\tau} \le 5$, and

$$\Delta\bar{P} = [0.187\lambda^{0.5}/(1 + 2.1\lambda^{0.59})]\bar{\tau}^{0.67}$$

for $2.5 \le \bar{\tau} \le 25$. In both cases, the applicable range of λ is still $0.04 \le \lambda \le 5$ and the approximation error is within 10%.

(a)   (b)
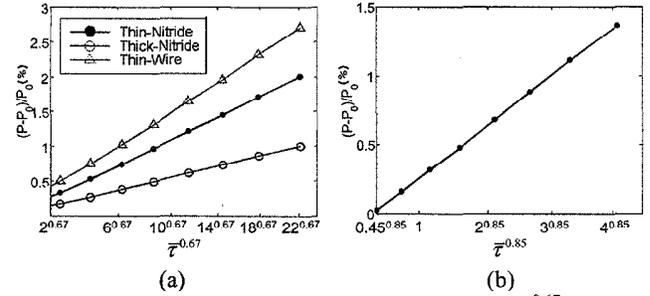
Fig. 8. Measured shear-induced power vs. (a) $\tau^{0.67}$ for the nitride sensors ($\tau = 1$ Pa gives $\bar{\tau} \approx 22$), and (b) $\tau^{0.85}$ for the Parylene sensor ($\tau = 1$ Pa gives $\bar{\tau} \approx 4$).

We now plot our experimental data in the appropriate range of $\bar{\tau}$ in Fig. 8. It can be seen that the nitride sensors, and the Parylene sensor, indeed approximately follow the 0.67th- and 0.85th-power laws, respectively. This confirms that our 2D theory correctly predicts the trend of MEMS sensor operation.

THREE-DIMENSIONAL MODELING

Practical MEMS sensors often exhibit significant 3D effects. To obtain quantitative information on such sensors,

(a)   (b)

Fig. 7. Numerically obtained dimensionless (a) power in still fluid and (b) shear-induced power change.
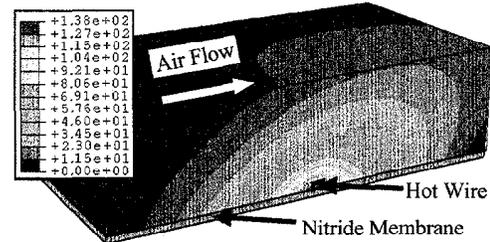
Fig. 9. Temperature distribution in the air flow past the thin-nitride sensor operating in constant-tempeature mode with $T_h - T_\infty = 100°C$. ($\tau = 1$ Pa and $P \approx 5.4$ mW.)
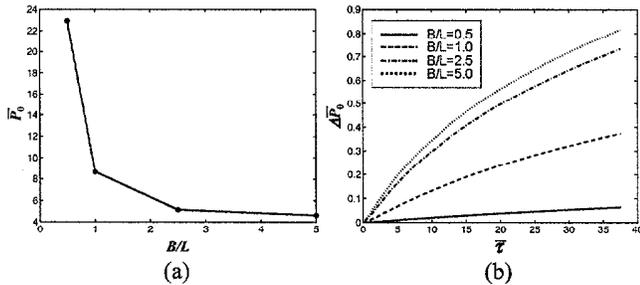
Fig. 10. (a) Power in still air $\overline{P}_0$ and (b) Shear-induced power vs. $\overline{\tau}$ for MEMS sensors for different B/L.
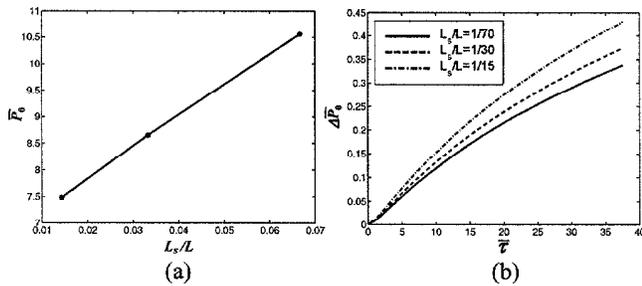
Fig. 11. (a) Power in still air $\overline{P}_0$ and (b) Shear-induced power vs. $\overline{\tau}$ for MEMS sensors for different L_s/L.

we perform 3D coupled conduction-convection simulation using the ABAQUS finite element analysis package.

Fig. 9 shows a typical calculated fluid temperature distribution, which clearly indicates that a thin thermal boundary layer does not exist and therefore the classical theory is inappropriate. Figs. 10 and 11 depict the effects of the geometric parameters B/L and L_s/L with material properties fixed. Shear stress and sensor power are still nondimensionalized by (6). We see that as the sensor depth B is increased, the dimensionless power consumption in still fluid decreases while the shear-induced power increases. This is consistent with reduced 3D effects for increasing B. On the other hand, as the hot wire width increases, both still-fluid and shear-induced powers increase. This is due to the increased contact area between the hot wire and the fluid.

Finally, Fig. 12 compares modeling and experimental
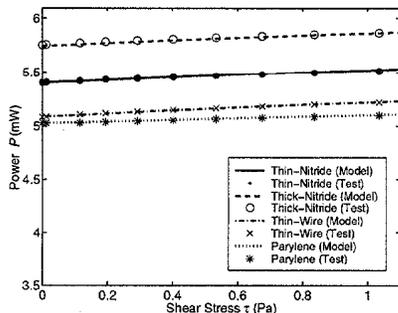
Fig. 12. Comparison of simulation and experimental results (the power for the thick nitride sensor was shifted down by 5.5 mW for convenience).

results for our four MEMS sensors (see the section Experiments). We allow the thermal conductivity of silicon nitride and the height of the Parylene-covered cavity (20 to 40 μm) to be adjusted in the simulation. It can be seen that the agreement between numerical and testing results is excellent. This confirms that our 3D model provides valid quantitative description of practical MEMS sensors.

CONCLUSION

We conducted an experimental and theoretical analysis of MEMS thermal shear-stress sensors. The classical theory for macroscale thermal shear-stress sensors was shown to be invalid. A 2D theory was developed to correctly predict the trend of MEMS sensor operation, and 3D simulation has been performed to provide quantitative description of MEMS sensors. The 2D and 3D models were both compared with experimental data, and yielded excellent qualitative and quantitative agreement.

ACKNOWLEDGMENTS

REFERENCES

[1] M. Kimura, S. Tung, J. Lew, C.M. Ho, F. Jiang and Y.C. Tai, "Measurements of Wall Shear Stress of a Turbulent Boundary Layer Using a Micro-Shear-Stress Imaging Chip", *Fluid Dynamics Res.*, Vol. 24, pp. 329-342, 1999.

[2] C. Liu, Y.C. Tai, J.B. Huang and C.M. Ho., "Surface Micromachined Thermal Shear Stress Sensor", *1994 ASME International Mechanical Engineering Congress and Exposition*, Chicago, IL, pp. 9-15, 1994.

[3] F. Jiang, Y.C. Tai, B. Gupta, R. Goodman, S. Tung, J.B. Huang and C.M. Ho, "A Micromachined Shear Stress Sensor Imager", *MEMS'96*, pp. 110-115, 1996.

[4] F. Jiang, Y.C. Tai, K. Walsh, T. Tsao, G.B. Lee and C.M. Ho, "A Flexible MEMS Technology and Its First Application to Shear Stress Sensor Skin," *MEMS '97*, pp. 465-470, 1997.

[5] X.Q. Wang, Z. Han, F. Jiang, T. Tsao, Q. Lin, Y.C. Tai, V. Koosh, R. Goodman, J. Lew, C.M. Ho, "A Fully Integrated Shear Stress Sensor", *Transducers'99*, pp. 1074-1077, 1999.

[6] B.J. Bellhouse and D.L. Schultz, "Determination of Mean and Dynamic Skin Friction, Separation and Transition in Low-Speed Flow with a Thin-Film Heated Element", *J. Fluid Mechanics*, Vol. 24, pp. 379-400, 1966.

[7] F.M. White, *Viscous Fluid Flow*, 2nd ed., McGraw-Hill, 1991.

SPUTTERED SILICON FOR INTEGRATED MEMS APPLICATIONS

Kenneth A. Honer and Gregory T. A. Kovacs
CIS-218X Stanford University
Stanford, CA 94305-4075

ABSTRACT

This paper describes a new fabrication process for creating electrostatic microstructures that are compatible with prefabricated aluminum-metallized CMOS circuitry. The process uses sputtered silicon to make released microstructures similar to those commonly made using LPCVD polysilicon but does so at much lower temperatures (~350 °C). The low-temperature nature of sputter deposition makes it possible to use polyimide sacrificial layers that can be released in an oxygen plasma. This dry-release process eliminates the need for critical point drying or similar methods. Average strain gradients in released sputtered silicon cantilevers were found to vary with the inverse square of thickness. At a thickness of 5.0 μm, the radius of curvature of released cantilevers was in excess of 80 mm. Improvements in the electrical conductivity of completed structures were realized by cladding the sputtered silicon structural layers in symmetric, 50 nm thick layers of titanium-tungsten. Underlying CMOS transistors showed no more than a 3% increase in their maximum saturation current after mechanical layer processing. As a demonstration of the integratibility of the sputtered silicon process, electrostatically actuated variable-capacitors were fabricated above CMOS capacitance detection circuitry.

INTRODUCTION

Integration is vital to the performance of many types of microsensors, since as dimensions decrease, sensitivity often falls off precipitously. The sensitivity of a torsional capacitive accelerometer, for example, scales as the fifth power of the lateral dimension. Large arrays of transducers also benefit from integration since the number of bondpads and the external circuitry needed to control each device can be reduced through the use of on-chip multiplexers.

LPCVD polysilicon has been the material of choice for most surface micromachined structures since the early 1980's [1] and has been integrated with on-chip circuitry for nearly as long [2]. Unfortunately, the high deposition and annealing temperatures of LPCVD polysilicon (~580-630 °C and >900 °C, respectively), have required either depositing metal layers after the polysilicon structural layers have been deposited in etched pits [3] or using refractory metals such as tungsten instead of aluminum for the circuits [4]. Both approaches increase the overall fabrication complexity and may require re-engineering the CMOS process to accommodate the new thermal budget, metal layers, and/or lithography changes.

In contrast, sputtered silicon can be used to fabricate released microstructures atop standard, aluminum-metallized CMOS at temperatures below 350 °C. By avoiding the temperature extremes of LPCVD polysilicon, greater flexibility is allowed in creating or selecting a circuit process, resulting in potentially significant cost-savings and performance advantages.

Most LPCVD polysilicon processes have used oxide sacrificial layers and wet-releases based hydroflouric acid (HF). Previous high temperature work on sputtered silicon microstructures [5] and our own low-temperature process [6] have followed in a similar vein. Figure 1 shows a sputtered silicon comb fabricated using an oxide sacrificial layer.
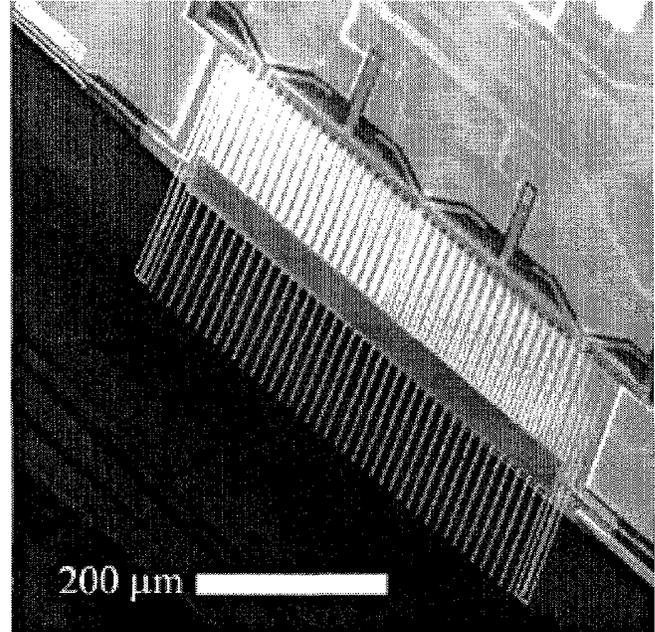
Figure 1: A comb structure fabricated in sputtered silicon atop silicon dioxide. The wafer was cleaved beneath the structure after it was released in buffered HF.

Investigating oxide sacrificial layers has led to an interesting discovery. HF porosity in sputtered silicon is evident at thicknesses far in excess of those reported for LPCVD polysilicon [7], making it a potential encapsulating material as well as a structural material. Figure 2 shows a buried cavity formed in an SiO_2 sacrificial layer underneath a 2.0 μm sputtered silicon layer.

One issue of concern is that the attack of the sputtered silicon during exposure to HF may weaken the film structurally. This is in addition to the problems associated with protecting the underlying CMOS during the release and capillary forces that can cause the structure to stick to the substrate if critical point drying [8] or self-assembling monolayers (SAMS) [9] are not used.

However, because sputtered films avoid the temperature excursions of LPCVD polysilicon, they are not only compatible with conventional oxide sacrificial layers, but also with organic sacrificial layers, such as polyimide. These organic layers can be removed in an oxygen plasma, thus avoiding the stiction and selectivity problems associated with wet releases and alleviating the concerns over chemical attack in HF.
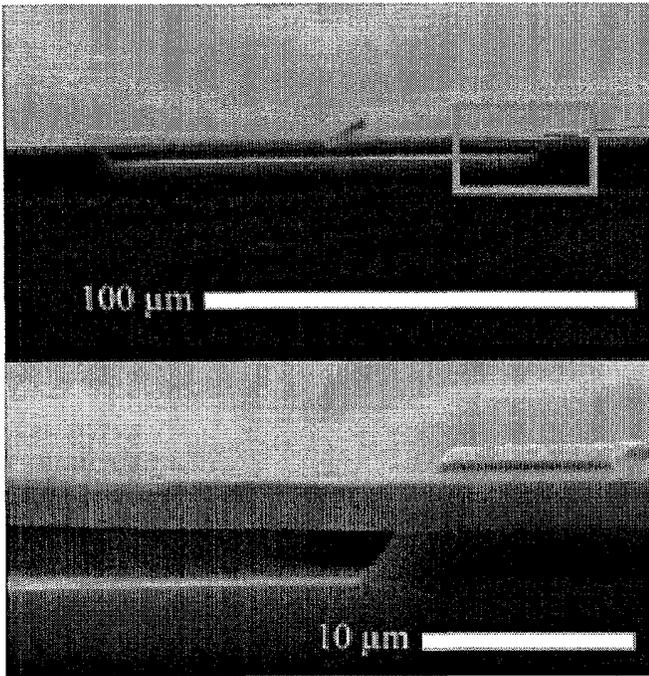
Figure 2: SEM image of a cleaved cavity formed in SiO₂ by selectively masking the top sputtered silicon layer with resist and exposing it to 20:1 buffered HF for approximately 1 hour.

anisotropy of this etch was increased by reducing the pressure to the minimum practical limit of the plasma etcher (in this case, 100 mTorr). The photoresist used to pattern the aluminum mask was intentionally left in place to provide a more uniform loading during the etch. Care must be taken to avoid overetching the polyimide at this step, since once the resist and the polyimide in bottom of the vias is cleared away, the lateral etch rate increases significantly due to changes in loading.

Figure 3: Dry release process flow incorporating conductive TiW cladding: Lower metallization (a); Polyimide sacrificial layer and patterned aluminum hardmask (b); Etched polyimide layer (c); Structural layer (TiW, Si, TiW tri-layer shown) (d); Oxygen plasma release (e).

DRY-RELEASED FABRICATION SEQUENCE

The basic process flow is illustrated in Figure 3 and begins with completed and tested CMOS circuitry. For convenience, the second metal layer of the underlying CMOS was used as the lower electrode layer for the mechanical structures. This electrode layer consisted of a bi-layer of 400 nm of aluminum and 100 nm of titanium. The titanium capping layer serves to minimize hillocking during the sacrificial polyimide cure cycle.

After patterning the lower electrode layer (Figure 3-a) and opening contact holes in the titanium over the bondpads, an adhesion promoter (*Dupont PIQ-13 coupler, Willmington, DE.*) was spun on and cured at 350°C for 30 minutes. Dupont PIQ-L100 polyimide resin was chosen for the sacrificial layer because of its low thermal expansion coefficient (3ppm/K) and clean release characteristics. This is in contrast to some other polyimides which, as a result of incorporating adhesion promoters within the resin, often leave behind a visible residue when etched in oxygen plasma. The PIQ-L100 resin was spun on for 30 s at 2200 rpm and cured at 350 °C for 1 hr. These spin parameters produced a polyimide film that was approximately 2.5 μm thick. Thicknesses were established using an optical measurement system (*Nanometrics, Sunnyvale, CA.*) which was calibrated relative to the results from a profilometer (*Veeco, Plainview, NY.*).

The polyimide was then etched back to 2.0 μm in oxygen plasma and coated with 200 nm of aluminum, which served as a hard mask. Etching back the polyimide both improved the adhesion of the aluminum mask and provided more precise control of the sacrificial layer thickness.

After patterning the aluminum mask (Figure 3-b), contact vias were etched in the polyimide using an oxygen plasma. The

During the course of the sacrificial via etch, the photoresist mask was etched away, leaving the aluminum mask exposed to the oxygen plasma. The oxidized aluminum mask was then removed by immersing the wafers in 50:1 HF for 10 s followed by a wet aluminum etch (Figure 3-c). This HF dip also helped remove any oxidized titanium from the bottom of the vias.

In order to improve the electrical contacts, the wafers were back sputtered in an argon plasma immediately prior to depositing the structural layer. This structural layer consisted of either 0.6 – 5.0 μm of boron-doped sputtered silicon or a tri-layer of 50 nm of TiW, followed by 2.0 μm of doped sputtered silicon and a second, symmetric, layer of TiW. By making the TiW layers symmetric, thermal bi-morph effects were minimized. Stresses in the films were controlled by varying the power and working gas pressure of the deposition. Films with stress values lower than 100 MPa were routinely obtained.

The TiW layers decrease the electrical resistivity of the resultant film significantly, as can be seen from Table 1.

TABLE 1: Comparison of electrical resistivity

Technology	Resistivity
Unannealed sputtered Si	6-50 x 10⁶ Ω/□
Annealed sputtered Si	7-50 x 10⁹ Ω/□
Sputtered TiW-Si-TiW	25 Ω/□

The wafers were annealed at 300 °C for 1 hr in a dry nitrogen ambient to sinter the contacts between the mechanical layer and the lower electrode and to stabilize the stresses in the structural layers.

The structural layers were patterned using an SF_6-C_2ClF_5 plasma (Figure 3-d). This chemistry etches both TiW and silicon

so there is no need for separate plasma steps when etching the tri-layer.

Following this etch, the wafers were diced and the individual chips were release in an oxygen plasma. The release etch took approximately 4 hrs to undercut 15 µm at an input power level of 200W and a pressure of 3 Torr (Figure 3-e). Release times could be decreased by using a more aggressive plasma etcher, but circuit damage may become a concern.

STRAIN GRADIENTS

Control of warping in released structures is a major concern during microstructure fabrication. This warping is caused by the presence of residual strain gradients throughout the thickness of the structural layer. The sputtered silicon in this work was not completely free of these strain gradients, as evidenced by the curvature of the released cantilevers in Figure 4. However, the gradients were observed to be a strong function of film thickness, with thicker films having smaller average values for the gradient. At a 5.0 µm film thickness, the average radius of curvature of a cantilever beam was in excess of 80 mm.
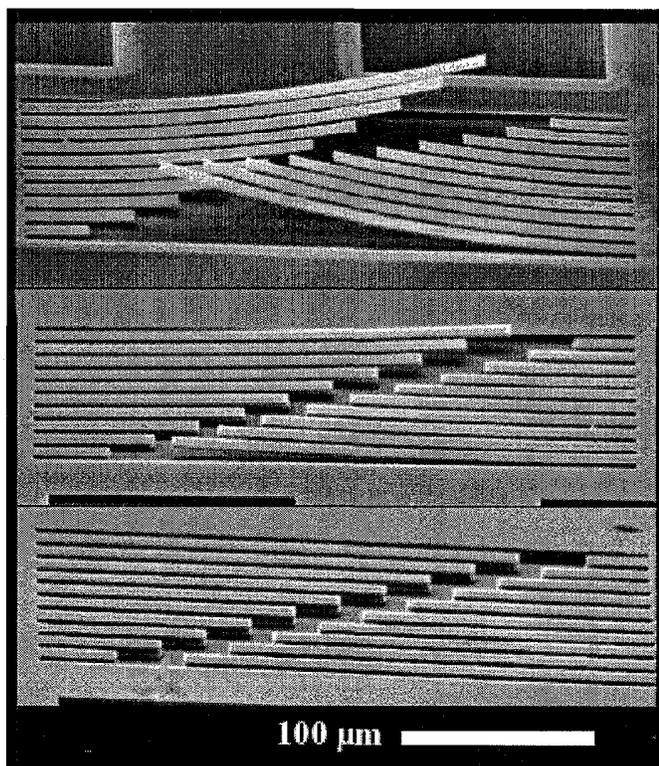
Figure 4: Dry-released pure sputtered silicon cantilever beams. The top set of beams is 0.6 µm thick. The center set is 2.0 µm thick. The bottom set is 5.0 µm thick and has the least curvature. All beams were released in a 1.3 W/cm² oxygen plasma.

Figure 5 shows data from interferometer scans of six cantilever arrays which appears to match a theoretical curve based on surface strains as the dominant contributor to the gradient. The moment caused by these proposed surface strains is proportional the thickness of the film and acts against the

stiffness of the cantilevers which is proportional to the cube of thickness. The radius of curvature is thus proportional to the inverse square of the thickness in this model.
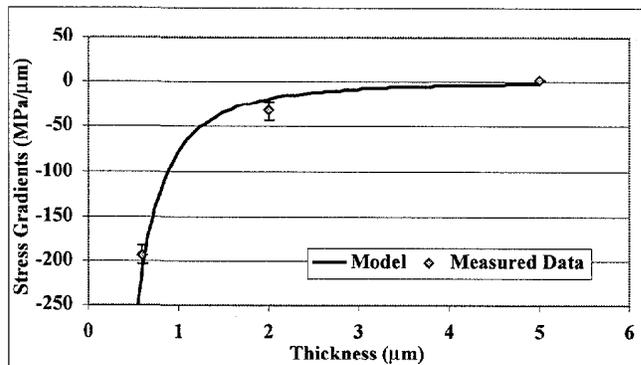
Figure 5: Sputtered silicon stress gradient as a function of thickness for dry-released cantilever beams Each data point represents typical beams from six arrays. Error bars represent two standard deviations. The solid line represents a theoretical curve based on surface effects, which are postulated to dominate the stress gradients in the films.

The TiW layers used to increase the conductivity of the sputtered silicon can also be used to adjust the curvature of the released structures. By adjusting the stress in the top layer to be more compressive, the released structures can be made more convex. Similarly, by making the stress more tensile, the structures can be made more concave.

CMOS COMPATIBILTY

As expected, the post-CMOS fabrication steps have little measurable effect on the underlying CMOS circuitry. Figure 6 shows transistor curves from transistors fabricated at the Stanford University Center for Integrated Systems before and after the mechanical layer processing. Out of the 30 NMOS and PMOS transistors measured, none showed more than a 3% change in maximum saturation current.
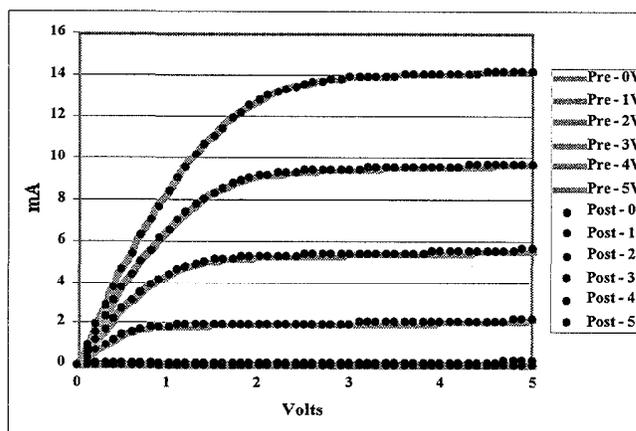
Figure 6: NMOS transistor curves before and after mechanical layer processing. Not more than a 3% increase in maximum saturation current was observed in any of the 30 nmos or pmos transitors tested.

Some of this change can be attributed to sintering of the contacts during the curing and annealing steps since no barrier layers were used between the silicon and the Al-1%Si metallization layers.

INTEGRATED VARIABLE CAPACITOR

As a demonstration of the process, TiW-clad sputtered silicon was used to create variable capacitors with integrated capacitance detection circuitry. This type of device was chosen because most integrated surface micromachined sensors operate using some variation on capacitance detection. Figure 7 shows an optical scan of one such device, an array of 100 μm plates, each suspended by four springs. A simple buffer circuit, with a reset switch to control charging, was used to measure changes in capacitance caused by electrostatically deflecting one set of capacitors. For small deflections, the relationship between deflection, capacitance change, and the output voltage should be linear.

(a)   (b)
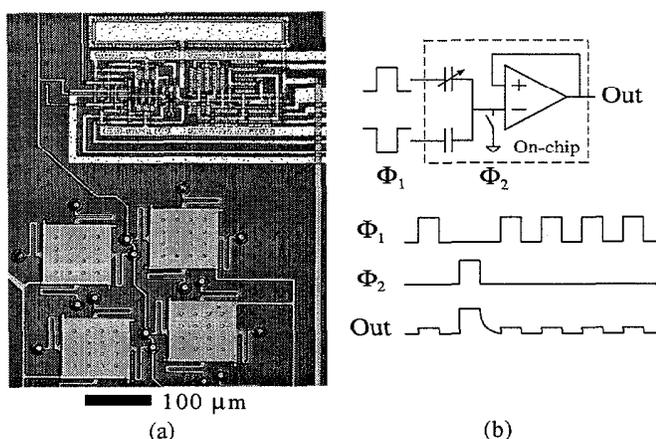
Figure 7: *Variable capacitors with integrated detection circuitry. (a) An optical image of the two reference and two variable capacitors next to the capacitance detection circuit. (b) Schematic representation of the capacitance detection scheme.*
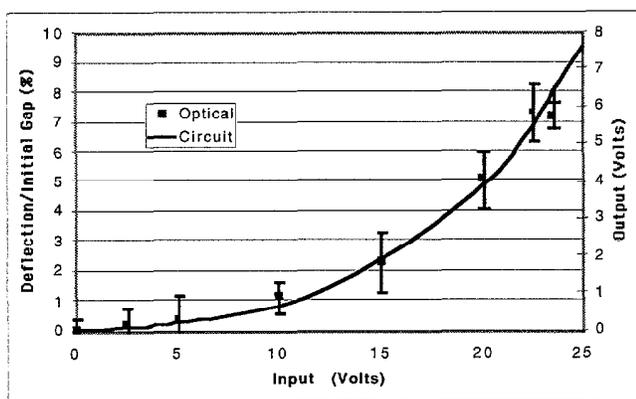
Figure 8: *Deflection as a function of voltage for a 100 μm square variable capacitor. The solid line shows the output of on-chip capacitance detection circuitry. The discrete points show the deflection as measured by six scans of an interferometric microscope. Error bars represent the standard deviation.*

Figure 8 shows a graph of the amplified output of the capacitance detection circuit and the deflection of the plates, as simultaneously measured by an interferometer. The measured deflection and circuit output correspond to each other, indicating that the circuit is transducing as expected.

CONCLUSIONS

Sputtered silicon is a promising material for integrated microstructures. When clad in TiW, the resistivity is only 25 $\Omega/\square$. The radius of curvature of released structures depends on the film thickness and can be made lower by using thicker films or by tailoring the stress in the TiW cladding layers. Curvature radii in excess of 80 mm were demonstrated with 5.0 μm thick samples. The process does not adversely affect threshold voltage or saturation currents of underlying CMOS transistors. Because the sacrificial layer is removed in an oxygen plasma, there are none of the stiction problems associated with wet-releases. Finally, the TiW-clad sputtered silicon process can be used to make many of the same types of devices currently made using LPCVD polysilicon.

ACKNOWLEDGEMENTS

REFERENCES

[1] R. T. Howe and R. S. Muller, Abstract No. 118, *Extended Abstracts of the Electrochemical Society Meeting*, Montreal, Canada, May, 1982, pp. 184-185

[2] R. T. Howe and R. S. Muller, "Integrated Resonant-Microbridge Vapor Sensor," *Proc. IEEE Int. Elect. Dev. Mtg.*, San Francisco, Dec., 1984, pp. 213-216

[3] J. H. Smith, et al., "Embedded micromechanical devices for the monolithic integration of MEMS with CMOS," *Proc. Int. Electron Devices Mtg.*, Washington, DC, Dec., 1995, pp. 609-612

[4] J.M. Bustillo, et al., "Process technology for modular integration of CMOS and polysilicon microstructures," *Microsystem Technologies*, Vol. 1, 1994, pp. 30-41

[5] T. Abe and M. L. Reed, "Low Strain Sputtered Polysilicon for Micromechanical Structures," *Proc. of Ninth International Workshop on Micro Electro-Mechanical Systems*, San Diego, Feb., 1996, pp. 258-262

[6] K. A. Honer and G. T. A. Kovacs, "Sputtered Silicon for Microstructures and Microcavities" *International Mechanical Engineering Congress and Exposition - MEMS*, Nashville, Nov., 1999

[7] K. S. Lebouitz, R. T. Howe, and A. P. Pisano, "Permeable Polysilicon Etch-Access Windows for Microshell Fabrication," *Transducers '95*, pp. 224-227

[8] G. T. Mulhern, D. S. Soane, and R. T. Howe, "Supercritical carbon dioxide drying of microstructures", *Transducers '93*, pp. 296-299

[9] M. R. Houston, et al., "Self-assembled monolayer films as durable anti-stiction coatings for polysilicon microstructures" *Technical Digest Solid State Sensors and Actuator Workshop*, Hilton Head, Jun., 1996, pp. 42-47

TEMPERATURE COEFFICIENTS OF MATERIAL PROPERTIES USING DIFFERENTIAL CAPACITIVE STRAIN SENSORS

Larry L. Chu, Long Que, and Yogesh B. Gianchandani[*]

Department of Electrical and Computer Engineering, University of Wisconsin, Madison

ABSTRACT

This paper demonstrates the use of micromachined differential capacitive strain sensors for determining the temperature dependence of the thermal expansion coefficient, Young's modulus, and residual strain of microstructure materials. The device and methodology presented are suitable for both tensile and compressive materials and can be used before or after packaging the device. It is shown by this method that for a particular set of fabrication conditions for electroplated nickel devices, the thermal expansion coefficient is 11-13 ppm/K over 23-50°C; the residual strain changes from neutral to −880 microstrain over 23-100°C; the Young's modulus has a temperature coefficient of −1590 ppm/K with an estimated value of 35 GPa at room temperature; consequently, the residual stress changes from neutral to −29 MPa over 23-100°C. Using passive bent beam strain sensors, the thermal expansion coefficient is found to increase from 13.5-16.5 ppm/K over 50-150°C.

I. INTRODUCTION

Monitoring the mechanical properties of structural materials is a critical challenge in MEMS research and manufacturing. Device performance parameters are sensitive to variations in Young's modulus, residual strain, residual stress, and the thermal expansion coefficient of structural materials. Unfortunately, these properties can vary with manufacturing, packaging, and deployment conditions. Diagnostic microstructures for rapid characterization of these material properties are useful for monitoring processes. Additionally, if these structures provide electronic readout, they can be integrated within device packages, improving overall performance by offering post-deployment calibration capability.

In past efforts to electronically determine material properties, electrostatic forces have been used to cause out-of-plane deflections of cantilevers, beams, and plates [1-4]. However, a number of sensors and actuators are based on lateral deflections that occur parallel to the plane of the substrate. A laterally deflecting test device is preferred for characterizing the materials used in these devices, which may be anisotropic and may exhibit different properties for in-plane and out-of-plane deflections. The differential capacitive strain sensor [5] is a suitable tool for such measurements. It provides the added benefit of a wide dynamic range, being able to accommodate not only tensile stresses, but also compressive stresses that are exploited in certain actuator applications [6].

In the differential capacitive strain sensor, a series of ribs, suspended above the substrate by bent-beams, support interdigitated tines, forming sidewall capacitors (Fig.1). The suspensions and tines are arranged such that the capacitance on one side of a rib increases as the other decreases. The residual strain in the structural material is determined from the differential capacitance, defined as ΔC =

$C_{AB} - C_{AC}$, which is >0 in tension, and <0 in compression. The Young's modulus of the material is determined from the curvature of the plot of C_{AB} (or equivalently, C_{AC}) versus an applied bias voltage between elements A and B (or C). The bias voltage applies an electrostatic force, resulting in a displacement that is measured by the change in capacitance. The displacement is related to the stiffness of the suspension, which is proportional to the Young's modulus. The curves obtained are referred to as CV curves.

In this effort, a procedure is presented for determining the temperature dependence of the Young's modulus, residual strain, residual stress, and expansion coefficient of microstructure materials, using differential capacitive strain sensors. It is demonstrated using electroplated metal strain sensors fabricated on a silicon substrate.
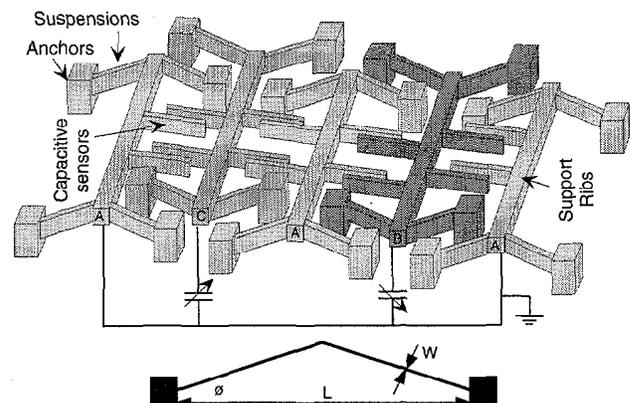
Fig. 1: Schematic of the differential capacitive strain sensor and definition of bent-beam dimensions.

II. MODELING

The modeling approach was conceptually similar to those described in [2-4]. The capacitance between tines was numerically modeled using FastCap™ [7]. The capacitance was determined for a range of separations between tines, and stored in a look-up table. The net mechanical spring constant of the bent-beam suspensions was calculated as noted in [6]. Starting with the zero-bias gap, spacing between the tines was decremented in small intervals. At each position, the capacitance was determined from the look-up table, while the corresponding bias voltage was determined from force balance by equating the electrostatic attractive force to the mechanical restoring force:

$$k_{eff} y = \frac{V_{Bias}^2}{2} \frac{\partial C_{AB}}{\partial y} \tag{1}$$

where k_{eff} is the effective mechanical spring constant of the system, and y is the in-plane displacement from the zero-bias position. Figure 2 shows calculated CV curves obtained by this method, assuming L=198 μm, W= 5 μm,

[*]*Corresponding author:* 1415 Engineering Drive, Madison, WI 53706-1691; Tel: (608) 262-2233; Fax: 262-1267, E-mail: yogesh@engr.wisc.edu

ø=0.1 rad., 4.4 μm thickness, and 10 pairs of tines exist with 167 μm interdigitated length. The sidewalls of the tines were assumed to have a re-entrant profile that was 22° off vertical for reasons described in the following section. The zero-bias gap between the upper edges of adjacent tines (where the tines were closest) was 1.85 μm. The tines were assumed to be 2.2 μm above the substrate.

The effect of certain structural non-idealities on the CV curves is shown in Figs. 3-5. Figure 3 shows variations in the CV curve caused by changing the gap between the tines. The device dimensions are as for Fig. 2, the Young's modulus is 100 GPa, and the gap between the upper edges of the tines is varied from 1.65 μm to 2.05 μm in 0.1 μm increments. Figure 4 shows the impact of changing the sidewall angles of the tines, with the Young's modulus at 100 GPa, and the remaining parameters as for Fig. 2. Figure 5 shows the impact of upward and downward deflections of one tine in each pair. Clearly, each of these variables can have an impact on the CV curves.

III. EXPERIMENTAL RESULTS

The strain sensors were fabricated on silicon wafers insulated with 1 μm thick thermal oxide and 0.5 μm thick LPCVD nitride. A 2 μm thick sputtered Ti sacrificial layer was patterned and covered with a Cr/Ni seed layer. The devices were electroplated into a photoresist mold from a nickel sulphamate solution (which contained trace impurities.) At 54°C temperature, using 5-10 mA/cm² current density, a thickness of 4.4 μm was achieved in 9.5 min. The sidewalls of the plated structures were 22° off vertical due to resist reflow during hard bake. The photoresist mold was subsequently stripped and the sacrificial material etched away. Following this, the sample was coated with self-assembled monolayers using ODS [5]. A SEM image of a fabricated structure is shown in Fig. 6.

The coefficient of thermal expansion for the structural metal was first measured by passive bent-beam strain sensors [8] which were fabricated on the same die as the differential capacitive stain sensors. The strain was measured as a function of temperature by observing their deformation through a microscope. The dimensions of these devices were L= 198 μm, W= 6 μm, and bending angle ø= 0.1 rad. The expansion coefficient of structural metal can be written as:

$$\alpha_{Metal}(T) = \alpha_{Si}(T) - \frac{d(e(T))}{dT} \qquad (2)$$

where e(T) is the strain observed by the strain sensor. The expansion coefficient of silicon as a function of temperature is provided in [9]. It increases in an approximately linear manner from 2.5 ppm/K at room temperature (23°C) to 4 ppm/K at 500°C. The residual strain in the structural metal at 23°C was 1.1×10^{-10}. Its expansion coefficient as obtained by this method increased from about 13.5 ppm/K at 50°C to about 16.5 ppm/K at 150°C. This compares well with previously published results. One report indicates that the expansion coefficient of pure bulk nickel increases from 10.2 ppm/K at 20°C to 16.3 ppm/K at 300°C, and retains the latter value at 400°C as well [10]. Another report indicates that the expansion coefficient of nickel electroplated under particular conditions may increase from 8.5 ppm/K averaged over the temperature range of 25-50°C to 15.1 ppm/K averaged over the temperature range of 25-367°C [11]. Thus, the measured results are in agreement with prior reports, given the contextual differences.
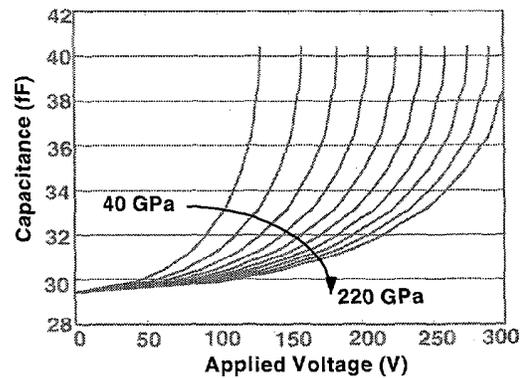
Fig. 2: Theoretical CV curves for k_{eff} values representing Young's moduli of 40-220 GPa in 20 GPa increments.
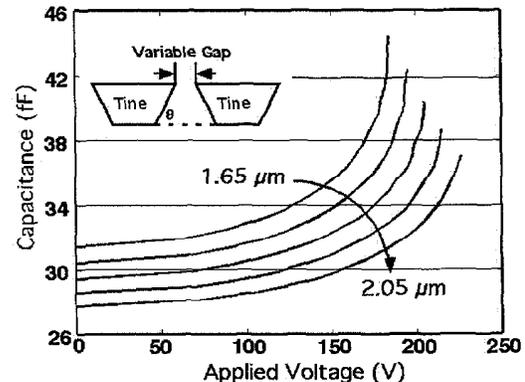
Fig. 3: Impact of gap variations between tines.
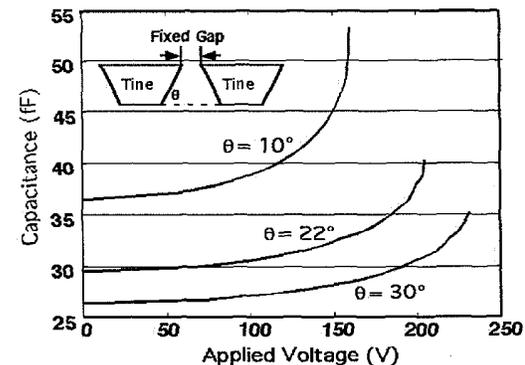
Fig. 4: Impact of variations in the sidewall angles of tines.
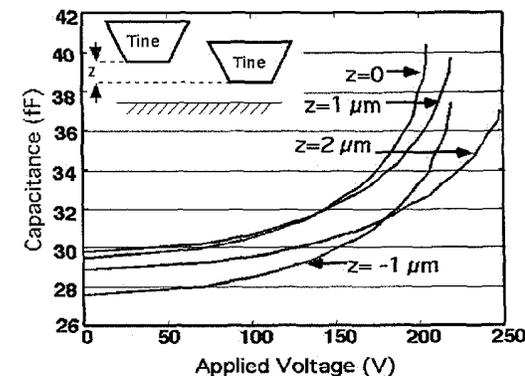
Fig. 5: Impact of out-of-plane deflections of tines.

A capacitive strain sensor was then used monitor the residual strain in the structural metal on the silicon substrate as a function of temperature for several samples (Fig. 8). The measured results of show a linear increase in the differential capacitance readout as temperature is increased from 20°C to 60°C. Measurements were taken with ±0.5 fF precision using 10 tine pairs with 168 μm overlap length and 5 μm nominal gap; for the suspensions, L=198 μm, W=6 μm, and ⌀=0.2 rad. The response of the differential capacitance readout to increasing temperature was theoretically estimated using FastCap™. For this simulation it was assumed that the expansion coefficient of the electroplated metal exceeds that of silicon by 10 ppm/K. This result is superimposed on the measured data in Fig. 8. Based on the passive strain sensor measurements (Fig. 7) and previously published results, the expansion mismatch between nickel and silicon over 20-60°C is 6-11 ppm/K, consistent with the differential capacitance measurements. At temperatures >60°C, however, the capacitive strain sensor was affected by out-of-plane buckling of the support ribs, and the capacitance model must be changed accordingly. As a design guideline, the lengths of support ribs should be minimized for highly compressive materials.

Figure 9 shows CV curves that were measured at various temperatures for devices with 10 tine pairs of 167 μm overlap length, and suspensions with L=198 μm, W=6 μm, and ⌀=0.1 rad. Numerically simulated curves were fitted to measurements using the zero bias capacitance and the Young's modulus as fitting parameters. A comparison of a measured and fitted curve pair is shown in Fig. 10. The best fit of k_{eff} corresponded to a Young's modulus of 35 GPa at 23°C for the electroplated metal. Although the Young's modulus of pure bulk nickel is 208 GPa, electroplated nickel has a substantially lower modulus and can demonstrate significant variability, underscoring the need for such measurements. For pure nickel electroplated under particular conditions, a value of 150 GPa was reported in [12]. Trace levels of contamination and variations in plating conditions can significantly affect the mechanical properties of electroplated materials. For the structural material used in these experiments, the relatively low value of Young's modulus may be related to a combination of effects, including trace impurities, aging of the solution, and possibly the plating rate.

The measured temperature coefficient of the Young's modulus, which is the same as the temperature coefficient of k_{eff}, was found by fitting measured CV curves at elevated temperatures to be -1590 ppm/K. This parameter is also strongly dependent on fabrication conditions, and values ranging from -550 ppm/K [12] to -952 ppm/K [10] have been reported for pure nickel.

Using the measured values of strain and Young's modulus the thermal stress in the metal microstructures can be calculated in a piecewise linear manner. At 23°C it is essentially neutral, at 60°C it is -14 MPa, and at 100°C it is -29 MPa. In the 23-60°C temperature range, the Young's modulus of the metal and the expansion coefficients of the metal and silicon are averaged at 34 GPa, 13.5 ppm/K, and 2.5 ppm/K respectively. In the 60-100°C temperature range, these values are 32 GPa, 15.0 ppm/, and 2.8 ppm/K, respectively. These variations in mechanical properties occur within a range of temperatures commonly encountered in microelectronics and MEMS applications, and can have significant impact on device performance.
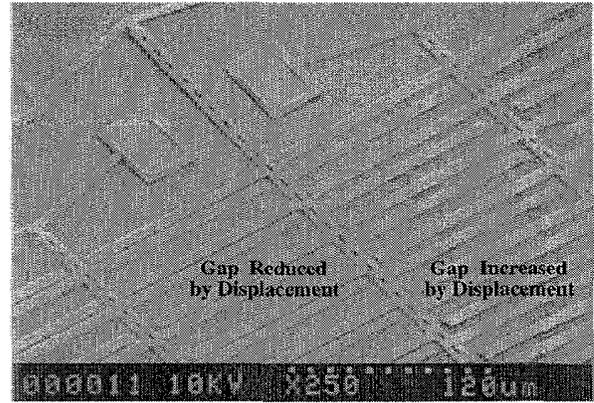
Fig. 6: SEM of a released device showing displaced tines. All tine spacings were equal before release.
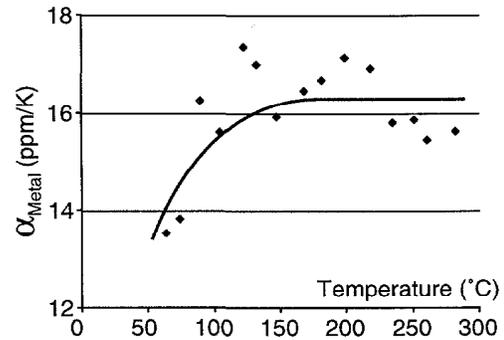
Fig. 7: The thermal expansion coefficient of the structural metal as measured by passive bent-beam strain sensors.
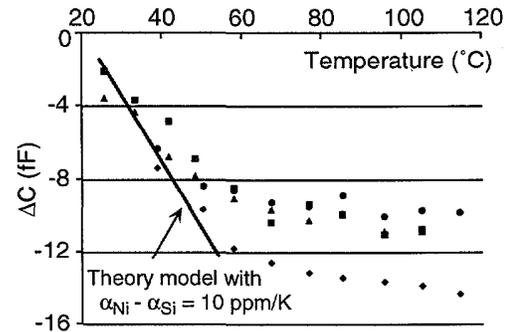
Fig. 8: The thermal expansion coefficient of the structural metal as measured by differential capacitive for four adjacent devices. The support ribs buckled near 60°C.
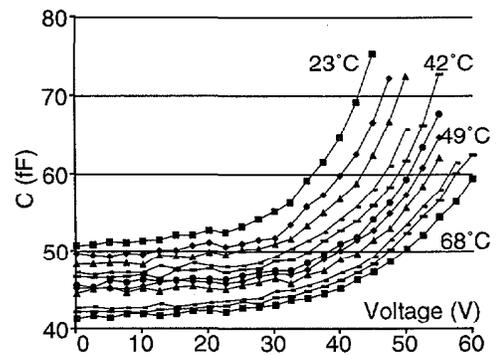
Fig. 9: Measured CV as a function of temperature. The buckling temperature for this device was >70°C.
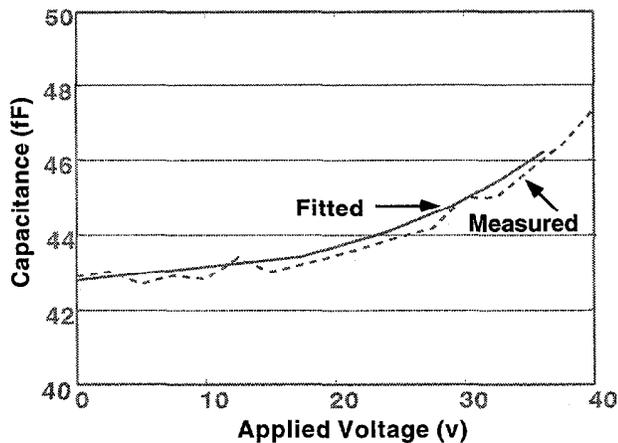
Fig. 10: Comparison of a measured CV curve from Fig. 9 to a fitted theoretical one. T =60°C.

IV. CONCLUSIONS

In summary, a procedure for determining temperature dependent mechanical properties of microstructure materials using differential capacitive strain sensors was presented in the context of electroplated nickel microstructures fabricated on a silicon substrate. For the fabrication conditions used, it was found that the thermal expansion coefficient was 11-13 ppm/K over 23-50°C; the residual strain changed from neutral to −880 microstrain over 23-100°C; the Young's modulus had a temperature coefficient of −1590 ppm/K with an estimated value of 35 GPa at room temperature; consequently, the residual stress changed from neutral to −29 MPa over 23-100°C. Using passive bent beam strain sensors, the thermal expansion coefficient was found to increase from 13.5-16.5 ppm/K over 50-150°C.

Theoretical calculations indicate that CV measurements can be affected by variations in device dimensions. In most commercial processes, however, these are held to tight tolerances. Within these constraints, differential capacitive strain sensors offer a wide dynamic range that spans both tensile and compressive materials; electronic readout, which can be used even after packaging; and measurements based on lateral deflections.

ACKNOWLEDGEMENTS

REFERENCES

[1] K. Najafi, K. Suzuki, "A novel technique and structure for the measurement of intrinsic stress and Young's modulus of thin films," *Proc., IEEE Intl. Workshop on Micro Electro Mechanical Systems (MEMS '89)*, Jan. 1989, pp. 96-97

[2] B.E. Artz, L.W. Cathey, "A finite element method for determining structural displacements resulting from electrostatic forces," *Proc., Solid-State Sensor and Actuator Workshop*, Hilton Head, SC, 1992, pp. 190-3

[3] P.M. Osterberg, S.D. Senturia, "M-TEST: A test chip for MEMS material property measurement using electrostatically actuated test structures," *IEEE Journal of Microelectromech. Sys.*, 6(2), Jun. 1997, p. 107-118

[4] E.K. Chan, K. Garikipati, R.W. Dutton, "Characterization of contact electromechanics through capacitance-voltage measurements and simulations," *IEEE Journal of Microelectromech. Sys.*, 8(2), June 1999, pp. 208-217

[5] L. Que, M.-H. Li, L. Chu, Y.B. Gianchandani, "A micro machined strain sensor with differential capacitive readout," *IEEE Intl. Conf. on Micro Electro Mechanical Systems*, Orlando, FL, Jan. 1999, pp. 552-7

[6] J.-S. Park, L.L. Chu, E. Siwapornsathain, A.D. Oliver, Y.B. Gianchandani, "Long throw and rotary output electro-thermal actuators based on bent-beam suspensions," *IEEE Intl. Conf. on Micro Electro Mechanical Systems*, Miyazaki, Japan, Jan. 2000

[7] K. Nabors, S. Kim, J. White, S. Senturia, *FastCap User's Guide*, M.I.T., Sep. 92; also K. Nabors, S. Kim, J. White, *IEEE Trans. Microwave Theory and Techniques*, 40(7), pp. 1496-1506, July 1992

[8] Y.B. Gianchandani, K. Najafi, "Bent-beam strain sensors," *Journal of Microelectromechanical Systems*, 5(1), pp. 52-58, Mar. 1996

[9] W.H. Ko, J.T. Suminto, G.J. Yeh, "Bonding techniques for micro sensors," *Micromachining & Micropackaging of Transducers*, Elsevier 1985; also *Microsensors*, pp. 198-208, Ed. Muller *et al*, IEEE Press, 1990

[10] E.M. Wise, R.H. Schaeffer, "The properties of pure nickel – I," *Metals and Alloys*, Sep. 1942, pp. 424-428

[11] J.W. Dini, H.R. Johnson, "Coefficient of thermal expansion of sulphamate nickel electrodeposits," *J. Materials Sc.*, (10)7, 1975

[12] M. Putty, *A micromachined vibrating ring gyroscope*, Ph.D. Thesis, Univ. of Michigan, Ann Arbor, pp. 211-212, Mar. 1995

RESIDUAL STRESS CHARACTERIZATION OF THICK PECVD TEOS FILM FOR POWER MEMS APPLICATIONS

X. Zhang[†], R. Ghodssi[††], K-S Chen[†††], A. A. Ayon[†] and S. M. Spearing[†]

[†]Massachusetts Institute of Technology, Cambridge, MA 02139

[††]University of Maryland, College Park, MD 20742   [†††]National Cheng-Kung University, Taiwan 701, Taiwan

ABSTRACT

This paper reports residual stress characterization of thick oxide films. Engineering approaches for creating 10-20μm oxide films without cracking and wafer bow were demonstrated.

INTRODUCTION

Thick insulating films are very important for MEMS-based devices to achieve higher power levels [1]. The deleterious effects of residual stress, however, tend to increase with film thickness. In particular, excessive wafer bow and cracking can prohibit integration within a micro-device.

The present work is focused on thick oxide films fabricated using a five-station continuous plasma processing system (Concept One™, Novellus™ Inc.) from a tetraethylorthosilicate (TEOS) precursor [2]. Conformality of TEOS in high aspect ratio features has been found to be better than that achieved by other means of growing/depositing oxide. For instance, we have successfully filled 10-30μm wide, 100-150μm deep, insulating trenches with TEOS oxide, which enables novel through-wafer interconnect schemes for igniters and temperature sensors in a micro-gas turbine device [3]. However, such thick deposited layers result in severely bowed wafers and even cracking which complicates or prevents subsequent processing.

In this paper we report a process for creating very thick (10-20μm) TEOS oxide films without cracking and wafer bow. This builds on previous work by our group to develop processes for integration of thick oxide films for Power MEMS applications [4-5]. The motivation for this work is to elucidate the factors contributing to residual stresses in TEOS films and to optimize the deposition process so as to reduce wafer bow and to avoid film cracking.

DEPOSITION AND CHARACTERIZATION

A variety of film depositions were conducted. PECVD oxide and nitride films were both deposited at 400°C while LPCVD nitride was deposited at 840°C. The deposition of PEVCD TEOS oxide was performed at low temperature (350°C). The deposition rate is nominally 0.25 μm/min and the refractive index is 1.46. The nominal gas flow rates are 2.3 ml/min of TEOS and 9500 sccm of O_2. The wafer-level TEOS thickness, deposited over a 4" silicon wafer, has non-uniformity better than 1%. However, there is a 3% reduction in film thickness across the wafer after densification.

The hydrogen content of TEOS films is about 4%, as measured with Rutherford Backscattering Spectroscopy (RBS). Subsequent densification was accomplished by subjecting the films to a higher temperature for one hour while a steady N_2 flow was maintained in the furnace. Table I tabulates the hydrogen content in 1μm TEOS films in six different conditions: as deposited, and after one hour densification at 700-1100°C. The hydrogen content

in TEOS films falls below 0.2% after densification at temperatures higher than 800°C.

Curvature measurements were performed using a KLA-Tencor™ FLX-2320 system, and the corresponding residual stress was calculated, as a function of temperature. Thermal cycling of a variety of films with a heating rate of 1°C/min was conducted and *in situ* wafer curvature was measured between room temperature and 500°C.

Table I. Hydrogen concentration of PECVD TEOS oxide.

T (°C)	350	700	800	900	1000	1100
H_2(%)	4.00	0.67	0.13	0.20	0.17	0.17

RESULTS AND DISCUSSION

The overall residual stress state is determined by the superposition of two primary effects. Thermal stress develops in thin films in which high temperature deposition or high temperature annealing is used, due to thermal expansion mismatch, and is usually unavoidable. Intrinsic stress is generated during the film growth process and is strongly dependent on deposition techniques and process parameters.
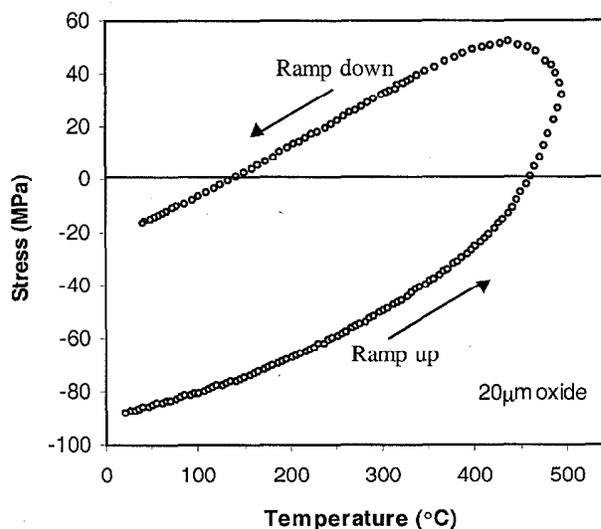
Figure 1. In situ residual stress variation of 20μm thick oxide film as a function of temperature during a 500°C thermal cycle.

Figure 1 shows the residual stress in a 20μm thick PECVD oxide film as a function of temperature during a thermal cycle between room temperature and 500°C. The dependence of residual stress on temperature is non-linear with significant hysteresis. Figure 2 shows the decomposition of thermal stress and intrinsic stress in an oxide film. The variation of stress with temperature is

apparently due to two superimposed effects: a linear thermal stress component, and a non-linear, irreversible intrinsic stress component. In this paper, thermal stress was estimated based on the assumption of constant material properties. For the case of oxide on silicon substrate: $E_f/(1-\nu_f) \sim 70GPa$, $\alpha_f \sim 0.5 \times 10^{-6} K^{-1}$, $\alpha_s \sim 3 \times 10^{-6} K^{-1}$. It is important to note that stress free temperature after thermal cycling between room temperature and 500°C is approximately 275°C, lower than the oxide deposition temperature.

In situ residual stress variation with temperature in 12μm thick PECVD TEOS oxide films is given in Figure 3. Residual stress as a function of temperature in TEOS oxide films exhibited a similar non-linearity with hysteresis behavior during the thermal cycle between room temperature and 500°C. The large hysteresis loop indicates that the intrinsic stress of TEOS oxide films was significantly reduced after thermal cycling.

There are many mechanisms that may be responsible for the generation of intrinsic stress. Typical examples include incorporation of atoms (e.g., residual gases), chemical reactions, recrystallization, dislocation rearrangements, lattice mismatch, excess vacancy annihilation, grain boundary relaxation, and phase transformations [6]. In the present case, given that the RBS analysis of TEOS films (Table I) shows two orders of magnitude difference in hydrogen concentration before and after annealing, it is likely that the initial compressive intrinsic stress of TEOS oxide was caused by dissolved gases.

Furthermore, the densification step plays an important role in determining stress and wafer bow in TEOS oxide films. Figure 4 shows room temperature wafer bow versus TEOS film thickness, before and after 1100°C densification. It is seen that the wafer bow after densification is much larger than that of the as-deposited films. This is because TEOS oxide at 1100°C is viscoelastic and stress-free (residual stress can be relieved by material flow). As a result, the stress in TEOS oxide films becomes more compressive after high temperature densification.

Figure 2. Decomposition of thermal stress and intrinsic stress in 20μm thick oxide film as a function of temperature during a 500°C thermal cycle.
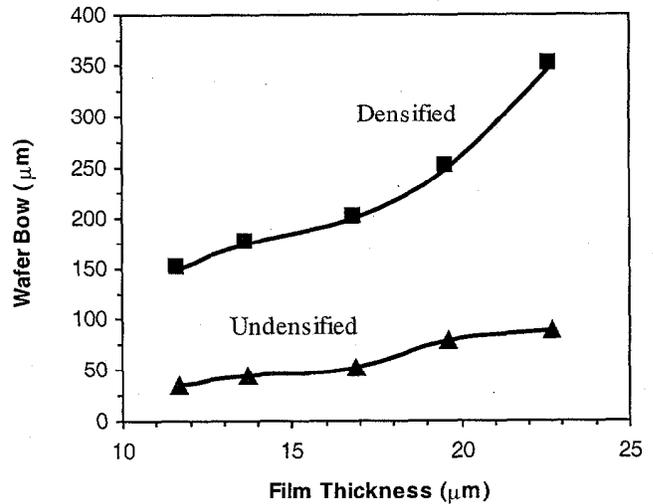
Figure 4. TEOS oxide wafer bow vs. TEOS oxide thickness before and after 1 hr, 1100°C densification.

Figure 3. In situ residual stress variation of 12μm thick TEOS oxide film (on 300μm thick silicon substrate) as a function of temperature during a 500°C thermal cycle.
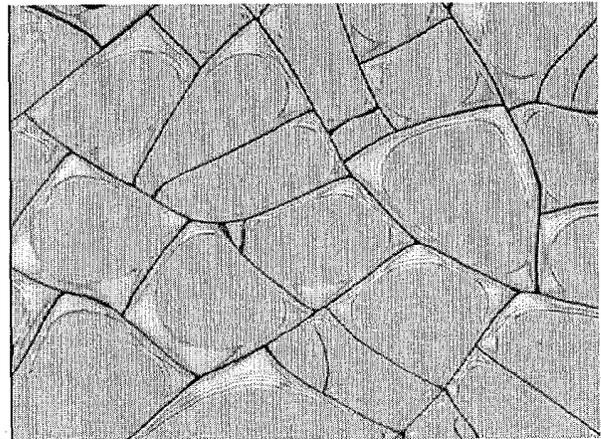
Figure 5. A cracked 25μm thick PECVD TEOS oxide film after 24 hrs annealing at 700°C.

The TEOS oxide films were observed to crack after being exposed to the annealing cycle. The tendency to form cracks is a strong function of film thickness. Figure 5 shows a cracked 25μm thick TEOS film after 24 hours annealing at 700°C. It is well known that film cracking only occurs when the film experiences a tensile stress. Since the thermal expansion coefficient of the oxide is less than that of silicon, tensile stresses will arise when the temperature is raised above the stress-free temperature. In addition, the intrinsic tensile stress generated during temperature ramp up also contributes to the failure of films. Cooling from the deposition temperature, however, leads to compressive stresses in the films and will not result in fracture.
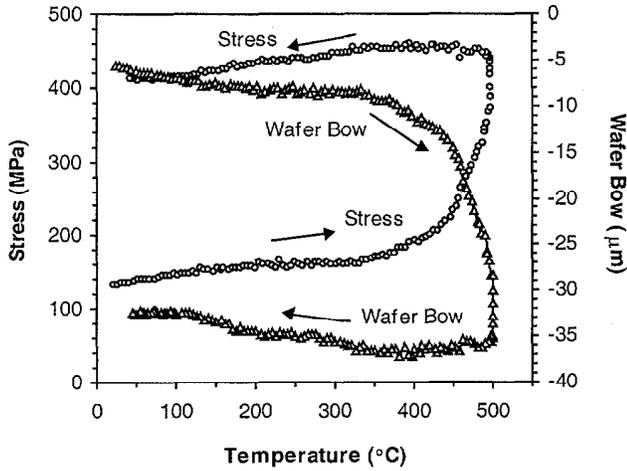
Figure 6. Residual stress and wafer bow variation with temperature in 1μm thick PECVD-nitride films.
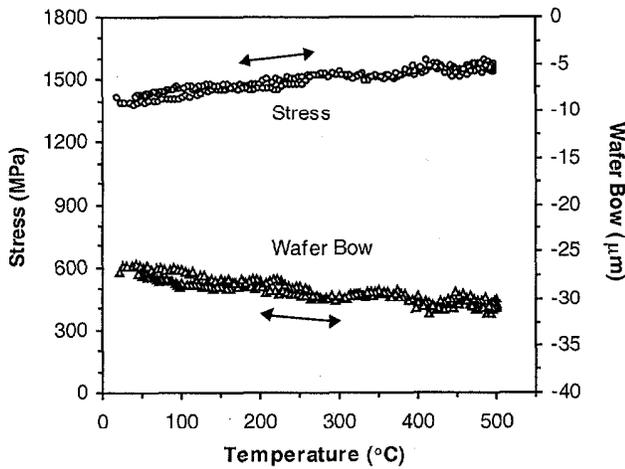
Figure 7. Residual stress and wafer bow variation with temperature in 0.25μm thick LPCVD-nitride films.

The fracture mechanics of thin films has been analyzed previously. A comprehensive summary was provided by Hutchinson and Suo [7]. The strain energy release rate for a crack propagating in a thin film is given by expressions of the form:

$$G = Z \frac{(1-v_f)\sigma^2 h_f}{E_f} \qquad (1)$$

Where h_f, E_f, and v_f are thickness, Young's modulus, and Poisson's ratio of the film, respectively, and σ is the stress in the film. Z is a dimensionless parameter which depends on the particular cracking pattern. The crack pattern of TEOS oxide in Figure 5 corresponds to a channeling crack [7]. In this case, Z takes a value of 1.976. Fracture occurs when the strain energy release rate exceeds the intrinsic fracture energy of the film material G_c. The critical temperature for fracture can be expressed as a function of film thickness, i.e.:

$$T_c = T_0 + \frac{1}{(\alpha_s - \alpha_f)} \sqrt{\frac{(1-v_f)G_c}{ZE_f h_f}} - \frac{\sigma_I (1-v_f)}{E_f (\alpha_s - \alpha_f)} \qquad (2)$$

Using the nominal material properties of the TEOS oxide, the calculated critical temperature to form cracks is about 950°C, which is significantly higher than the observation (about 700°C). This is because the contribution of intrinsic stress was not included. However, the critical temperature would be reduced if the effect of intrinsic stress is considered. For example, if the intrinsic stress was 40 MPa, the critical temperature would be reduced to 720 °C. This calculation is only approximate because no detailed material properties and intrinsic stress data are available.

The excessive wafer bow after densification causes difficulties for subsequent processing. To counteract the compressive stress inherent to the TEOS film, a highly tensile silicon nitride film was deposited prior to deposition of the thick TEOS oxide film. Residual stress and wafer bow behavior of the nitride alone under thermal cycling is shown in Figures 6 and 7. Compared to 1μm thick PECVD-nitride behavior given in Figure 6, 0.25μm LPCVD-nitride (Figure 7) exhibited virtually no hysterisis.
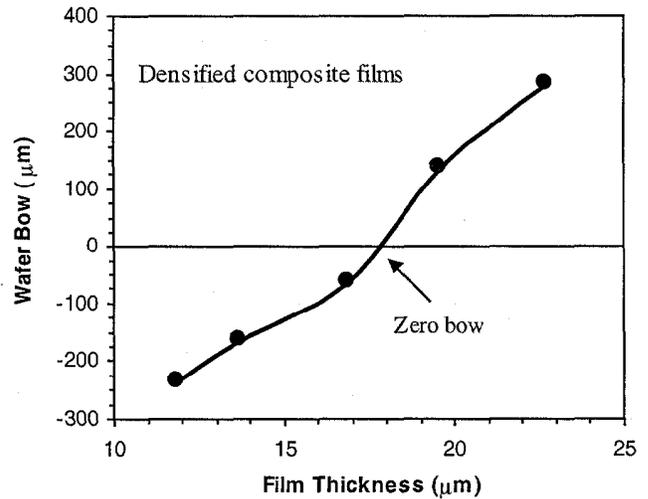
Figure 8. Wafer bow of composite films (0.25μm thick LPCVD-nitride + 10-25μm thick PECVD-TEOS) vs. film thickness after 1hr densification at 1100°C.

The objective of using a composite film is to introduce a film with tensile residual stress to counterbalance the effect of TEOS oxide. With proper calculation of thickness ratio and process control, it should be possible to create a wafer with low curvature. The wafer of composite films (0.25μm LPCVD-nitride + 10-25μm PECVD-TEOS) versus film thickness after one hour, 1100°C densification is given in Figure 8. The composite exhibited less

and in some case even zero wafer bow for a given film thickness. Further deformation of such composite films was not observed after they were subjected to a 500°C thermal cycling. The stability of the composite films under thermal cycling is illustrated by Figure 9.
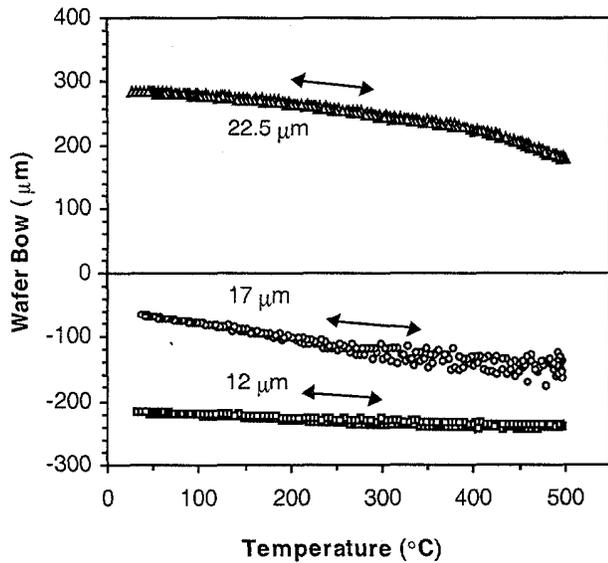
Figure 9 Stability of wafer bow of 1100°C densified wafers coated with composite films during a 500°C thermal cycle.
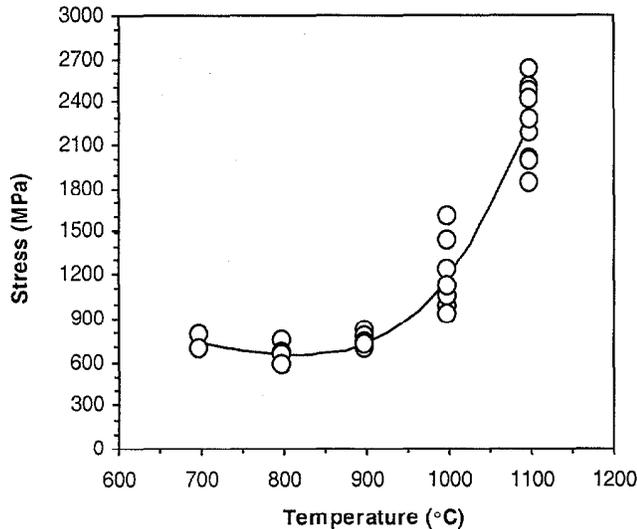
Figure 10. Residual stress variation with annealing temperature in 1 μm thick PECVD-nitride films.

It should be emphasized that although there is no deflection macroscopically, the local residual stress was not reduced. As a result, this approach may still encounter problems caused by high residual stress such as delamination. Residual stress variation with annealing temperature in 1μm thick PECVD-nitride films is given in Figure 10. The very high tensile stress (~2500 MPa) is sufficient to cause delamination of the film from the substrate after high temperature annealing, as shown in Figure 11.
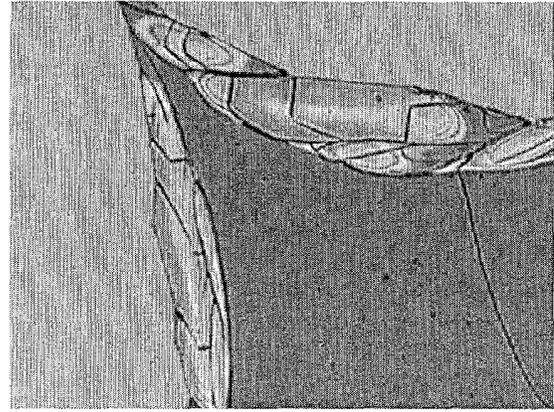
Figure 11. A delaminated 1μm PECVD-Nitride + 12μm PECVD-TEOS oxide film from its silicon substrate.

CONCLUSIONS

Residual stress characterization of thick TEOS oxide films was performed. Dissolved gases were found to play an important role in governing intrinsic stress. The tendency to form cracks is a strong function of film thickness and annealing temperature. Engineering approaches for creating 10-20μm oxide films without cracking and wafer bow were demonstrated.

ACKNOWLEDGEMENTS

REFERENCES

1. A. H. Epstein and S. D. Senturia, "Macro Power from Micro Machinery", *Science, 276,* 1211 (1997).
2. *Novellus Systems Inc.*, 81 Visa Montana, San Jose, CA 95134.
3. A. Mehra, X. Zhang, A. A. Ayon, I. Waitz, and M. A. Schmidt, "A Through-Wafer Electrical Interconnect for Multi-Level MEMS Devices", *Journal of Vacuum Science & Technology A,* in press, (2000).
4. R. Ghodssi, L. Frechette, S. F. Nagle, X. Zhang, A. A. Ayon, S. D. Senturia, and M. A. Schmidt, "Thick Buried Oxide in Silicon (TBOS): An Integrated Fabrication Technology for Multi-Stack Wafer-Bonded MEMS Processes", *Proceedings of the 1999 International Conference on Solid-State Sensors and Actuators,* Sendai, Japan, June 7-10, 1999, pp. 1456-1459.
5. R. Ghodssi, X. Zhang, K-S Chen, S. M. Spearing, and M. A. Schmidt, "Residual Stress Characterization of Thick PECVD Oxide Film for MEMS Application", *The 46th International Symposium of the American Vacuum Society,* Seattle, WA, 10/25-29/99.
6. W. Buckel, "Internal Stresses", *Journal of Vacuum Science & Technology, 6,* 606 (1969).
7. W. Hutchinson and Z. Suo, "Mixed Mode Cracking in Layered Materials", *Advanced in Applied Mechanics, 29,* 63 (1991).

ALKENE BASED MONOLAYER FILMS AS ANTI-STICTION COATINGS FOR POLYSILICON MEMS

W. Robert Ashurst, Christina Yau, Carlo Carraro, Roger T. Howe[*], and Roya Maboudian

Berkeley Sensor & Actuator Center
Department of Chemical Engineering
* Departments of Electrical Engineering and Computer Sciences and Mechanical Engineering
University of California at Berkeley, Berkeley CA 94720-1492 USA
tel (510) 643-3489, fax (510) 642-4778, ashurbr@uclink4.berkeley.edu

ABSTRACT

This paper describes a new class of anti-stiction coatings for silicon MEMS that is based on the free radical reaction of a primary alkene (e.g., 1-octadecene, $C_{16}H_{33}CH=CH_2$) with hydrogen terminated silicon [1,2]. The new monolayer coating has several key advantages over the previously reported octadecyltrichlorosilane (OTS) and 1H,1H,2H,2H-perfluorodecyltrichlorosilane (FDTS) based self-assembled monolayers (SAMs) [3,4]: 1) The coating does not produce HCl at any stage in the monolayer formation whereas chlorosilane-based chemistry does. 2) The coating does not require the formation of an intervening oxide layer. 3) The film formation procedure for alkene-based monolayers is simpler than for chlorosilane-based SAMs for two main reasons. First, the surface oxidation step is eliminated. Second, the coating solution does not need to be conditioned before use, since water is not a reagent in this process. 4) The coating process is much more robust since it is essentially insensitive to relative humidity. 5) The coated surfaces have much fewer particulates in comparison to those coated with OTS. 6) The coating process can be made selective to coat only exposed silicon by generating radicals using a radical initiator.

The coating is evaluated in several ways, including X-ray photoelectron spectroscopy (XPS), atomic force microscopy (AFM), contact angle analysis (CAA), work of adhesion by cantilever beam array (CBA) technique and coefficient of static friction using a sidewall testing device. The octadecene film is compared to the OTS SAM. XPS data confirm the absence of oxygen in both freshly prepared samples and in samples aged for four months in laboratory ambient. AFM shows that the samples which receive 1-octadecene films accumulate far fewer particles during processing than those which receive the OTS SAM treatment. These improvements have been achieved without sacrificing the stiction characteristics of coated micromachines, in that water and hexadecane contact angles, apparent work of adhesion, and coefficient of static friction data are found to be similar to those of OTS.

INTRODUCTION

The large surface-area-to-volume ratios of surface and bulk micromachined micromechanisms bring the role of stiction into the foreground, as adhesion of these mechanisms to adjacent surfaces is a major failure mode for MEMS [5-7]. Stiction is a term that has been applied to the unintentional adhesion of compliant microstructure surfaces when restoring forces are unable to overcome interfacial forces such as capillary, van der Waals and electrostatic attractions. Release stiction, the adhesion of surface-micromachined structures to the underlying substrate following the final sacrificial layer etch, is caused primarily by liquid capillary forces. Engineering solutions to this problem include a variety of techniques which have been reviewed elsewhere [5-7]. Most of

these techniques, however, do not prevent adhesion from occurring during micromachine operation. Microstructure surfaces may come into contact unintentionally through acceleration or electrostatic forces, or intentionally in applications where surfaces impact or shear against each other. When adhesive attractions exceed restoring forces, surfaces permanently adhere to each other causing device failure -- a phenomenon known as in-use stiction. In addition, it is known that on the microscale, friction is strongly dependent upon adhesion [8].

In order to alleviate these adhesion-related problems, both the topography and the chemical composition of the contacting surfaces must be controlled. Several approaches have been developed to address these tribological problems and have been reviewed elsewhere [5-7]. Our group has investigated both OTS and FDTS SAMs for alleviating adhesion in polycrystalline silicon microstructures. These SAMs have been shown to effectively eliminate release stiction, and significantly reduce in-use stiction, friction, and to some extent wear [9-11].

Although these SAM coatings have been shown to significantly reduce both release and in-use stiction, they possess a number of limitations intrinsically related to their chemistry. As with any chlorosilane-based SAM precursor molecule, the first step in the reaction sequence for binding the molecule to the substrate is the hydrolysis of one or more Si-Cl bonds. This hydrolysis thus results in one equivalent of HCl for each Si-Cl bond that is hydrolyzed. The presence of HCl in the SAM coating solution poses a threat to exposed aluminum interconnects, as well as other metals or metal compounds that may be present. Additionally, silicon surfaces must first be oxidized in order for any chlorosilane SAM to form. This is usually accomplished by growing a thin chemical oxide (~20 Å) with an oxidizing agent such as hydrogen peroxide. This oxidation step further complicates the coating process and poses a barrier for any strategy to achieve selective coatings. Another limitation arises from the ability of the precursor molecule to polymerize. As long as the precursor molecule has a functionality greater than one, bulk polymerization can occur, producing particulates that are up to several microns in diameter. These particulates can mechanically interfere with the device operation and pose reliability concerns. Unfortunately, there is no satisfactory method for removal of the polymerized clusters once they have agglomerated on the surfaces of the micromachines.

Other limitations of the chlorosilane SAM coatings arise from the coating procedure. The coating process is somewhat cumbersome in that the SAM solution must be freshly made and appropriately conditioned immediately before each coating. This is due to the sensitivity of the SAM solution to ambient humidity, and the ability of the SAM precursors to polymerize. As mentioned earlier, a hydrolysis step is required for SAM formation. The water, which is a reagent, necessary for this step is introduced to the organic SAM solution by mass transfer from the

ambient air. Hence it is necessary to control the ambient humidity such that there will be enough water present in solution to carry out the hydrolysis, but an insufficient concentration to promote bulk polymerization of the SAM precursor molecules. The tolerable relative humidity range is from 30% to 70%.

Based upon the desired anti-stiction film characteristics, in particular surface hydrophobicity, and the considerations of the chemistry and coating processes, we have investigated the feasibility of employing a new class of monolayer coatings as anti-stiction films. This new class of films abandons the chlorosilane chemistry and adopts a free radical reaction of a primary alkene to bind the precursor molecule to a hydrogen terminated silicon substrate [1,2]. The reagent we have examined is 1-octadecene, $C_{16}H_{33}CH=CH_2$. Due to the substantially different chemistry of this precursor molecule, no HCl is produced during the reaction. Additionally, there is no need for oxidation of the surface, since the molecule binds *directly* to the silicon. Water is not a reagent in this reaction sequence, and hence the coating procedure is insensitive to ambient humidity. The alkene is very stable at room temperature; the coating solution can be stored indefinitely. The coating solution can even be re-used or recycled from previous treatments if it is so desired.

The coating procedure is described in detail below. We also report on the characteristics of the resulting film, using XPS, AFM, and CAA. Furthermore, we have determined the values of work of adhesion by cantilever beam array technique, as well as coefficient of static friction and thermal stability in vacuum and oxidizing environments.

EXPERIMENTAL

The films studied in this work are the 1-octadecene monolayer and the OTS SAM for comparison. Both precursor molecules contain a straight 18-carbon chain, and bind to the substrate at only one end. Hence, the films created with these particular molecules are directly comparable in terms of anti-stiction behavior. For reference, structures were also prepared without any anti-stiction coating, i.e., oxidized surfaces. The procedure used for the OTS SAM coating was outlined in ref. 9 with the following exceptions. The solvent for the OTS was anhydrous iso-octane (Aldrich Chemical Co.) and the SAM formation was carried out for 60 minutes instead of 15 minutes. The procedure that was carried out for the oxidized microstructures consisted of a standard release followed by surface oxidation with 30% H_2O_2 at 90 °C for 10 minutes. The structures were then rinsed with deionized water and either supercritically dried or removed directly from water and air-dried.

In order to address the issue of large volumes of waste solvent generated by the displacement rinses commonly found in monolayer preparation procedures, a solvent "fill/drain" approach was developed instead of an aspiration approach. A more complete description of this system can be found in ref. **12**. This device consists of a Teflon FEP separatory funnel modified by cutting the top such that the rounded portion is removed. In addition, a Teflon ETFE plug is machined to fit inside the modified funnel so as to leave sufficient volume above the cone for process liquids. The plug reduces the amount of solvents needed, by reducing the stagnant volume of the funnel. This device was used for all displacement rinses.

The general coating procedure for the alkene-based films is summarized in Table 1. As with any microstructure release procedure, the first step is the sacrificial layer etch. Typically, this is done with concentrated HF for an amount of time that depends of the fabrication history of the structures being released. After the sacrificial layer etch, the silicon surfaces will be hydrogen terminated. However, the etch products that result from removal of the sacrificial layers need to be thoroughly rinsed away, and the

etch solution displaced. At this point, a deionized water rinse (18-MΩ-cm) serves both of these purposes. However, during the time that the rinsing is taking place, the water can oxidize the hydrogen terminated surface to the extent that the integrity of the alkene coating may be jeopardized. Since the hydrogen termination is needed for the reaction, the silicon surfaces are treated again with HF to recover the hydrogen termination. This time, however, the HF is displaced with isopropyl alcohol (IPA) which does not oxidize hydrogen-terminated silicon. Additionally, the IPA serves as a useful intermediate between the organic 1-octadecene coating solution and the aqueous HF. The structures are then rinsed with the 1-octadecene. Now the structures are ready to be transferred into a reactor containing the octadecene coating solution. This solution is a 9:1 v/v solution of hexadecane:1-octadecene. The hexadecane is anhydrous spectroscopic grade (Aldrich Chemical Co.) and the 1-octadecene is technical grade, 90%, anhydrous (Aldrich Chemical Co.). The octadecene is diluted to reduce the potential for bulk polymerization of the difunctional alkene. After the structures are placed inside the reactor and the lid is sealed, a dry nitrogen gas is turned on in order to purge oxygen from the vapor space above the coating solution. This gas flow was maintained throughout the heating process. The reactor is heated to 180 °C and held at temperature for 30 minutes. Heating was chosen as the method to generate free radicals; this is the step during which the monolayer forms. The reactor is then allowed to cool. The structures are then removed from the coating solution and *thoroughly* displacement rinsed into petroleum ether (Fisher, b.p. 38-54 °C). The structures are then displaced into IPA and rinsed further. Next, the IPA is completely washed away with DI water, while the liquid surface is aspirated. At this point, the structures can be removed from the water.

Step	Procedure	Approximate Time
Sacrificial Layer Etch	HF	10 min.
Octadecene Coating	Water rinse to pH neutrality	10 min.
	Secondary HF for hydrogen termination	2 min.
	Displacement rinse to IPA	5 min.
	Displacement rinse to 1-octadecene solution	5 min.
	1-Octadecene treatment at 180 °C	30 min.
Rinse and Dry	Displacement rinse to petroleum ether	5 min.
	Displacement rinse to IPA	5 min.
	Displacement rinse to water	10 min.
	Removal from water	

Table 1. *The basic processing steps for a 1-octadecene anti-stiction coating.*

Contact angle data were taken with DI water (resistivity 18-MΩ-cm) and spectroscopic grade hexadecane (Aldrich Chemical Co.) according to the sessile droplet method with a Ramé Hart 100 A goniometer. Droplet size was approximately 4 μl, and the measurement reproducibility was ±2°. A Digital Instruments Nanoscope III atomic force microscope was used to image sample surfaces and quantify their roughnesses. XPS data were taken with a Kratos XSAM 800 system. The probing system used for all micromachine actuation was a Lucas-Signatone S-1160 probe station with a Mitutoyo FineScope 60 microscope, equipped with a Sony CCD-IRIS camera. A Nomarski style differential interference contrast prism (manufactured by Olympus) was used with a Mitutoyo M Plan Apo 10x objective on the microscope.

All micromachine actuation was done under normal laboratory ambient conditions, 20 °C and 50% relative humidity. Work of adhesion by cantilever beam array technique was performed in accordance with the procedure outlined in ref. 9. The CBA test structures used in this study were fabricated at Sandia National Labs, Albuquerque, NM. The beams were of lengths varying from 150 to 1700 microns, 2.5 microns thick and 2 microns above the substrate. The RMS roughness of the groundplane underneath the beams was about 2.3 nm. Coefficient of static friction was determined from a sidewall testing device described in ref. 10. Actuation of this device and analysis of data were in accordance with the procedures outlined in ref. 13.

The thermal stability of the octadecene film has been studied on Si(100) in vacuum and in oxidizing ambients. In the latter case, the sample was exposed to high temperature, in laboratory air, for five minutes, and then removed to a room temperature aluminum slab for cooling. The five minute exposure time was chosen based upon estimates of integrated circuit packaging times. The film stability was checked by contact angle goniometry, and the water contact angles were taken immediately after the samples cooled.

RESULTS AND DISCUSSION

Contact angle data are summarized in Table 2. Also listed are the representative values for oxide and OTS coated surfaces for reference. The contact angle data indicate that the molecular packing of film produced by the octadecene coating process is not as high as that produced by the OTS coating process. However, this difference appears to have little effect on the anti-stiction properties of the film. XPS analysis on Si(100) confirmed the absence of silicon oxide in the freshly made 1-octadecene samples. To gauge the passivation effect of the coating, samples were prepared and left in laboratory environment for various durations. Even after four months of storage, no silicon oxide was observed with XPS.

Although it has been observed that the octadecene reaction is not selective if the initiation is by heating, it can be made selective to hydrogen terminated silicon by changing the method of free radical formation. Ultraviolet excitation and radical initiators (such as 2,2'-azobisisobutyronitrile) are two other ways in which free radicals may be formed. Ultraviolet excitation may not be an appropriate means of initiation due to conformality concerns for micromachines. It has been observed that the use of a radical initiator results in a selective coating process. Although not selective, heating remains the simplest and most uniform means of causing the monolayer formation and hence, all data reported here arises from monolayers formed by this method.

AFM was used to study particulates that accumulate on the surface during the coating procedure. Figures 1 (a) and (b) are AFM images for OTS and 1-octadecene on Si(100) respectively. These samples have been treated alongside microstructures, and hence, undergone every processing step exactly as the structures. They were not cleaned in any way after final removal from water. The scan size is 30μm × 30μm and the z scale (height) is 100 nm. Note that in contrast to octadecene, the OTS sample is cluttered with large particulates that are on the order of microns in diameter. The RMS roughness based on image statistics is 25 nm for the OTS sample and 0.4 nm for the octadecene sample.

In CBA test structures, beams as long as 900 microns were consistently found to be free standing after release. The CBAs were subsequently used to find apparent work of adhesion values for the films. Due to the inherent variation from chip to chip, beam sticking probability data were accumulated over many samples prepared independently by two operators. The OTS sticking probability function exhibits a sharp increase in the fraction stuck as a function of beam length near the detachment

length, whereas the octadecene curve exhibits a more gradual transition. More work is underway to determine the nature of the difference between these coatings in this respect. The average values of apparent work of adhesion are summarized in Table 2. Note that the work of adhesion for the octadecene coating is similar to that of the OTS SAM. Additionally, the work of adhesion for both coatings is roughly four orders of magnitude lower than that of the standard oxide coating.

Sidewall friction testers were used to determine the coefficient of static friction on micromachine sidewalls. Again, due to chip to chip variation, a number of samples prepared by different operators were tested for each surface treatment. The results for μ_S are summarized in Table 2. Note that the 1-octadecene treatment and the OTS SAM have similar coefficients of friction, and that their values are roughly a factor of 20 less than that of an oxidized surface.

Surface Coating	Water Contact Angle	Hexadecane Contact Angle	Work of Adhesion (mJ/m²)	μ_S
Octadecene	104 °	35 °	0.009	0.050
OTS	109 °	38 °	0.012	0.073
Oxide	0 – 30 °	0 – 20 °	30	1.1

Table 2: Property data comparison for different surface preparations.
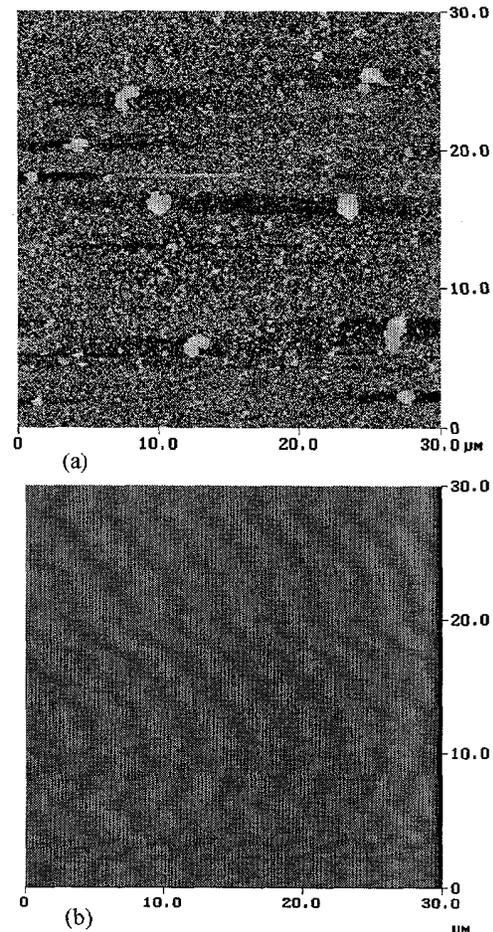
Figure 1: AFM image of (a) an OTS SAM and (b) 1-octadecene coated Si(100) chip. The samples have undergone identical treatments as micromachine chips. The particulates on the OTS surface are attributed to bulk polymerization of the precursor

molecules. The z scale in both images is 100 nm and the RMS roughness is 25 nm for (a) and 0.4 nm for (b).

Thermal behavior of the octadecene monolayer in vacuum has been studied by Sung et al. [2]. The film was found to degrade in vacuum at about 350 °C, and by 525 °C evidence of a thin SiC film was observed. The same authors have investigated the thermal stability of the OTS coating. The film degradation begins at about 450 °C, and by about 620 °C, only the underlying oxide remains [2]. In air, the thermal stability is approximately the same for both the octadecene and OTS films. Both films begin to degrade at about 225 °C, where the octadecene film shows a sharper falloff in contact angle than does the OTS film. Note that the FDTS films were found to be stable in air up to 400 °C [9].

CONCLUSIONS AND FUTURE WORK

Based upon the data presented here, we find that the anti-stiction properties of films produced with the alkene chemistry are indeed comparable to those produced with the trichlorosilane SAMs, but without many of the limitations imposed by the chlorosilane chemistry. The apparent work of adhesion for 1-octadecene coated polysilicon is nearly four orders of magnitude lower than that of the standard oxide coating and slightly less than that of the OTS SAM. The sidewall friction testers showed that the octadecene treatment and the OTS SAM have similar coefficients of static friction, and that they are roughly a factor of 20 less than that of the standard oxidized surface. XPS data confirms that the alkene-based monolayers are effective at passivating the silicon surface to oxidation. It was also shown that films produced with alkene precursors accumulate fewer particulates than those produced by the OTS precursor. The thermal stability of the octadecene films is nearly the same as that for the OTS SAM.

Thermal annealing of octadecene films in vacuum has been shown to produce silicon carbide [2]. Current work suggests that an octadecene treatment followed by rapid thermal annealing in vacuum could become a viable route to further carbidization by some other method.

ACKNOWLEDGEMENTS

REFERENCES

1. M. R. Linford, P. Fenter, P. M. Eisenberger, and C. E. D. Chidsey, "Alkyl Monolayers on Silicon Prepared from 1-Alkenes and Hydrogen-Terminated Silicon", *J. Am. Chem. Soc.* 117, 3145-3151 (1995).

2. M. M. Sung, G. J. Kluth, O.W. Yauw, and R. Maboudian, "Thermal Behavior of Alkyl Monolayers on Silicon Surfaces", *Langmuir*, vol. 13, no 23, 6164-6168 (1997).

3. M. R. Houston, R. T. Howe, and R. Maboudian, "Self-Assembled Monolayer Films as Durable Anti-Stiction Coatings for Polysilicon Microstructures", *Technical Digest of the 1996 Solid-*

State Sensor and Actuator Workshop, Hilton Head '96, pp. 42-47 (1996).

4. U. Srinivasan, M. R. Houston, R. T. Howe, and R. Maboudian, "Self-Assembled Fluorocarbon Films for Enhanced Stiction Reduction", *Technical Digest, Proceedings of the 9th International Conference on Solid-State Sensors and Actuators*, Transducers '97, pp. 210-213 (1997).

5. R. Maboudian and R.T. Howe, "Critical review: Adhesion in surface micromechanical structures," *J. Vac. Sci. Technol. B*, vol. 15, pp. 1-20, (1997); and references therein.

6. N. Tas, T. Sonnenberg, H. Jansen, R. Legtenberg, and M. Elwenspoek, "Stiction in surface micromachining", *J. Micromech. Microeng.*, vol. 6, pp. 385-397, (1996); and references therein.

7. K. Komvopoulous, "Surface Engineering and Microtribology for Micromechanical Systems", *Wear*, vol. 200, pp. 305-327, (1996).

8. B. Bhushan, *Tribology and Mechanics of Magnetic Storage Devices*, Springer, New York, (1990).

9. U. Srinivasan, M. R. Houston, R. T. Howe, and R. Maboudian, "Alkyltrichlorosilane-Based Self Assembled Monolayer Films for Stiction Reduction in Silicon Micromachines" *J. Microelectromechanical Systems*, vol. 7, no. 2, (1998).

10. D. C. Senft and M. T. Dugger, "Friction and Wear in Surface Micromachined Tribological Test Devices", *SPIE* vol. 3224, pp. 31–38, (1997).

11. U. Srinivasan, J. D. Foster, U. Habib, R. T. Howe, and R. Maboudian, "Lubrication of Polysilicon Micromechanisms with Self-Assembled Monolayers", *Technical Digest of the 1998 Solid-State Sensor and Actuator Workshop*, Hilton Head '98, pp. 156-161 (1998).

12. R. Maboudian, W. R. Ashurst, and C. Carraro, "Self-Assembled Monolayers as Anti-Stiction Coatings for MEMS: Characteristics and Recent Developments", *Sensors and Actuators A*, in press.

13. W. R. Ashurst, C. Yau, C. Carraro, M. T. Dugger, R. T. Howe, and R. Maboudian, Manuscript in preparation.

Application of Micromachining Technology to
Lateral Field Emission Devices

Veljko Milanović[1], Lance Doherty, Dana A. Teasdale, Chen Zhang, Siavash Parsa[†], Kristofer S. J. Pister

Berkeley Sensor and Actuator Center
497 Cory Hall; Berkeley, CA 94720

[†]UC Berkeley Microfabrication Laboratory
406 Cory Hall; Berkeley, CA 94720

Abstract - We demonstrate a range of novel applications of micromachining and microelectromechanical systems (MEMS) in achieving efficient and tunable field emission devices (FEDs), and in improving their ease of integration. In each case, arrays of lateral field emission tips are fabricated with submicron spacing utilizing deep reactive ion etch (DRIE). Tip density of over $150 \cdot 10^6$ tips/cm^2 can be achieved, and current of over 150 A/cm^2. By utilizing sacrificial sidewall spacing, electrodes can be placed at arbitrarily close distances. We further utilize MEMS actuators to laterally adjust electrode distances. To improve integration capability of the FEDs, we demonstrate batch bump-transfer integration of working FEDs onto a quartz target substrate.

I. INTRODUCTION

The research field of vacuum microelectronics strives to combine the best features of highly advanced solid-state fabrication technology and vacuum tube electronics [1],[2]. The goals and potential applications vary from high-temperature and high-radiation amplifiers; RF oscillators and amplifiers and field emitter displays [2]-[7], to a variety of sensors such as ion gauges [6]. To date, the proposed and fabricated field emission devices (FEDs) have been largely divided into two groups: vertical and lateral structures. Vertical FEDs were frequently fabricated by the use of silicon or Spindt-type metal tips [5]. To our knowledge, the achievable currents per device with silicon as emitting material have been limited to below or the order of 1 μA. Hence, in most proposed applications, large field emitter arrays (FEAs) of up to tens of thousands of tips are required to achieve milliamps of electron current. Significantly higher currents have been achieved by the use of metals with lower work functions [5]. FEAs are commonly characterized by turn-on voltages above 60 Volts.

Lateral field emission devices (LFEDs) [7]-[9] may have many advantages in high-speed and RF applications owing to the simple fabrication, precise control of electrode distances, and lower capacitance. They have been demonstrated with relatively high current densities and turn-on voltages as low as 22 Volts [9]. Lateral silicon emitters with currents on the order of tens of microamps per tip have been demonstrated.

Utilization of micromachining technology presented in this work allows for more than an order of magnitude increase in achievable current for each individual device than in the previous work with silicon emitters. With the use of deep reactive ion etch [10],[11] (DRIE) which cyclically alternates between etch and deposition steps forming scalloped sidewalls, each fabricated lateral device can consist of dozens of tips vertically stacked with <0.5 μm distance. The methodology is very simple, and requires only 1 mask and 1 trench etch with DRIE. It also provides for very accurate electrode distance control via photolithography, oxidation self-alignment, or sacrificial sidewall spacer self-

alignment. These are discussed in detail in Sec. III. Diodes with turn-on voltages between 25 and 40 V emitting currents up to 400 μA were fabricated in a 2 μm process.

The new approach is moreover directly integrateable with MEMS actuators which provide a *mechanical tuning* feature and open numerous possibilities in the area of inertial sensing. Electrodes can be placed at closer, arbitrary distances by lateral actuation using comb drive, gap-closing, or thermal actuators as described in Sec. IV.

Finally, the approach provides integration capability for these types of devices by implementing batch-transfer methodology for fully fabricated diodes and triodes. This methodology, described in detail by Maharbiz *et al* [12] has the advantage of decoupling circuit and MEMS processes. In our case, CMOS fabrication can therefore be completed independently of FED fabrication on separate wafers and in separate, optimized, processes. After completion of both processes, FEDs are batch-transferred onto the CMOS wafer for integration with sense/drive and digital circuits. The methodology and results are detailed in Sec. V.

II. THEORETICAL BACKGROUND

Emission of electrons from a cold metal cathode is characterized by the Fowler-Nordheim tunneling equation [13]. In most cases the equation closely approximates current-voltage characteristics for relatively highly doped semiconductors and is frequently used to extract physical parameters such as cathode sharpness, material work function, and field enhancement factor. We utilize an approximation by Spindt *et al* [4] of the current I in terms the operating voltage V, work function ϕ *[eV]*, emission area α *[m^2]* and the local field conversion factor β *[m^{-1}]*:

$$I = aV^2 \, exp(-b/V), \qquad (1)$$

where

$$a = \frac{\alpha A \beta^2}{1.1\phi} \cdot exp\left(\frac{1.44 \times 10^{-7} B}{\phi^{1/2}}\right), \qquad (2)$$

$$b = 0.95 B \phi^{3/2} / \beta$$

and A and B are dimensionless constants, $A=1.54 \times 10^{-6}$ and $B=6.87 \times 10^7$. The parameter β can be expressed as the ratio of field enhancement to the cathode-anode distance

$$\beta = \frac{\kappa_{tip}}{d_{cathode-anode}}. \qquad (3)$$

The dimensionless κ_{tip} represents the geometrical enhancement of electric field E *[V/m]* at the emitter tip due to its small radius of curvature, $\kappa_{tip}=E/(V/d)$. In the high voltage regime, the exponential term in the current equation may be neglected and current is simply $I=aV^2$. Integration of FEDs with MEMS actuators allows for electrode distance adjustment and modulation. To understand these effects on current, we differentiate current in high voltage regime with respect to d,

---

[1] veljko@eecs.berkeley.edu

yielding:

$$\frac{\partial I}{\partial d} \approx V^2 \cdot \frac{\partial a}{\partial d} = \frac{2\alpha\beta A V^2}{1.1\phi}\exp\left(\frac{1.44\times10^{-7}B}{\phi^{1/2}}\right)\frac{\partial\beta}{\partial d}$$

$$= -\frac{2\alpha A V^2}{1.1\phi}\exp\left(\frac{1.44\times10^{-7}B}{\phi^{1/2}}\right)\frac{\kappa_{tip}^2}{d^3} \qquad (4)$$

As evident in the equations above, there are five physical ways of increasing the emitted current:

1) Increasing voltage
2) Decreasing inter-electrode distance
3) Decreasing material work function by increasing the dopant concentration in the n-type silicon or by changing materials
4) Increasing the effective field at the tip through sharpening
5) Increasing the number of tips in a given area, i.e. tip density

Our fabrication efforts, as detailed in the Sec. III, are directed at simple and low-cost ways of optimizing 2), 4), and 5).

III. DEVICE FABRICATION AND CHARACTERIZATION

A. Deep reactive ion etch for tip fabrication

This method of tip fabrication is based on the cyclical nature of the time multiplexed inductively coupled plasma etch (DRIE) [10], i.e. the patented Bosch process [11]. The process etches deep, high aspect ratio trenches in silicon by alternating an SF_6 plasma etch with a teflon-like C_4F_8 deposition. As a consequence, trench sidewalls are not smooth but undulate with the process' periodicity with so-called sidewall *scallops*, as illustrated in Fig. 1b and also shown in Figs. 2 and 3. The scallop is commonly an undesired property of DRIE. However, when the mask layout defines an acute angle as shown in Fig. 1a, two scalloped sidewalls meet at the corner of the structure and form relatively sharp tips. Therefore, with a simple layout as shown in Fig. 1a and a single etch of an SOI wafer down to the oxide etch-stop, a field-emission diode is obtained with dozens of tips vertically stacked at submicron distance. Varying etch and deposition time in the etch cycle results in varying vertical tip density. Three of the recipes used were characterized in a SEM and yielded the results listed in Table 1. The highest vertical tip density achieved to date is ~3 tips/micron. When mask layout includes many such devices at small lateral spacing (e.g. in Fig. 12,) extremely high overall tip density is achieved. If, for example, a 50 μm thick SOI wafer is used each edge includes ~150 tips, and since there would be 5000 such edges/cm length (at 2 μm spacing), a total of 750,000 tips would result. This methodology, based on present achievements can therefore fabricate $150\cdot10^6$ tips/cm^2. We believe that further optimization of the vertical etch process and most recent photolithography can far exceed this estimate.

Tip depth is a function of layout tip sharpness. Even with high density recipe, tip depths in excess of 900 nm have been achieved. Radii of tips following DRIE vary between 15 nm and 30 nm. While apparently sharp, such tips have turn-on voltages well over 100 V rendering them inadequate for low-power FED applications. Images of tips formed with each of the three recipes are shown in Figure 2. All three SEMs were taken at a magnification of 40k.

For a given recipe, tip density and depth varies visibly down the device. In a final device, this affects the cathode-anode distance and consequently the turn-on voltage of individual tips. The variation is up to the order of 100 nm, and its overall effect is still being investigated. We believe, however, that when devices are operated in the high voltage regime all tips contribute
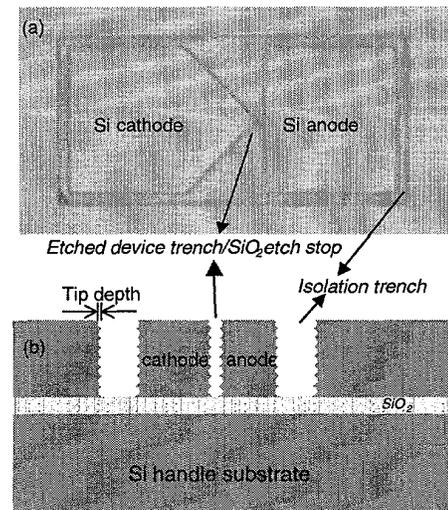
Figure 1. Fabrication of LFEDs by DRIE etching of SOI wafers: (a) microphotograph of the top view of a diode test structure, and (b) schematic cross-section of the SOI wafer.

DRIE tip recipe	SF₆ etch time [s]	C₄F₈ dep. Time [s]	Vertical tip dens. [μm⁻¹]	Tip Depth [nm]
High density	6	5	2.9 - 3.1	350 ± 70
Std. density	9	7	2.0 - 2.2	510 ± 50
Low density	14.3	11	1.3 - 1.5	700 ± 100

Table 1. Details of different DRIE tip recipes. In all recipes, the overlap time between etch and deposition is 0.5 seconds. Quantities are for layout of 60 degree tips. In all cases, coil power=600 W, platen power=120W, and SF_6 flow=130 sccm.
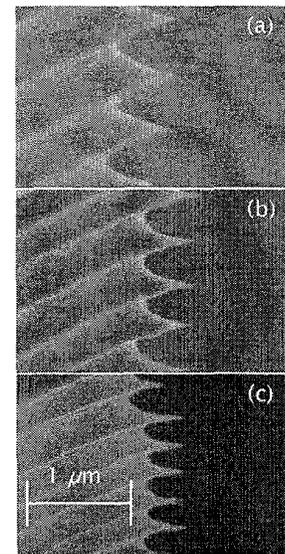
Figure 2. Three SEM images taken at the same magnification: (a) low density tips, (b) standard density tips and (c) high density tips.

to the total current nearly equally.

Depending on the mask layout, two types of geometry can result after the etch. In one case, as in Fig. 2, the anode is separated from the cathode in the layout and does not affect the formation of the tips. The electrode distance in that case is defined photolithographically. In the case where there is a slight (<0.5 μm) overlap in the mask, the etch results with the form shown in Figure 3, where cathode and anode electrodes are not separated, but connected via silicon nanowires. Oxidation and subsequent oxide removal, as detailed in Sec. IIIB, results in self-aligned cathode and anode where both sides are lined with sharp
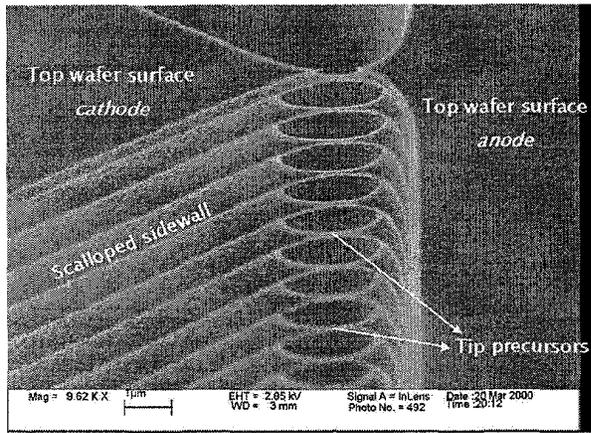
Figure 3. Silicon nanowires extend between blocks overlapping in layout. DRIE alone does not result in tips with submicron spacing.

tips at a distance which is a function of the oxidation process.

B. Oxidation sharpening

To increase the electric field at the cathode surface, the tips are sharpened by thermal oxidation [14], i.e. by repeatedly growing and removing (wet etching) a thermal oxide at temperatures not exceeding 1000 °C. Since tip radii after DRIE are already <30 nm, wet oxidation of ~700Å is sufficient to achieve atomically sharp tips. Due to the stress-dependent nature of oxidation, tip depth is relatively unaffected.

Additionally, since oxidation consumes silicon, the oxidation sharpening cycle can be exploited to separate the aforementioned self-aligned tip geometries. This is similar to the methodology in [9]. The nanowires in Fig. 3 become symmetrically sharpened tips on the cathode and anode in Fig. 4. Not only does this result in electrode distances closer than those possible through lithographic definition and etch aspect ratio alone, but sharpened tips on both ends of the gap effectively double the electric field. In terms of (3), this doubles parameter β for the same tip sharpness resulting in a fourfold increase in current.

Devices fabricated via above methodology were characterized in a vacuum probing station with an HP4145 *IV* curve tracer. Results of a typical device from a 22 μm SOI wafer are shown in Fig. 5. The diode shows hundreds of microamps of current at relatively low voltages. To our knowledge, this is the highest achieved current from a single silicon FED. The data was fit to (1) extracting $a=4.76 \cdot 10^{-6}$ and $b=247$.

C. Reducing distance by VHARM methodology

The above methodology does not provide a means of reducing electrode separation below 0.5 μm. Further reduction of the cathode-anode gap is possible using the VHARM methodology [15] by depositing polysilicon over the sharpened and aligned tips. Deposition thickness can be accurately controlled to nanometer distances allowing for gap distances that greatly surpass lithographic and etching limits. Following DRIE, approximately 3000 Å of doped n-type polysilicon is deposited in an LPCVD furnace. Oxidation sharpens the tips and a blanket RIE etch removes the oxide and polysilicon from the top and bottom of the electrode structures. A typical result of VHARM is shown in Fig. 6. While the tips are extremely close, there is a tip radius penalty associated with the polysilicon deposition. Additionally, there is less tip uniformity following the deposition and further sharpening. Tip radius was further decreased with the same sharpening cycle as detailed previously (Sec. IIB).
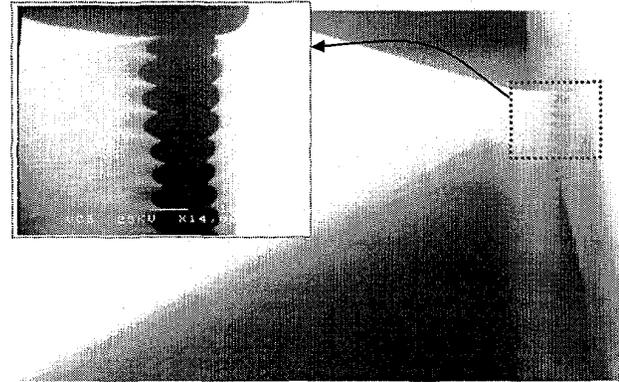
Figure 4. SEM images of an LFED after DRIE etching and oxidation sharpening, showing multiple tips along the vertical edge. The main image at 2,500 x, and the inset at 15,000x magnification.
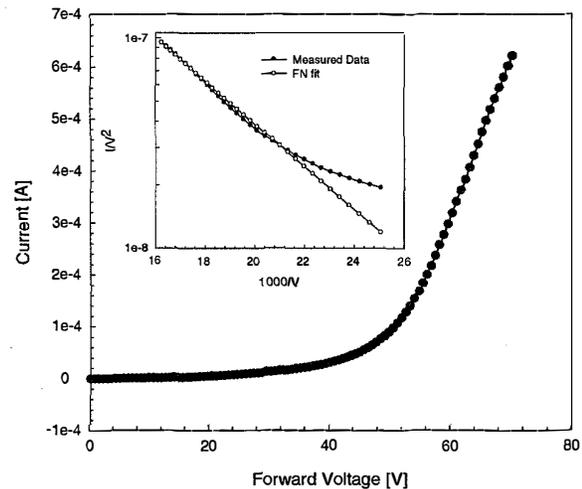
Figure 5. Measured current-voltage characteristics of a typical diode. Cathode-anode separation of 2.2 μm separation was measured on a SEM.
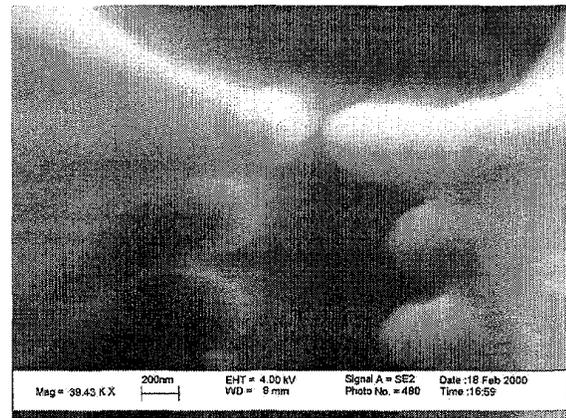
Figure 6. Electrode tips following VHARM processing and one oxidation sharpening step. While tip radius is decreased, distance between the top tips is reduced to less than 50nm.

Large vertical non-uniformity is apparent with tips of differing depth (Fig. 7).

Emission has been observed following VHARM and repeated oxidation sharpening similar to the results in previous methodology. In one device, current of 618 μA was observed at
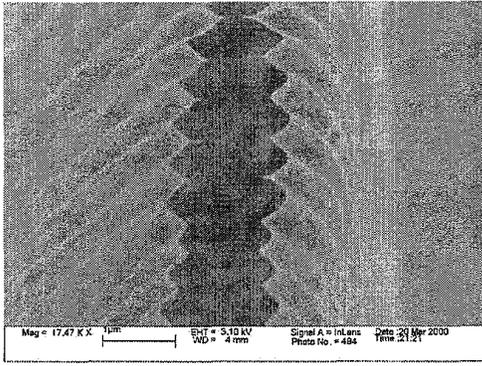
Figure 7. Multiple oxidation sharpening cycles restore some sharpness to the tips following VHARM.

50 V, and after FN fitting, parameters $a=1.41 \cdot 10^{-5}$ and $b=210$ were extracted. That device was fabricated on a 22 μm SOI wafer using the standard density recipe (~45 tips.)

Assuming a current of 1 μA/tip for reliable device operation, and using the estimated tip density from Sec. IIIA, we estimate a total of 150 A/cm^2.

D. Submicron self-alignment of electrodes by oxidation spacing

Minor advantages and a large increase in non-uniformity in resulting devices render the VHARM methodology inadequate for our goals. Another approach was taken utilizing micromachining technology to preserve sharpness while maintaining a constant inter-electrode distance through the vertical length of the device. Moreover, this new approach offers arbitrary cathode-anode spacing. Namely, after DRIE and the oxidation sharpening cycle of the SOI wafer, an additional low-temperature oxide layer is deposited to define the inter-electrode spacing without removal of the thermal oxide. This is followed by an n-type doped polysilicon deposition which covers the top, the sidewalls and trenches surrounding the devices. Subsequent chemical mechanical polishing removes the polysilicon from the top of the devices to expose the underlying single crystal silicon (SCS) blocks. Over-polishing through the polysilicon, oxide and first micron of SCS can be exploited to remove the upper tips of the device which are often inconsistent. An HF removal of the sidewall oxide between the sharpened SCS tips and the polysilicon walls completes this method of fabrication. Typical resulting devices of this process are shown in Figs. 8 and 9, where cathode-anode spacing was set to 4000 Å and 8000 Å, respectively.

This single-mask three-dimensional fabrication methodology allows for arbitrary cathode-anode spacing with the polysilicon walls acting as the anode and SCS block as the cathode. Maintaining cathode tip sharpness while reducing electrode distance results in decreased turn-on voltages. One measured device (Fig. 8b) with 8000 Å distance resulted in turn-on voltage of ~22 V. Furthermore, the cathode-anode distance is now constant over the height of the device at exactly the oxide spacing that should result in all tips having a similar turn-on voltage.

Appropriately choosing the geometry of a self-separating SCS diode yields a triode in this process with the polysilicon walls acting as the gate. Application of potential to the polysilicon modulates the cathode-anode current. Furthermore, a sufficiently high potential can be used to extract electrons from the cathode in an electron source configuration. With the cathode grounded, different extraction potentials are controlled
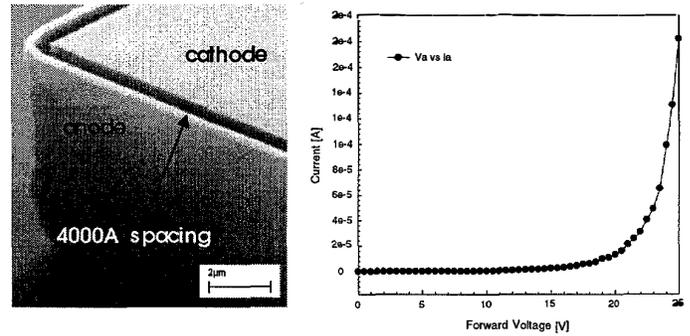
Figure 8. Diode with 4000 Å spacing between SCS cathode and poly anode (a) footing is apparent at the base of the structure, and (b) a similar device with 8000 Å spacing has a turn-on voltage of 22V.
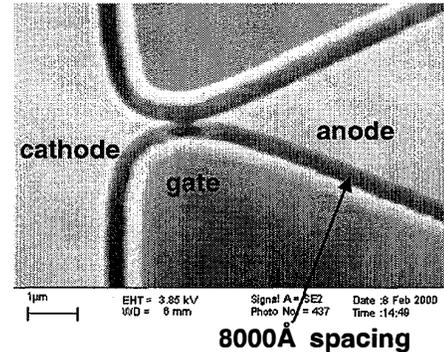
Figure 9. An SCS diode can operate as either a triode or an electron source using the polysilicon walls as gates.
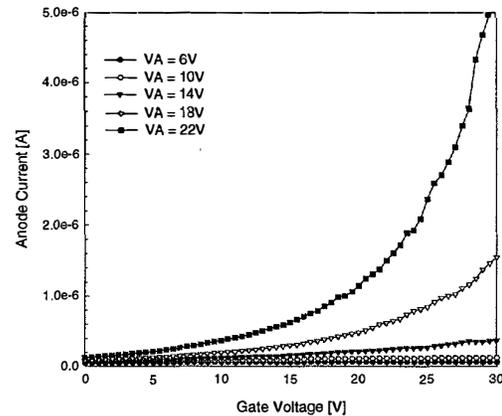
Figure 10. Triode operating in electron source mode – increased gate potential extracts more current from the cathode to the anode.

by the gate voltage as illustrated in Fig. 10.

Another attempt at controlling emission current, as in classical vacuum triodes, is with silicon microgrids fabricated in the same process. Grids are designed in layout as very thin and modulated photoresist patterns, such that the vertical scalloping effect is sufficient to periodically etch through the structure and create airgaps. The horizontal width modulation in layout results in the microgrid structure shown in Fig. 11. This same phenomenon can be observed in Fig. 3.

IV. MICROELECTROMECHANICAL FEDs (MEMFEDs)

MacDonald *et al* [17] have demonstrated utility of integration of cold cathodes with MEMS actuators. In their work, and that of Kenney *et al* [16], numerous advantages of
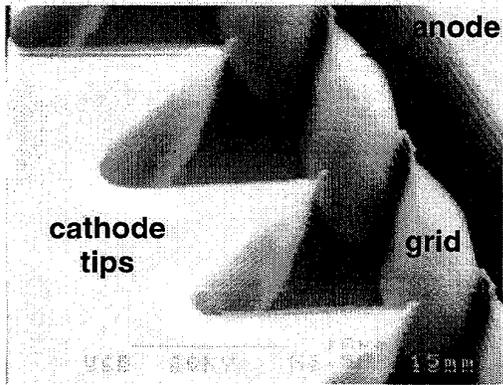
Figure 11. Multi-tip cathode and anode pictured with a microgrid in between to modulate emission current.
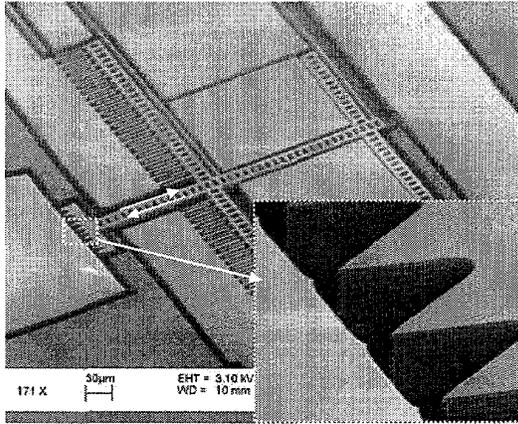
Figure 12. Microelectromechanical field emission diode with multiple columns of submicron spaced tips. The cathode-anode separation of ~3 μm can be adjusted with the comb-drive down to <0.5 μm.

Figure 13. Measured current-voltage characteristics of an electrostatically tunable field emission diode.

Figure 14. Triode with gate actuation in two dimensions using gap-closing actuators. The gates are controlled independently in both x and y to <0.5 μm relative to the fixed anode and cathode.

Figure 15. Accelerometer device with acceleration sensed by differential current change with gap distances.

sensors based on tunneling are discussed. Due to the lateral structure of our proposed devices, and high achievable currents, these FEDs lend themselves to numerous related applications. They are easily integrated with electrostatic actuators and fabricated in the same DRIE/oxidation process on SOI wafers. While the work is in initial stages, the result will be a variety of tunable diodes, triodes, and other vacuum devices. The example in Fig. 12 is a diode with ~600 atomically sharp tips with electrostatically adjustable cathode-anode distance. Preliminary results in Fig. 13 show an increase of diode current in this device with the application of actuation voltage to the comb-drive structure, which brings the cathode to a submicron distance from the anode. One goal of this research is a fully tunable RF-capable triode in which either placement of both gate electrodes (as shown in Fig. 14) or of all electrodes is adjustable electrically. Gap distances are varied electrostatically either with comb-drives or gap-closing actuators. In combination with methodology for arbitrary cathode-anode spacing discussed in Sec. IIID, effort is under way to develop highly sensitive and reliable accelerometers as in Fig. 15.

V. BATCH TRANSFER OF DEVICES

Fabrication of vacuum microelectronics is typically incompatible with other processes, e.g. CMOS, due to high-temperature oxidation sharpening. For most applications that may utilize the proposed MEMFED methodology it would be advantageous to have the capability of integration with CMOS drive/sense circuits via batch transfer of the former from separate wafers. We utilized the methodology of Maharbiz et al [12] on a variety of devices such as diodes, triodes, and electron sources.

The devices were fabricated in the following four-mask process: starting with a 100 mm silicon wafer (donor wafer), 1.8 μm of low-stress nitride is deposited and patterned as the structural material over 2 μm of sacrificial oxide. This structural pattern was used to etch the sacrificial oxide as well to expose the wafer surface outside of device areas. Next, 4 μm of doped polysilicon was deposited. This polysilicon material was

328
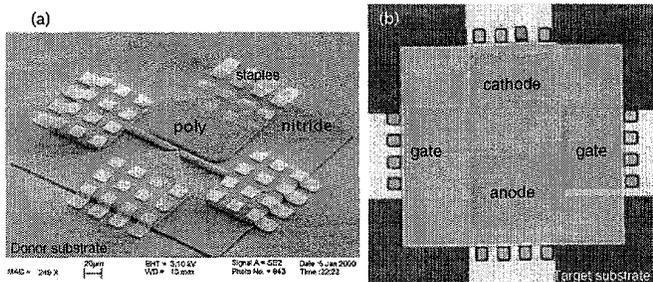
Figure 16. Poly triode fabricated for batch transfer to other substrates: (a) on donor substrate, released with gold staples holding it in place, and (b) after transfer to quartz target substrate, upside down; electrodes still visible through the nitride.
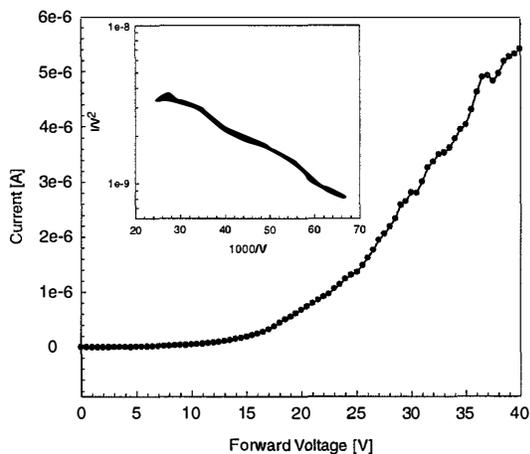
Figure 17. Transferred polysilicon diode performance.

subsequently patterned and etched (DRIE) with a "FED" mask to achieve diodes and triodes over the structural nitride islands.

After cycles of oxidation sharpening, as described in Sec. IIIB, the thermal oxide on top surfaces was removed in a blanket etch. At this point the device fabrication is complete and bumps are added for release and bonding. A 150 Å/600 Å Cr/Au layer is evaporated as a seed layer; 6-8 μm gold bumps are then electroplated in a resist mold. The donor wafers are then diced. After resist-stripping, the structures on donor wafers are then released in concentrated HF for ~4 minutes. The released structures remain on the individual chips due to the small gold bump *staples* [12], which were plated over the structures and surrounding silicon as shown in Fig. 16a. At that point, the devices are batch-transferred using a flip-chip bonder after careful alignment of *donor* and *target* chips, as illustrated in Fig. 16b. Target quartz wafer for the demonstration was plated with gold and patterned with pads in areas where FEDs would be transferred to. Although the process is designed for wafer-level transfers, at this time only chip-level transfers were considered and implemented. The target and donor substrates are aligned and pressure is applied to form a gold-gold compression bond; the donor substrate is then removed, leaving the switches in place over the pads. We verified that the devices continue to operate, as shown on a measured transferred diode in Fig. 17. This emission curve is similar to those found on pre-transfer polysilicon diodes. Diodes fabricated in polysilicon are significantly less consistent in emission characteristics and tip shape than those fabricated in SCS suggesting that batch transfer methods should be explored for the latter material.

VI. CONCLUSIONS

The use of the scalloping formed in DRIE for sharp emission tips has been reported, resulting in achievable tip density of over $150 \cdot 10^6$ tips/cm^2, and current of over 150 A/cm^2. By utilizing sacrificial sidewall spacing, electrodes were placed at 4000Å. We further utilized MEMS actuators to laterally adjust electrode distances. To improve integration capability of the FEDs, we demonstrated batch bump-transfer integration of working FEDs onto a quartz target substrate.

VII. ACKNOWLEDGEMENTS

REFERENCES

[1] I. Brodie, P.R. Schwoebel, and H. Falk, "Vacuum Microelectronic Devices," Proceedings of the IEEE, vol. 82, no.7, pp.1005-34, July 1994.

[2] D. Temple, W. D. Palmer, L. N. Yadon, J. E. Mancusi, D. Vellenga, and G. E. McGuire, "Silicon Field Emitter Cathodes: Fabrication, Performance, and Applications," J. Vac. Sci. Technol. A, vol. 16, no. 3, pp. 1980-1990, May/June 1998.

[3] C. A. Spindt, C. E. Holland, A. Rosenberg, and I. Brodie, "Field Emitter Arrays for Vacuum Microelectronics," IEEE Tran. on Electron Devices, vol. 38, no. 10, pp. 2355-2363, Oct. 1991.

[4] C. A. Spindt, I. Brodie, L. Humphrey, E. R. Westerberg, "Physical properties of thin-film field emission cathodes with molybdenum cones", Journal of Applied Physics, vol.47, (no.12), Dec. 1976. pp. 5248-63.

[5] C. A. Spindt, C. E. Holland, P. R. Schwoebel, and I. Brodie, "Field-Emitter-Array Development for Microwave Applications," Proc. of Vacuum Microelectronics Conference, St. Petersburg, pp. 638-639, 1996.

[6] B. R. F. Kendall, E. Drubetsky, "Cold Cathode Gauges for Ultrahigh Vacuum Measurements," J. Vac. Sci. Technol. A, vol. 15, no. 3, pp. 740-746, May/June 1997.

[7] B. R. Johnson, A. I. Akinwande, D. Murphy, "Characterization of Lateral Thin-Film-Edge Field Emitter Arrays," Proc. of Vacuum Microelectronics Conference, St. Petersburg, pp. 663-667, 1996.

[8] C-M. Park, M-S. Lim, and M-K. Han, "A Novel In Situ Vacuum Encapsulated Lateral Field Emission Triode," IEEE Electron Device Letters, vol. 18, no. 11, pp. 538-540, Nov. 1997.

[9] J-H. Park, H-I. Lee, H-S. Tae, J-S. Huh, and J-H. Lee, "Lateral Field Emission Diodes Using SIMOX Wafer," IEEE Tran. On Electron Devices, vol. 44, no. 6, pp. 1018-1021, Jun. 1997.

[10] A. A. Ayon, R. Braff, C. C. Lin, H. H. Sawin, M. A. Schmidt, "Characterization of a time multiplexed inductively coupled plasma etcher," Journal of the Electrochemical Society, vol. 146, no. 1, pp. 339-49, Jan. 1999.

[11] R. Bosch Gmbh, patents 4855017 and 4784720 (USA), and 4241045C1 (Germany.)

[12] M. M. Maharbiz, M. B. Cohn, R. T. Howe, R. Horowitz, A. P. Pisano, "Batch Micropackaging by Compression-Bonded Wafer-Wafer Transfer", MEMS '99, Orlando, FL, January 17-21, 1999, pp. 482-9.

[13] R. H. Fowler and L. Nordheim, "Electron Emission in Intense Electric Fields," Proc. Royal Soc. (London), vol. A119, p. 173, May 1928.

[14] T. S. Ravi, R. B. Marcus, D. Liu, "Oxidation sharpening of silicon tips," Journal of Vacuum Science & Technology B, vol.9, no. 6, pp. 2733-2737, Nov.-Dec. 1991.

[15] J. M. Noworolski and M. Judy, "VHARM: Sub-Micrometer Electrostatic MEMS", Transducers '99, Sendai, Japan, June 7-10, 1999, pp.1482-1485.

[16] Kenny, T.W.; Waltman, S.B.; Reynolds, J.K.; Kaiser, W.J. A micromachined silicon electron tunneling sensor. Proceedings. IEEE Microelectromechanical systems, Napa Valley, CA, USA, pp. 192-6, Feb. 1990.

[17] W. Hofmann, L-Y Chen, N. C. MacDonald, "Fabrication of integrated micromachined electron guns," Journal of Vacuum Science &Technology B, vol.13, no. 6, pp. 2701-2704, Nov.-Dec. 1995.

Real-Time Etch-Depth Measurements of MEMS Devices

Sylvie Bosch-Charpenay, Jiazhan Xu, John Haigis, Peter A. Rosenthal and Peter R. Solomon
On-Line Technologies, Inc
East Hartford, CT 06108

James M. Bustillo
Berkeley Sensor & Actuator Center, University of California,
Dept. EECS, 497 Cory Hall, Berkeley, CA 94720-1770

ABSTRACT

An *in-situ*, real-time process control tool was developed for MEMS deep-reactive-ion-etch (DRIE) fabrication. DRIE processes are used to manufacture high-aspect-ratio silicon structures up to several hundred microns thick, which would be difficult or impossible to produce by other methods. DRIE MEMS technologies promise to deliver new devices with increased performance and functionality at lower cost. A major difficulty with DRIE is the control of etch depth. Our research shows that it is possible to monitor the etch depth of various MEMS structures (holes, pillars, trenches, etc.) through measurement and analysis of the infrared reflectance spectrum. Depths as large as 150 μm have been measured. Excellent correlation is found between the etch depths determined by analysis of these measurements and those measured with an SEM. In addition to etch depth, other parameters such as the photoresist thickness (e.g. mask erosion) can be simultaneously extracted. Based on these results, an infrared-reflectance etch monitor was integrated onto a reactive ion etcher at the Berkeley Sensor & Actuator Center for real-time monitoring and end-point determination. The integrated optical metrology system demonstrated accurate real-time monitoring of the etch depth and photoresist mask erosion.

INTRODUCTION

DRIE processes are used to manufacture high-aspect ratio MEMS structures up to several hundred microns thick (see Figure 1). The resulting three-dimensional structures typically have 25-50 times the thickness of most surface-micromachined devices [1]. DRIE methods have the potential to yield advantages such as: 1) an improvement in z-axis stiffness and robustness, z being the etch direction, 2) improved sensitivity, 3) increased robustness of devices to packaging stresses, 4) improvements in manufacturability, and 5) possibly lower cost. Examples of MEMS to benefit from DRIE high-aspect-ratio features include micro-inertial instruments (accelerometers, gyroscopes), for which the sensitivity can be increased by two orders-of-magnitude [2]; micro-disk-drive armatures, which can be produced with significantly lower cost and higher manufacturing throughput [3]; micro-assembly actuators, which have the required high out-of-plane and low in-plane stiffnesses [4], etc. Other DRIE applications include fabrication of SCALPEL mask membranes, requiring accurate and reproducible membrane thickness and buried-layer etch stops for thick silicon-on-insulator (SOI) designs.

While DRIE and other MEMS fabrication technologies promise new devices with increased performance and functionality at lower cost, one major difficulty is the control of etch depth during the DRIE process. Optical end-point detectors currently exist for DRIE of SOI wafers, but they require large open area test patterns (300 μm^2 area is typical) that do not etch at the same rate as the desired devices (which nominally have features sized in the micron range). For silicon (non-SOI) wafers, etch depth is measured using different systems depending on the pattern width, but always after the etch step is completed. Wide feature measurements are performed using profilometry or near-IR transmission measurements. For narrow structures (e.g., those less than ~ 100 μm-wide) the measurements are performed using *ex-situ* destructive SEM. The lack of nondestructive metrology for narrow MEMS structures, as well as the need for *in-situ* process control, are the driving forces motivating this work.
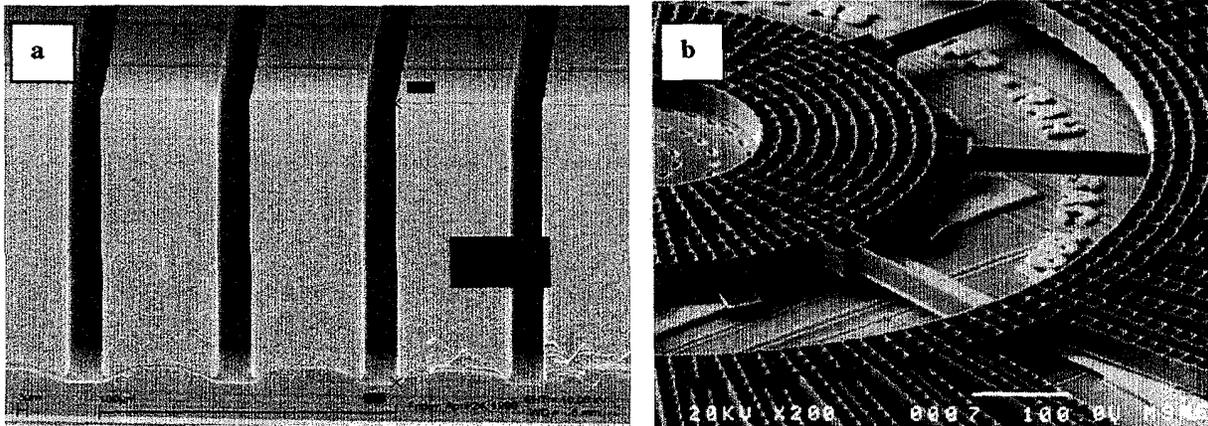
Figure 1. SEM photographs of a MEMS high-aspect-ratio, DRIE trench structure: a) cross-section of a grid of 10 μm-wide trenches whose depths have been measured by the FTIR method developed in this work, and which are the first step in the production of a 'Hexsil' micro-molded angular micro-actuator, b) finished 100 μm-thick angular actuator transferred and mounted onto a CMOS integrated circuit (Photos: Berkeley Sensor & Actuator Center, University of California, Berkeley).

PRINCIPLE OF OPTICAL ETCH-DEPTH MEASUREMENTS

When light reflects from a patterned and etched feature, interference occurs between those wavefronts returning from the top and the bottom of the feature. The resulting interference pattern contains information about the etch depth as well as the patterning masking layer. The analysis of such interference patterns in the visible spectrum has been used for the characterization of IC devices [5,6,7]. We demonstrate that it is possible to make similar analyses with an infrared reflectance spectrum to measure the etch depth of MEMS trench structures.

Infrared reflectance of patterned MEMS structures - Figure 2 shows the infrared reflectance of MEMS 'Hexsil' structures with 10 μm-wide trenches forming an x-y grid (similar to the structure shown in Figure 1a); the unetched areas are covered with photoresist. The reflectance in the spectral region 2-10 μm (1000-5000 cm^{-1}) was taken using an On-Line Technologies FTIR (Fourier Transformed Infrared) spectrometer at an incident angle of 4° from normal, a resolution of 16 cm^{-1} and a measurement time of <1 sec. Two types of interference fringes are evident in the spectrum; the large low-frequency fringe is caused by interference in the photoresist, while the small superimposed high-frequency fringes are caused by interference in the trenches.
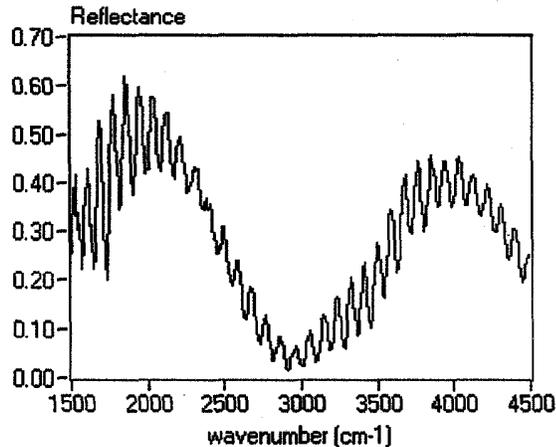
Figure 2. *Infrared reflectance of MEMS 'Hexsil' structures with 10 μm-wide trenches forming an x-y grid. Unetched areas are covered with photoresist. The large, low frequency fringe is caused by interference in the photoresist, while the small, high frequency superimposed fringes are caused by interference in the trenches.*

In order to extract parameters from the reflectance spectra of an etched structure, analysis models need to account for: 1) pattern effects, resulting in interference between wavefronts being reflected from different film stacks, and 2) thin film effects, resulting in interference between wavefronts reflected at various films interfaces.

Frequency analysis - For the case of a single material with etched regions (e.g. features etched in silicon without the photoresist masking layer), one can account for pattern effects by simply calculating the path difference, Δx, between the wavefront reflected by the unetched regions and the wavefront reflected by the etched regions. This relation can be expressed as: $\Delta x = 4\pi h \nu/\cos\theta$, where ν = wavenumber of probing light (ν = 1/wavelength), h = trench depth, and θ = angle of incidence. Constructive interference will occur for two consecutive wavenumbers ν_1 and ν_2 when the path difference is equal to 2π:

$$4\pi h\nu_1/\cos\theta = 4\pi h\nu_2/\cos\theta + 2\pi \Rightarrow (\nu_1 - \nu_2) = \cos\theta/2h$$
$$\Leftrightarrow \Delta\nu = \cos\theta/2h \Leftrightarrow F_{\text{trenches}} = 2h/\cos\theta \quad \text{(Equation 1)}$$

with F_{trenches} = fringe frequency in the reflectance spectrum. Equation 1 indicates a linear relation between fringe frequency and trench depth. A similar approach can be used to account for thin film effects and extract the thickness of a blanket film on a substrate [8]. In that case, the path difference is between the beams reflecting from the top and the bottom of the film, and the effect of the index of refraction of the film needs to be accounted for, which yields the following relationship for the fringe frequency F_{film}:

$$F_{\text{film}} = 2zn\cos\theta_t \quad \text{(Equation 2)}$$

with z = film thickness, n = film index of refraction, and θ_t is the refracted angle with respect to normal where $\sin\theta_t = (\sin\theta)/n$.

In the case of MEMS structures, photoresist is present as a masking layer, and both pattern effects and thin film effects are present. A full spectrum fit (as opposed to a simpler frequency analysis) is usually required to simulate the compounding interference effects and extract etch-depth and film thickness [9]. However, in this case the interference patterns due to the photoresist film and that due to the trenches can be easily separated because they yield fringes of very different frequencies (see Figure 2). The different frequencies result from the very different values for trench depth (typically few tens of μm's), and photoresist thickness (typically 1-2 μm). As a consequence, Equation 1 was used to extract the etch depth using the high frequency fringes, and Equation 2 was used to extract the photoresist thickness using the low frequency fringes. A more complete analysis including full spectrum fitting (useful for more complex film stacks) is given elsewhere [9].

Test of analysis model - Correlation between the etch depth as-extracted by the FTIR method and measured by SEM were obtained for various structures and are displayed in Figures 3 and 4. Figure 3 shows the FTIR vs. SEM etch depth correlation for two types of MEMS trench structures, DRIE orthogonally-patterned 'Hexsil' micro-mold trenches 10 μm-wide and parallel trenches with lines and spaces 20 μm-wide. As seen in the figure, the FTIR-extracted etch depth was found to agree very well with SEM measurements. The residual photoresist thickness is also plotted as a function of etch depth to show resist erosion (note: the photoresist thickness below ~1 μm is not plotted as the extracted thickness is not sufficiently reliable from the frequency analysis of Equation 2).
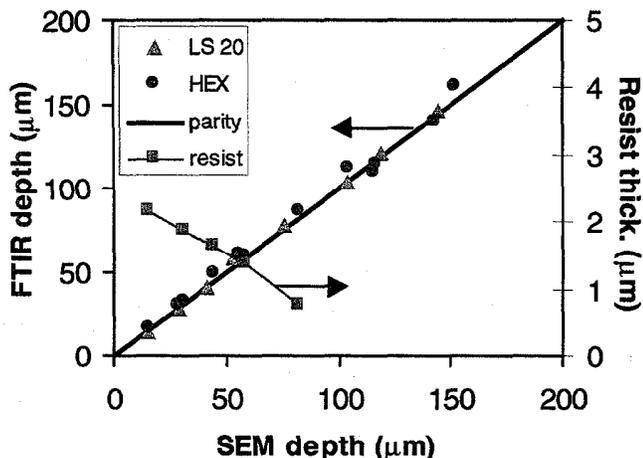
Figure 3. Correlation between the etch depth as extracted by the FTIR method to that measured by SEM. left axis: depths of orthogonally patterned 10 μm-wide 'Hexsil' micro-mold trenches and 20 μm-wide parallel trenches (lines and spaces), right axis: erosion of 2.5 μm thick photoresist masking-layer thickness extracted by FTIR.

Influence of trench width - Figure 4 shows the FTIR vs. SEM depth correlation for etched parallel trenches with line and space widths of 4, 8, and 20 μm. As can be seen from the figure, the model applies linearly, as predicted, for large features (20 μm), but requires an offset for smaller features (4 and 8 μm) - the offset magnitude depending on the feature width. This offset is thought to be due to trench edge effects, which may cause slightly different phase shifts, and results from the model limitations in the case of patterned features whose size are similar to the probing wavelength [10]. These limitations can only be eliminated by the use of
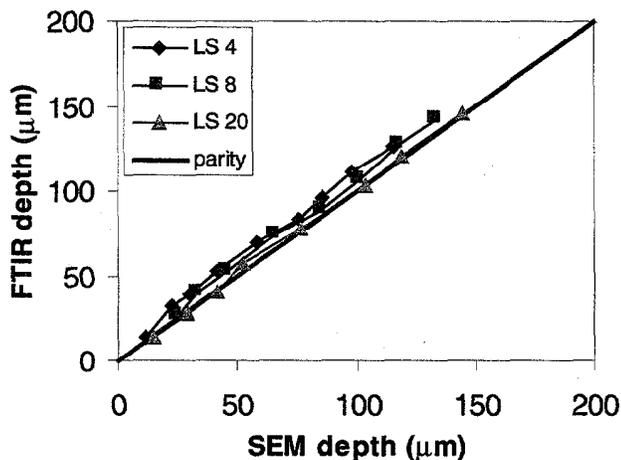
Figure 4. Correlation between the etch depth as extracted by the FTIR method to that measured by SEM for trench structures whose line/space (LS) width were 4, 8, and 20 μm respectively. A trench-width dependent offset is observed for smaller trench widths.

comprehensive electromagnetic vector-diffraction models for the exact simulation of the reflectance spectrum. Vector diffraction models solve the rigorous coupled-wave solution in that they account for all the interfering wavefronts [10]. The drawback of these models is that they are computationally intensive and usually require supercomputers or high-powered workstations [5], in contrast with the frequency model used in this work which has an extremely low computational load. Also, the vector diffraction models have been applied, so far, only to simple geometries which may not be relevant to MEMS applications.

The trench width limit at which the simple frequency model cannot be applied directly appears to be between 10 and 20 μm, as below these widths the correlation curves are progressively offset. However, the model can still provide good correlation with SEM, allowing an accurate determination of etch depth for narrow structures (4 μm features being the smallest analyzed with the FTIR probe) when the offset is determined through a calibration procedure derived directly from data such as in Figure 4. As a consequence, the frequency model appears to be a satisfactory approach to MEMS etch depth extraction.

Influence of structures shape - The model has also been successfully applied to various MEMS structures having different geometrical patterns (holes, pillars, trenches, trench grids, honeycombs, etc.). The most complicated case was found to be that of parallel trenches (i.e., trenches organized as a diffraction grating), because the fringe frequency varied throughout the analyzed infrared spectrum. This behavior is consistent with the behavior of diffraction gratings, where it was found that the 0th-order diffraction efficiency varies with wavelength in an oscillating manner, and can only be accurately simulated by vector diffraction analysis [10]. Other geometrical patterns did not display this behavior, presumably because of their more isotropic arrangement (e.g. orthogonally-patterned 'Hexsil' micro-molds). In spite of these 0th-order deficiencies it was observed that even for the parallel trench samples the fringe frequency, averaged over the whole IR spectrum, correlated very well with the etch depth measured by SEM (see Figure 4) - still allowing accurate etch depth extraction.

In-Situ, Real-Time Etch-Depth Measurements

The FTIR sensor has been integrated on a Surface Technology Systems (STS) Multiplex ICP deep reactive ion etcher at the Berkeley Sensor & Actuator Center at the University of California at Berkeley. Real-time measurements of etch depth and photoresist thickness (e.g. photoresist mask erosion during etch) were obtained on MEMS structures (as in Figure 1a). Etch depth as a function of etch time for this structure is shown in Figure 5. The etch process used was the switched Advanced Silicon Etch (ASE) process [11]. The models from Equation 1 and Equation 2 were used to extract the etch depth (using FTIR analysis of the fast-varying superimposed fringes), and the eroding photoresist thickness (using FTIR analysis of the slow-varying fringes). The single-point measurement time, which includes the FTIR reflectance measurement and analysis, is approximately 1 sec. The measurements are precise enough (when measuring depth) to see the sequential "passivating" and "etching" steps that are characteristic of the switched ASE process. The result is clearly
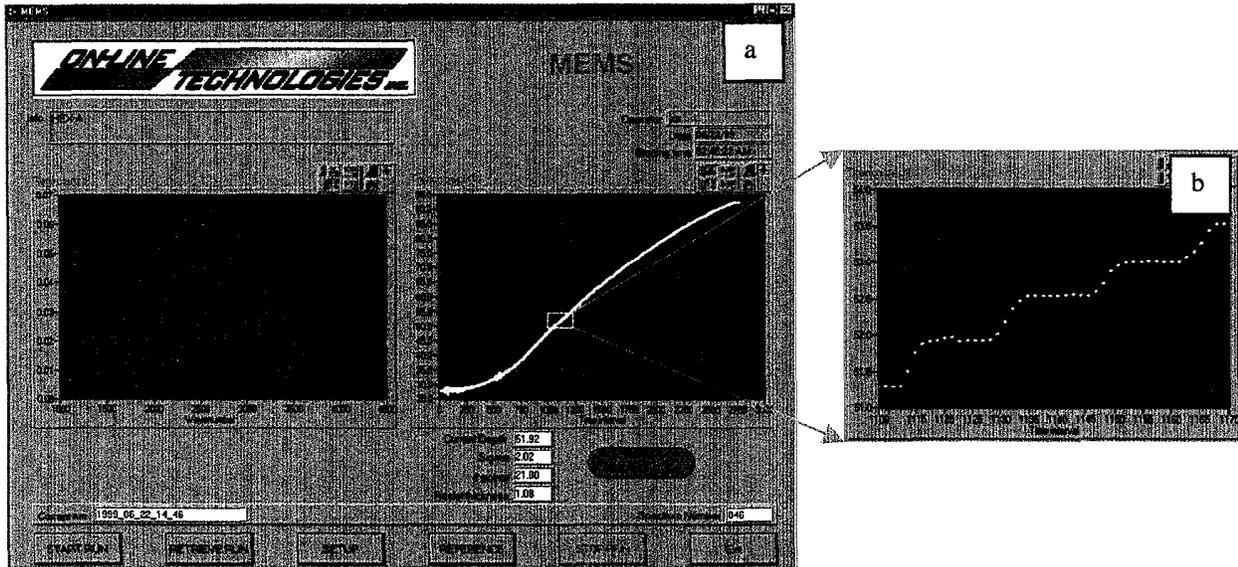
Figure 5. User interface with real-time data displayed for an ASE-etched MEMS with 10-µm wide trenches. a) left: IR reflectance versus wavenumber (ν = 1/wavelength) in cm⁻¹, and right: etch depth (µm) versus time (sec.) for a 92 µm trench. b) etch depth versus time with an expanded scale, clearly displaying the sequential etch and passivation layer deposition steps typical of the ASE process. The typical switching period is 10-20 sec.

Table 1. Summary of the sensor's capabilities.

MEMS structures which can be measured	Pillars, holes, trenches, trench grids, honeycombs, of etched width less than ~ 100 µm
Etch depth precision (reproducibility)	0.05-0.5 µm depending on the structure
Etch depth accuracy (as compared to SEM)	~ 1-2 µm (estimated)
Maximum measurable etch depth	> 200 µm
Current spot size [1]	~ 800 µm DIA
Measurement time	~ 1 sec.
Measurement spot range [2]	± 5 mm in both x and y direction
Probing IR wavelength	2-10 µm (1000 to 5000 cm⁻¹)

defined staircase-like etch-depth vs. time profile. Table 1 displays a summary of the integrated optical metrology system's capabilities.

CONCLUSIONS

In this work *in-situ* real-time DRIE etch depth and eroding photoresist thickness measurements were obtained on MEMS devices using infrared reflectance analysis. To our knowledge, this is the first time that these measurements have been made while DRIE plasma-etching MEMS structures. The method is non-destructive and provides real-time process analysis. Better real-time *in-situ* monitoring of MEMS processing will allow improved process accuracy and controllable process repeatability. The specific anticipated benefits of the new technique are: 1) to reduce cost through the reduction of destructive measurements and the improvement in process control, 2) to increase the reproducibility

of the MEMS structures through better process control, 3) to provide useful feed-back for process development, thus reducing development time.

ACKNOWLEDGEMENTS

---

[1] *Optical diffraction limits preclude spot sizes smaller than ~ 400 µm. MEMS devices smaller than this minimum spot size can however still be measured provided they are geometrically isolated from other DRIE etched features.*
[2] *Translation range of infrared spot over sample is strictly related to equipment geometry at the etch chamber – in this case a small overhead viewport into the STS Multiplex ICP system.*

REFERENCES

[1] J. Bhardwaj, H. Ashraf, A. McQuarrie, "Dry silicon etching for MEMS", *Proc. of the Third International Symposium on Microstructures and Microfabricated Systems*, Montreal, Canada, (1997), pp. 118-130.

[2] T. Juneau, M.A. Lemkin, T.A. Roessig, W.A. Clarck, R.T. Howe, J.M. Bustillo, T. Brosnihan, A.P. Pisano, "Commercialization of precision inertial sensors with integrated signal conditioning", *Sensors Expo*, May (1998).

[3] D.A. Horsley, A. Singh, A.P. Pisano, R. Horowitz, "Angular micropositioner for disk drives", *Proc. of the IEEE Micro-Electro-Mechanical Systems Workshop*, Nagoya, Japan, (1997), pp. 454.

[4] C.G. Keller, R.T. Howe, "HexSil tweezers for teleoperated micro-assembly", *Proc. of the IEEE Micro-Electro-Mechanical Systems Workshop*, Nagoya, Japan, (1997), pp. 72.

[5] M.E. Lee, C. Galarza, W. Kon, W. Sun, F.L. Terry, Jr., "Analysis of reflectometry and ellipsometry data from patterned structures", *Characterization and metrology for ULSI Technology: 1998 International Conference*, Seiler D.G. et al. Editors, the American Institute of Physics, Pub., (1998).

[6] P.A. Heimann, R.J. Schultz, "Optical etch-rate monitoring: computer simulation of reflectance", *J. Electrochem. Soc.: Solid State Science and Technology*, Vol. 131. No. 4 , (1984).

[7] M. Haverlag, G.S. Oehrlein, "In-situ ellipsometry and reflectometry during etching of patterned surfaces: experiments and simulations", *J. Vac. Sci. Technol.*, B 10 6, (1992).

[8] E. Hecht,. "Optics", Addison-Wesley Pub. (1987).

[9] S. Bosch-Charpenay, J. Xu, J. Haigis, P.A. Rosenthal, P.R. Solomon, and J.M. Bustillo, "Real-Time Etch-Depth Measurements of MEMS Devices" submitted to the *IEEE Journal of Microelectromechanical Systems* (2000).

[10] D.A. Gremaux, and N.C. Gallagher, "Limits of scalar diffraction theory for conducting gratings", *Applied Optics*, Vol. 32, No. 11 (1993).

[11] F. Larmer, A. Schilp, "Method of Anisotropically Etching Silicon", German Patent DE4241045.

IN-SITU MICROPLASMAS FOR RAPID DRY ETCHING OF SILICON

Chester G. Wilson and Yogesh B. Gianchandani[*]
Department of Electrical and Computer Engineering
University of Wisconsin, Madison

ABSTRACT

This paper reports on the generation of spatially confined plasmas and their application to silicon etching. The etching is performed using SF_6 gas and DC power applied between thin-film electrodes patterned on the silicon wafer to be etched. The typical operating pressure and power density are in the range of 1-20 Torr and 1-10 W/cm^2, respectively. The plasma confinement can be varied from <100 μm to >1 cm by varying the electrode area, operating pressure, and power. High power densities can be achieved at moderate currents because the electrode areas are small. Etch rates of 4-17 μm/min., which enable through-wafer etching and varying degrees of anisotropy, have been achieved. The etch rate increases with power density, whereas the etch rate per unit power density increases with operating pressure. Plasma resistance measurements and electric field modeling are used to provide an initial assessment of the microplasmas.

I. INTRODUCTION

Plasma processing is extensively utilized in semiconductor processing applications, and is the dominant silicon etching technique. Silicon etching is commonly performed in parallel plate reactors by applying RF power (typically at 13.56 MHz) between two electrodes placed several centimeters apart. The silicon wafer is located on the powered electrode for reactive ion etching. The operating pressure and power are in the range of 10-500 mTorr and 10-500 mW/cm^2, respectively. Since the plasma exists globally across the wafer, the etch is selectively masked by a thin film of photoresist, SiO_2, or metal which is patterned on the wafer surface. More recently, fast anisotropic etches have been demonstrated by alternative plasma etchers utilizing electron cyclotron resonance (ECR) [1] and inductively coupled plasmas (ICP) [2]. All of these options, however, employ a single plasma that acts over the entire surface area of a wafer. Creating several different etch depths or profiles in a single die mandates the use of a like number of masking steps.

In this effort we report on silicon etching using in-situ microplasmas. The microplasmas are powered by two metal electrodes separated by a dielectric in a tri-layer stack on the wafer surface (Fig. 1). The power density is approximately 100x higher than in conventional plasmas. In addition, DC power is used to etch silicon substrates. This eliminates the need for matching impedance networks

to associated with RF driven plasmas. Plasma confinement can be varied from a few tens of microns to more than a centimeter by changing operating conditions. The etch dimensions can also be confined to the openings in the mask, as with conventional etchers. Localized through-wafer etching has been demonstrated with this technique, and etch rates greater than 17 μm/min. have been recorded using SF_6.
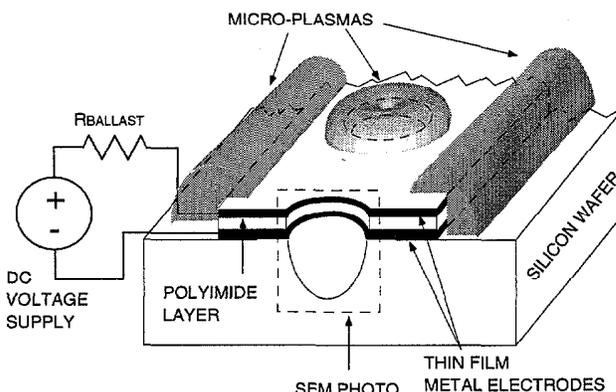
Fig. 1: Schematic illustrating the generation of an in-situ microplasma.

II. EXPERIMENTAL RESULTS

Test samples are fabricated by a two-mask process sequence. A metal-polyimide-metal electrode stack is deposited on the wafer. The first lithography mask is used to pattern the upper metal layer and the polyimide. The second mask is then used to pattern the lower metal layer and simultaneously re-pattern the upper metal as well. Metal layers that have been investigated include chrome, aluminum, titanium, and nickel. Following the microplasma etch, the electrode stack can be stripped by sacrificing the lower metal electrode in the appropriate wet etchant.

In the electrical set-up, the pads for the two in-situ electrodes are contacted by probes and connected to a DC power supply across a series ballast resistor, which provides a means to control the plasma current. The lower electrode is connected to the negative terminal of the supply, as shown in Fig. 1, in order to direct the positive ions in the plasma toward the Si substrate. Typical bias values range from 300-600 V, depending on the ambient gas used and target etch rate. The silicon is shielded from the electric field by the electrode stack itself. The relatively small

[*]Corresponding author: 1415 Engineering Drive, Madison, WI 53706-1691; Tel: (608) 262-2233, Fax: 262-1267, E-mail: yogesh@engr.wisc.edu

electrode areas for in-situ microplasmas allow power densities in the range of 1-10 W/cm² without drawing high currents. The relatively large operating pressures of 1-20 Torr serve to spatially confine the plasma. This allows several microplasmas with different etch characteristics to operate simultaneously on a wafer.

The spatial self-confinement of microplasmas is shown in Fig. 2 as a glow that is localized to the in-situ electrodes. In this experiment, the ambient gas was N_2, and the bias voltage was -360 V, which has opposite polarity to the bias normally used for etching. By varying the power and pressure the confinement can be changed from <100 μm to >1 cm.

Typical etch profiles achieved by microplasmas are shown in Figs. 3-5. In each of these cases the etch was performed though openings in the in-situ electrodes using a partial pressure of SF_6. Figure 3(a) shows a 92 μm deep etch through a circular opening of 150 μm diameter, that was achieved in 20 min. at 2.7 Torr, with a power density of 3.2 W/cm² averaged over the electrode area. The sidewall angle was 21° off vertical. The electrode metal used in this case is aluminum. Figure 3(b) shows a 233 μm deep etch that was achieved in 50 min. using the same parameters. Figure 4 shows a 33 μm deep etch through a 50 μm diameter opening achieved in 3 min. at 2.9 Torr and 7.3 W/cm². The sidewall angle is nearly vertical in certain locations of the profile. The electrode metal for this case was also aluminum. Figure 5 shows the cross-section of a 207 μm deep etch through a 280 μm wide, 2.2 mm long slit opening. The electrode metal used in this case was titanium. The etch was achieved in 24 min. at 5.2 Torr and 6.8 W/cm². The sidewall angle of this etch is also nearly vertical in certain locations of the profile. The profiles shown in Figs. 3-5 indicate that varying degrees of anisotropy can be achieved by changing the operating conditions of microplasmas.

The etch rate of Si in SF_6 microplasmas was studied as a function of several operating parameters (Figs. 6-8). Figure 6 shows the dependence of the etch depth and the cumulative average etch rate on the duration of the etch. This etches were performed under two sets of conditions: case A used 2.7 Torr pressure, 1.6 W/cm² power density, and aluminum electrodes with circular openings of 350 μm diameter; case B used 5.2 Torr, 6.8 W/cm², and titanium electrodes with 280 μm wide slit openings. Case A achieved etch rates of 4-7 μm/min., whereas Case B achieved etch rates of 9-12 μm/min. In both cases the etch rate was significantly higher for the first few minutes, and then rapidly settled at a lower value that was stable for 50 minutes. Although the reasons for the higher initial etch rate remain to be determined, it is noteworthy that the electric field above the exposed Si is highest for the first few minutes of an etch, during the time that the etched depth is relatively small. This is shown as a modeling result in the next section.

Figure 6(a) also indicates that under the conditions of case B, it was possible to etch though a wafer in less than

an hour. Through-wafer etches were routinely achieved with the use of titanium electrodes, which developed considerably less damage than aluminum from sputtering in exposed regions of the cathode. This is consistent with trends seen in conventional etchers [3].
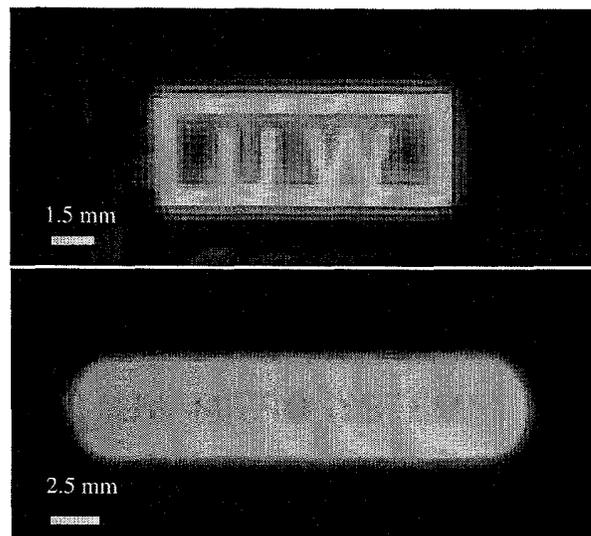
Fig. 2: Local confinement of N_2 microplasmas at 1-20 Torr and -360 V DC. (a-upper): Narrow electrodes with wide gaps; (b-lower): Wide electrodes with narrow gaps.
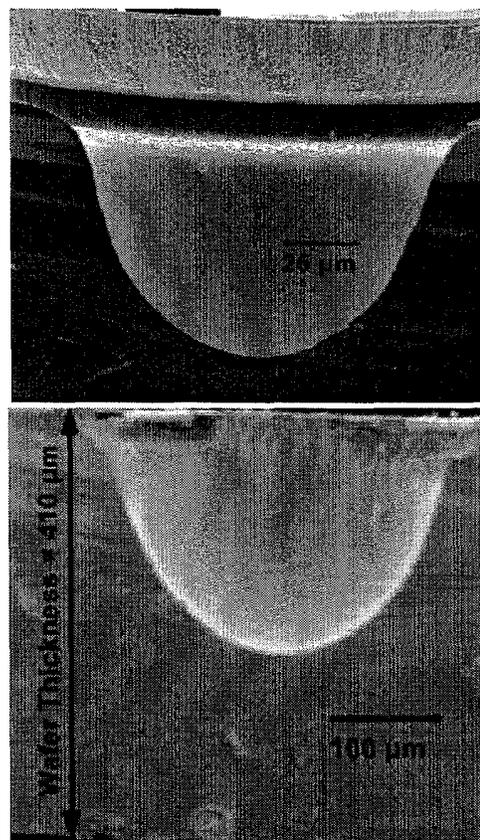
Fig. 3 (a-upper): A 92 μm deep etch through a 150 μm ø circular opening (2.7 Torr, 3.2 W/cm², 20 min.). (b-lower): A 233 μm deep etch (50 min.).
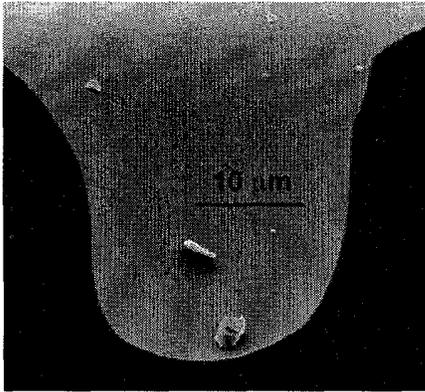
Fig 4: A 33 μm deep etch through a 50 μm ⌀ opening (2.9 Torr, 7.3 W/cm², 3 min.).
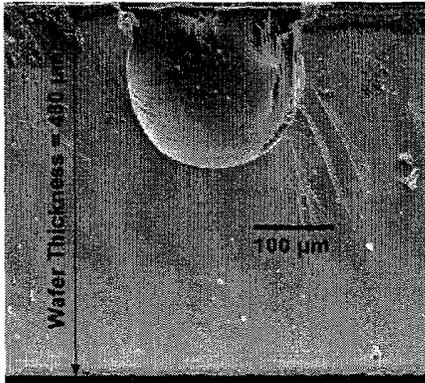
Fig 5: A 207 μm deep etch through a 280 μm wide slit opening (5.2 Torr, 6.8 W/cm², 24 min.).
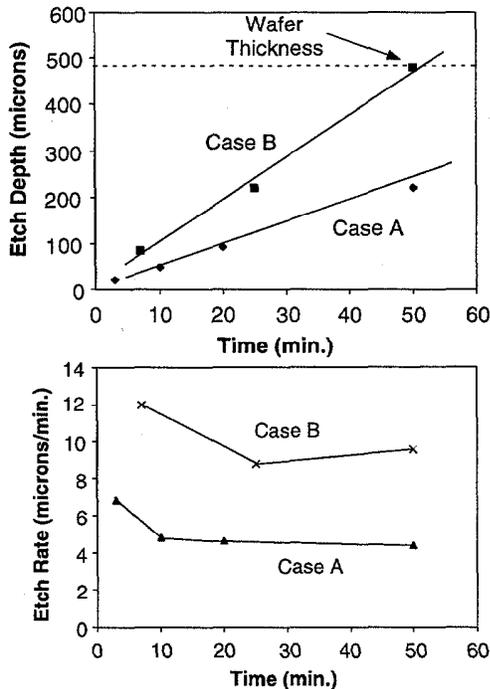
Fig 6 (a-upper): Etch progression for two cases described in the text. Through-wafer etches can be achieved in <1 hour. (b-lower): A cumulative average etch rate of 4-10 μm/min. has been sustained, although initial values are higher.

Figure 7 shows the dependence of the etch rate on electrode power density. All the etches shown in this figure were performed for three minutes at 2.7 Torr using aluminum electrodes with circular openings of 350 μm diameter. It is clear that the etch rate increases linearly with power density within the operating regime explored. The highest etch rate achieved is 17.4 μm/min. The relatively high etch rates are facilitated in part by the high power densities. Since the electrode area for microplasmas is generally <<1 cm², these power densities can be achieved at moderate current levels.

Since the etch rate is linearly related to electrode power density, it is worthwhile to evaluate the etch rate per unit power density as a figure of merit for the efficiency of the etch. Figure 8 shows this parameter as a function of chamber pressure for three minute long etches performed though a 280 μm wide slit opening in titanium electrodes. As the pressure is increased, the power necessary to sustain the plasma increases from 3-7 W/cm². Despite this, the ratio of the etch rate to power density increases from 0.15 to 4.23 μm/min per W/cm² as the pressure is changed from 2 to 20 Torr. Higher ratios achieved by other operating parameters are also shown in Fig. 8.
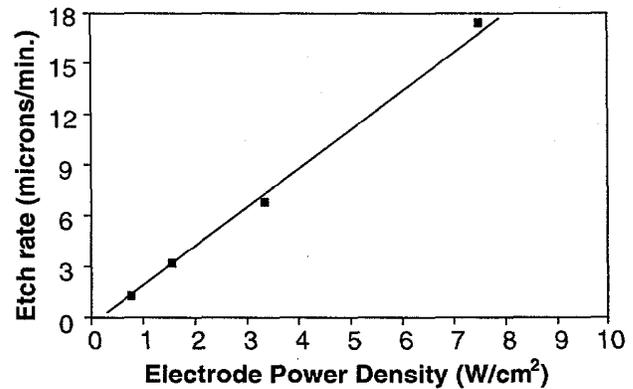
Fig 7: The etch rate is proportional to electrode power density. (All 3 min. etches at 2.7 Torr.)
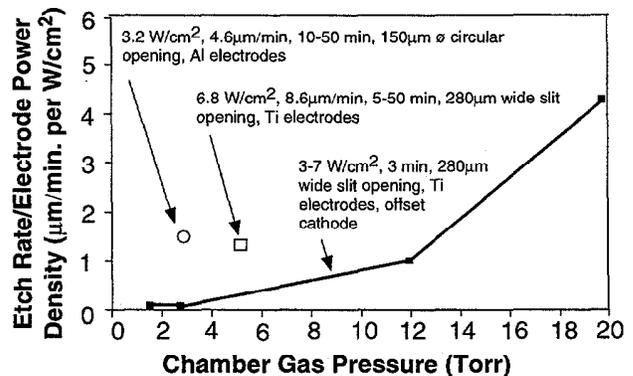
Fig 8: The ratio of the etch rate to the electrode power density increases with chamber pressure.

IV. PLASMA MODELING

The resistivity of a conventional plasma can be estimated by [4]:

$$\eta \approx \frac{\pi q^2 \sqrt{m}}{\left(4\pi\varepsilon_o\right)^2 (kT)^{3/2}} \ln \Lambda; \quad \Lambda = \overline{L_D / r_o}; \quad r_o = \frac{q^2}{4\pi\varepsilon_o m v^2}$$

where ε_o is the permittivity of free space, k it the Boltzmann constant, T is the electron temperature, q is the electron charge, m is electron mass, and v is electron velocity. For laboratory plasmas, $\ln\Lambda$ is close to 10. To the extent that the electrode power density is linearly related to kT, if the plasma dimensions remain relatively constant, eqn. (1) suggests that the plasma resistance should be proportional to $P^{-3/2}$, where P denotes the power density. Figure 9 plots the measured resistance of the microplasmas as a function of P. It also shows the best fit of these points to a curve proportional to $P^{-3/2}$, which is achieved using a proportionality constant of 2.2. The measurements for Fig. 9 were obtained at 2.7 Torr using aluminum electrodes of 0.2 cm^2 area.
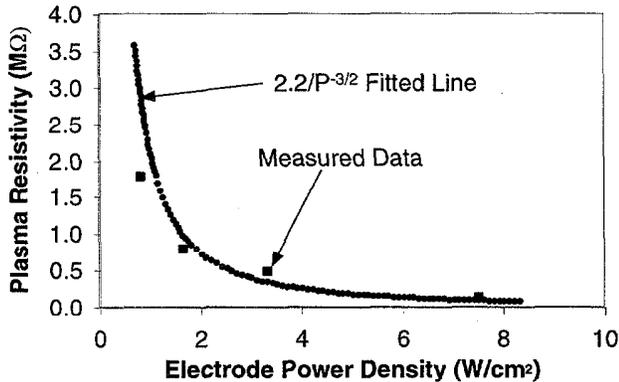
Fig. 9: Consistent with eqn. (1), the plasma resistance is proportional to $P^{-3/2}$, where P is the power density.
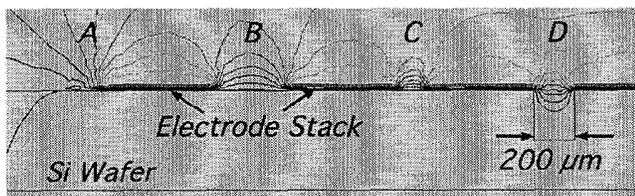
Fig. 10: Equipotential contours showing variations of the local electric field with mask openings of varying widths (A-C) and with etch progression (D).

Under the etching conditions described in the preceding section, the etching is believed to be performed by ions which are pulled away from the sheath that exists above the electrode stack by the electric field associated with the openings in the stack. Results from a preliminary effort at modeling the electric field are shown in Fig. 10. The simulations were performed using MAXWELL™, and neglect the conductivity and charge distribution of the plasma. The geometry assumes openings of various widths in a 30 μm thick polyimide-metal-polyimide electrode stack on a 500 μm thick silicon wafer. The relative dielectric constant of the polyimide is 3.5, and the resistivity of the metal is zero. The figure shows equipotential contours which are crowded progressively closer together in openings A, B, and C, suggesting that the electric fields are higher above narrow openings. At the same time, the local electric field extends to greater heights above the wider gaps. The difference between the contours at locations C and D, which are of the same width but represent shallow and deep etches, respectively, suggest that the electric field is reduced as an etch progresses. This may be correlated to the observation made in the preceding section that the initial etch rate is higher than the cumulative average rate.

V. CONCLUSIONS

This effort has demonstrated the use of microplasmas for the etching of silicon. The microplasmas were generated by DC power applied to in-situ electrodes patterned on the silicon wafer that also served as an etch mask and a shield against the applied potential. Short etches at rates exceeding 17 μm/min. and sustained etch rates in the range of 4-10 μm/min. were demonstrated, and through-wafer etches were achieved using SF$_6$. The typical operating pressure and power density were in the range of 1-20 Torr and 1-10 W/cm^2, respectively. Varying degrees of anisotropy were achieved by changing the etch conditions. The etch rate depended linearly on the electrode power density; the power efficiency of the etch rate increased with pressure. The change of the plasma resistance with the power density was found to comply with theoretical expectations.

ACKNOWLEDGEMENTS

REFERENCES

[1] W.H. Juan, S.W. Pang, "Released Si Microstructures Fabricated by Deep Etching and Shallow Diffusion," *IEEE J. Microelectromechanical Sys.*, 5(1), March 1996, pp. 18-23

[2] C. Marxer, C. Thio, M.-A. Gretillat, N.F. de Rooij, R. Battig, O. Anthmatten, B. Valk, P. Vogel, "Vertical Mirrors Fabricated by Deep Reactive Ion Etching for Fiber-Optic Switching Applications," *IEEE J. Microelectromechanical Sys.*, 6(3), Sep. 1996, pp. 227-285

[3] W.-H. Juan, S.W, Pang, A. Selvakumar, M.W. Putty, K. Najafi, "Using Electron Cyclotron Resonance (ECR) Source to Etch Polyimide Molds for Fabrication of Electroplated Microstructures," *Proc., Solid-State Sensors and Actuators Workshop*, Hilton Head, SC, June 1994, pp. 82-85

[4] F.F. Chen, *Introduction to Plasma Physics and Controlled Fusion*, Plenum Press, 1983, pp. 181

Ultra Deep Anisotropic Silicon Trenches Using Deep Reactive Ion Etching (DRIE)

A. A. Ayón, X. Zhang and R. Khanna
Department of Electrical Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139

Abstract. This paper reports solutions to the problem of profile control of narrow trenches in the vicinity of wider topographic features, as well as for etching high aspect ratio, anisotropic trenches with depths in the 300 to 500 μm range, and of widths between 12 to 18 μm.

Background.

Deep Reactive Ion Etching (DRIE) is being increasingly recognized as an enabling MEMS technology [1] and it is, in fact, playing a crucial role in the emerging field of Power MEMS [2]. An extensive characterization database has been reported recently [3, 4] for tailoring operating conditions in order to obtain prescribed profiles as well as targeted silicon etching rates and uniformity when using a DRIE tool that employs the Bosch process. This technique can be briefly described as consisting of sequential etching and passivating steps using an appropriate gas chemistry in each step (SF_6 for etching, C_4F_8 for passivating). However, solutions to the problem of etching high aspect ratio trenches located in the vicinity of wide trenches that locally deplete the availability of etching species are absent in the literature. Furthermore, the issue of etching highly anisotropic, narrow silicon trenches (measured widths of 12 to 18 μm) with etched depths in the range of 300 to 500 μm has not been scrutinized and no reported practical guidance can be found in this respect.

Narrow Trenches Adjacent to Wide Features

This exercise involved 4" single crystal silicon wafers, <100>, standard thickness, n-type, with resistivity 10-20 Ω-cm. The masking material was photoresist AZ4620 spun at 3000 rpm (thickness ≈ 6 μm). The mask consisted of arrays of trenches 3 to 5 μm wide adjacent to trenches 30 to 50 μm wide. The separation between the trenches was 10 to 15 μm, and the length of all trenches was fixed at 2000 μm. The exposed area was estimated at 10%. Etching was performed in a time multiplexed etcher from Surface Technology Systems that has been described elsewhere [3, 4].

While it is possible to locate reported operating conditions enabling users to obtain high silicon etching rates [3, 5], i.e., in excess of 3 μm/min and with reasonable etching selectivities, i.e., exceeding 70:1, and good anisotropy, it is frequently observed that it is problematic to etch high aspect ratio trenches in the vicinity of wide trenches because the latter locally decrease the availability of etching species as shown in figure 1.
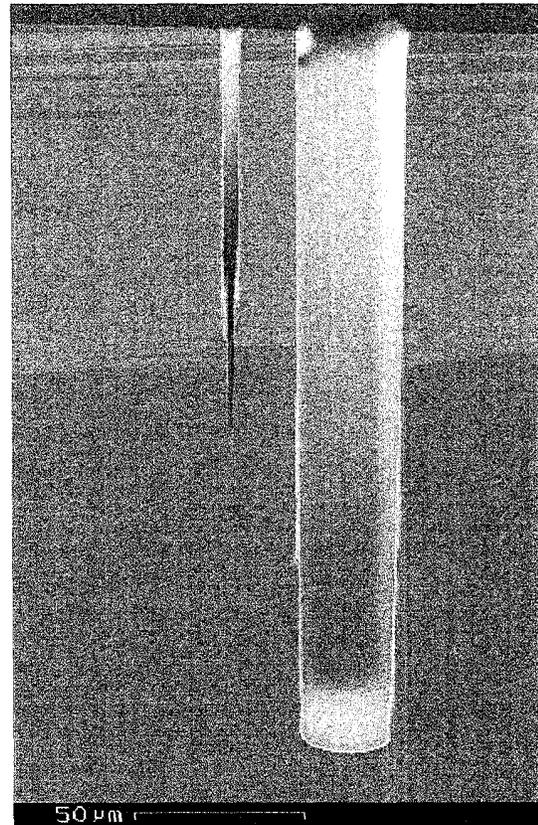
Figure 1. It is common to encounter problems while etching narrow trenches in the vicinity of wide features. Operating conditions were: 105 sccm of SF_6 (15 sec active cycle, 1 sec overlap, 12 W electrode power and 600 W of coil power), 20 sccm of C_4F_8 (11 sec of active cycle, 6 W electrode power and 600 W of coil power), the APC set at 65°. Etching conditions known as MIT47. This operating point has also been reported as producing smooth walls [5].

Additional steps have to be taken to regain control of the profile in applications exhibiting such geometries. One solution to this problem is to increase the ion flux reaching the wafer surface during the etching cycle to promote a more efficient removal of the protective polymer layer at the

bottom of the trench. The ion flux is controlled by the applied coil power during the corresponding cycle [3, 4], thus, subsequent process optimization studies concentrated on this variable. Figure 2 illustrates the dramatic improvement obtained by using the aforementioned approach.
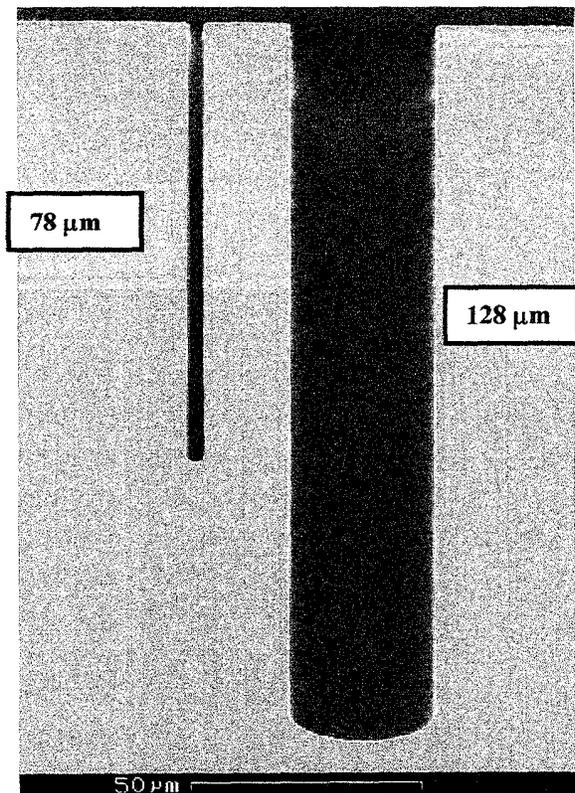
Figure 2. Results obtained after optimization of the operating conditions for etching narrow trenches (width = 3.7 μm) in the vicinity of wide trenches (width = 38 μm). Operating conditions were: 105 sccm of SF_6 (14 sec active cycle, 0.5 sec overlap, 12 W electrode power and 750 W of coil power), 40 sccm of C_4F_8 (11 sec of active cycle, no overlap, 6 W electrode power and 600 W of coil power), the APC set at 65°. Etching conditions known as MIT59. Trenches observed to widen at the rate of approximately 1 μm for every 100 μm etched down.

The set of operating conditions known as MIT59 described in figure 2, has found widespread application in accelerometers [6], micro-gas turbine engines [7], tethered motors [8], Power MEMS [9] and, in general, for etching through wafers.

Other variations of MIT59 have been used to etch through wafers with thickness in excess of 525 μm. In these instances the strategy is to lower the electrode power during the etching cycle from 12 W to 10 W (variation known as MIT69) [10], thereby increasing the selectivity at the

expense of a slight deterioration in etching rate and anisotropy. By increasing the coil power during the etching cycle to 800 W (variation known as MIT56) it has been possible to fabricate successfully 10 μm wide trenches up to 200 μm deep. These trenches are important in forming forward and aft thrust nozzles for microturbomachinery applications [7, 11].

Ultra Deep Trenches

The mask used in these experiments consists of an array of 12 circular trenches with a diameter of 4200 μm and measured width of 10 μm after photolithography. Standard vacuum theory applied to narrow (width ≤ 13 μm), very high aspect ratio (>15) silicon trenches, reveals that the conductance of these features decreases (see figure 3) as the depth of the trench increases [12, 13]. Decreasing the conductance of the trench impedes both the transport of etchant species to the surface bottom and the removal of etching byproducts [13-15]. The combined effect is reflected in the noticeable profile deterioration when depth increases beyond 200 μm, as shown in figure 4.
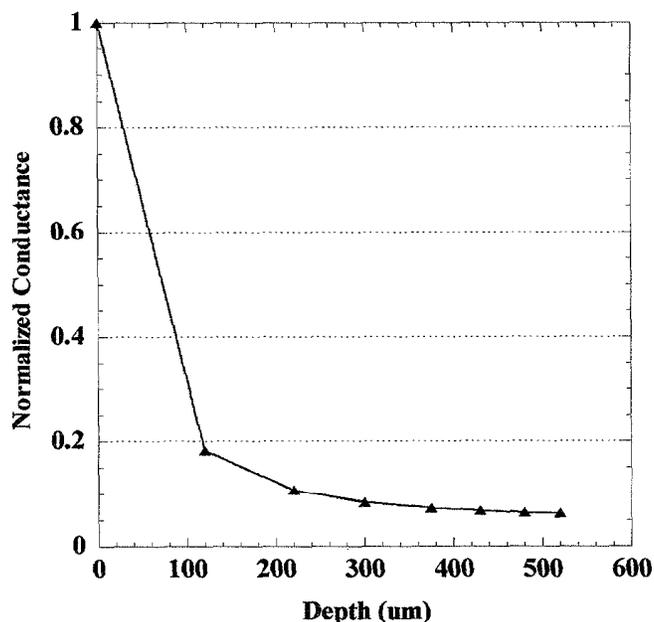
Figure 3. As the depth of etched trenches increases, the neutral flow conductance decreases drastically, effecting trench profile and silicon etching rate.

This problem can be alleviated by dynamically compensating the neutral flux reduction reaching the feature bottom during processing and by quickly removing the etching byproducts that dissociate and redeposit. This type of control is natural in tools employing the Bosch process because in the standard mode of operation, the throttle or

automatic pressure control (APC) valve remains fixed and the chamber pressure is controlled by the flow rate of the respective gas species during the etching or passivation cycles. Thus, the pressure during the etching cycle can be systematically increased throughout a particular process simply by increasing the SF_6 flow rate (see figure 5). The chamber pressure sets the neutral flux via the following,

$$N_f = p/\sqrt{(2\pi m K_B T)},$$

where N_f is the neutral flux, p the chamber pressure, m the mass of the neutral species and T its temperature. Additional advantages of operating at higher pressures include higher collisional and dissociation rates which increase the availability of etching species. The higher SF_6 flow rates associated with high pressure operation facilitate the replenishing of etching species and the quick removal of etching byproducts. Thus, increasing the SF_6 flow rate was the approach selected to locate processing conditions that permitted the repeatable microfabrication of trenches with depths in excess of 300 μm, with measured widths ≤ 13 μm and good profile control, as the trench shown in figure 6.
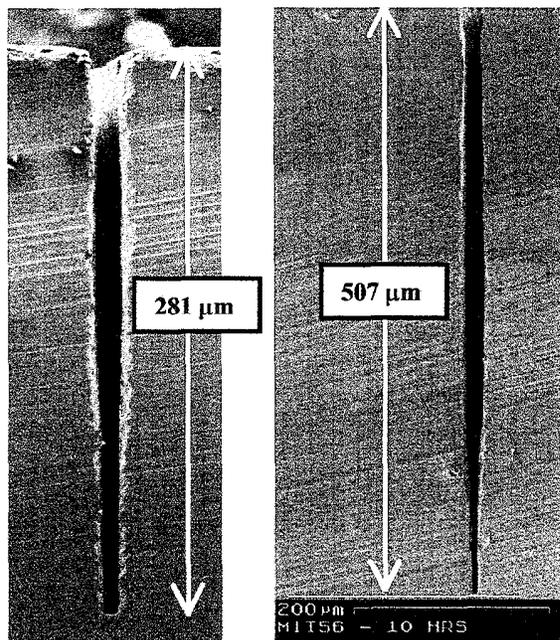
Figure 4. Trench profile quickly deteriorates as the depth of a high aspect ratio trench increases. The trench in the left SEM micrograph was obtained during a 4-hour etch using MIT56. Profile deterioration becomes more noticeable as the depth extends beyond 300 μm as can be observed in the right SEM micrograph. At its widest point, the trench on the right measured ~20 μm.

These trenches serve as journal bearings [7, 11] for microturbomachinery elements rotating at speeds in excess of 10^6 rpm, for which, the profile is critical for achieving the

desired performance. The manufacturer of the etching tool used in this work now offers software controlled upgrades that allow the automatic modification of operating conditions during etching using parameter ramping schemes [16]. However, the equipment utilized in this work lacked such capabilities and the processing variations were implemented by creating a process with a number of 60-minute steps where the SF_6 flow rate was monotonically increased by 2 sccm, every 1-hour step.
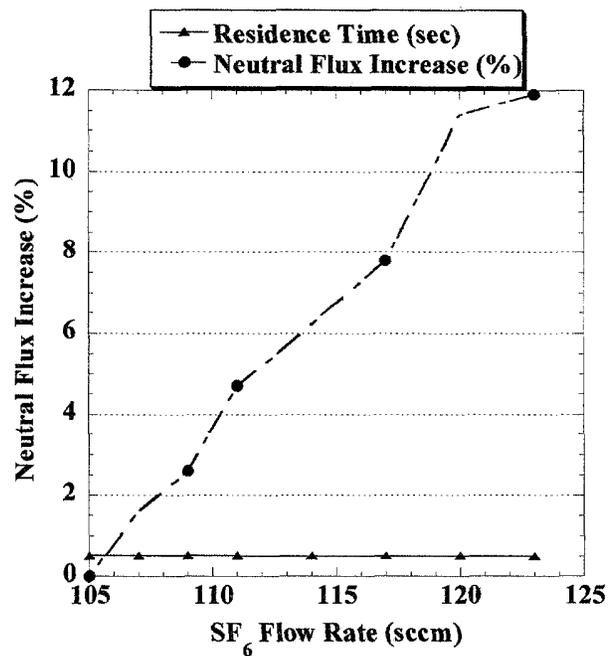
Figure 5. By increasing the SF_6 flow rate during processing, the chamber pressure is increased, and, consequently, the neutral flux to the wafer surface is increased. Residence time is given by pV/f, where V is the chamber volume (36 l) and f the SF_6 flow rate.

The carefully selected operating conditions described herein, have also been tested on trenches reaching depths in excess of 500 μm. One example of such an effort is shown on the right SEM micrograph in figure 6. As the depth of etched trenches extends beyond 300 μm, it is increasingly more difficult to control the anisotropy of narrow, high aspect ratio silicon trenches, note, however, that it is feasible to produce profiles with a slight positive taper that are acceptable for many applications (see figure 6, SEM micrograph on the right). It is thought that the profile variation for trenches with depths in excess of 500 μm, can be understood in terms of the widening of the ion angular distribution function at the entrance of the trench due to the increased chamber pressure, and also to the loss of ion flux reaching the feature bottom as the depth of the trench becomes a larger fraction of the mean free path typical at these operating pressures.

314 μm

515 μm

Figure 6. Operating conditions are based on MIT56, with +2 sccm increase in SF₆ flow rate every hour. The average width of the trench is about 11.8 μm. Normally, a 4-hour etch delivers a 315-320 μm deep trench.

Conclusions.

We have demonstrated the ability to control the profile of narrow trenches (≤ 5 μm) in the vicinity of much wider features by tailoring the ion flux reaching the wafer surface. We have also demonstrated the etching of ultra deep anisotropic silicon trenches in the range of 300 to 500 μm by dynamically controlling the neutral flux arriving at the wafer surface during processing.

Together, the manufacturing processes described herein are critical for the construction of micro-turbomachinery and bi-propellant micro-rocket engines.

References.

[1] E. Klaassen *et al*, in *Proc. of Transducers '95*, Stockholm, Sweden, June 1995, p. 556.

[2] A. H. Epstein and S. D. Senturia, *Science*, **276**, 1211 (1997).

[3] A. A. Ayón, R. Braff, C. C. Lin, H. Sawin and M. A. Schmidt, *J. Electrochem. Soc.*, **146**, 339 (1999).

[4] A. A. Ayón, R. Braff, R. Bayt, H. H. Sawin and M. A. Schmidt, *J. Electrochem. Soc.*, **146**, 2730 (1999).

[5] R. Bayt, A. A. Ayón and K. Breuer, 33rd AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, Seattle, WA, July, 1997.

[6] K. Ishihara, C.-F. Yung, A. A. Ayón and M. A. Schmidt, *J. Of MicroElectroMechanical Systems*, **8**, 403, (1999).

[7] C.-C. Lin, A. A. Ayón, R. Ghodssi, D.-Z. Chen, S. Jacobson, K. Breuer, A. Epstein and M. A. Schmidt, Proceedings of MEMS'99, Orlando, Florida, January 17-21, 1999.

[8] S. F. Nagle and J. H. Lang, 27th Meeting of the Electrostatics Society of America, Boston (1999).

[9] A. Epstein, S. D. Senturia, G. Anathasuresh, A. Ayon, K. Breuer, K-S Chen, F. E. Ehrich, G. Gauba, R. Ghodssi, C. Groshenry, S. Jacobson, J. H. Lang, C-C Lin, A. Mehra, J. M. Miranda, S. Nagle, D. J. Orr, E. Piekos, M. A. Schmidt, G. Shirley, M. S, Spearing, C. S. Tan and I. A. Waitz, in *Proceedings of Transducers '97*, Chicago (1997).

[10] A. Mehra, X. Zhang, A. A. Ayon, I. A. Waitz, M. A. Schmidt and A. H. Epstein, *submitted to J. Of MicroElectroMechanical Systems*, (2000).

[11] L. Frechette *et al*, in *Proc. of Solid State Sensor and Actuator Workshop*, Hilton Head, SC (2000).

[12] J. B. Hudson, in *Surface Science*, John Wiley and Sons, p. 143 (1998).

[13] J. W. Coburn and H. F. Winters, *Appl. Phys. Lett.*, **55**, 2730 (1989).

[14] C. Lee, D. B. graves and M. A. Lieberman, *Plasma Chemistry and Plasma Processing*, **16**, No. 1 (1996).

[15] D. C. Gray *et al*, *J. Vac. Sci. Technol.*, **B11**, 1243 (1993)

[16] J. Hopkins *et al*, in *Symposium Proc. of the MRS*, vol. 546, pp. 63-68 (1999).

Through-Wafer Electrical Interconnects Compatible
With Standard Semiconductor Processing

Eugene M. Chow, Aaron Partridge, Calvin F. Quate, and Thomas W. Kenny
Ginzton Laboratory, Stanford University
Stanford, CA, 94305-4085

ABSTRACT

The fabrication and characterization of high-density electrical through-wafer interconnects on silicon substrates are described. A 20 μm via is etched with a high-density plasma, followed by conformal oxide deposition, polysilicon deposition, and diffusion doping. The process uses high-temperature compatible materials and results in a filled via, making it suitable for standard semiconductor processing and wafer-handling equipment. Resistance (14 Ω), capacitance (0.6 pF) and noise measurements are presented.

INTRODUCTION

Electrical through-wafer interconnects (ETWI) in silicon have significant applications for integrated circuits (IC) and microelectrical mechanical systems (MEMS). New electrical structures are enabled, such as high-Q three-dimensional inductors or vertical diodes for charged particle detection [1-3]. Stacks of IC or sensor chips are possible, saving space, and adding packaging and system design flexibility [4]. If the ETWI are small and densely packed, they could enable the transfer of thousands of signals between stacked chips. Small ETWI benefit arrayed detectors, such as ultrasound imagers, and minimize non-sensing area and thus enhance performance [5]. In previous work on two-dimensional arrays of scanning probes, we demonstrated how small ETWI (30 μm) solve the geometric problem of aligning arrays of micron-sized tips to arbitrary planar samples in the presence of wire-bonds [6, 7]. ETWI can provide a solution to MEMS-IC integration problems, as they permit MEMS and ICs to be fabricated on separate wafers, and afterwards be bonded together to form a system [5].

Unfortunately, previous interconnects use techniques which are not generally compatible with standard semiconductor processing. Laser drilling can make 20 μm diameter holes, but the process is serial and not as geometrically flexible as standardized lithography and etching techniques [2]. A wet etch is most common, leaving large holes in the wafer (>200 μm), precluding future vacuum handling and seriously impeding the use of many etching, metalization, and lithography tools [8, 9]. Subsequent layers often require unconventional lithography, such as electrodeposited resist or shadow masks [1, 10]. High-density anisotropic plasma etching is attractive for via formation, as it enables smaller holes than wet etching and more flexible geometry. While we previously plasma etched a small (30 μm) via and used spin-on thick resist patterning, we could not use vacuum handling because the interconnects were not filled [7].

Our primary ETWI application is for integration with micromachined scanning probe arrays, for applications in imaging, data storage, and lithography [11-13]. These sensors are delicate, freestanding structures, so ETWI prior to device fabrication is preferred. Sensor packing density, and thus small ETWI, are critical. High-temperature thin film steps are necessary for subsequent device fabrication, so ETWI fabrication precludes the use of metals. This is acceptable because resistances of only tens of ohms are required, as they are in series with kilo-ohm piezoresistors. Signals are generally in the kilohertz range, so picofarad capacitances are more than adequate. To not degrade sensor sensitivity, noise smaller than the piezoresistor cantilever noise is desired.

In this work we have fabricated and characterized ETWI which meet these criteria. They can be integrated with subsequent standard semiconductor processing, as the interconnects are filled and use only standard high-temperature compatible materials.

FABRICATION

The basic fabrication process (Fig. 1), consists of etching a high-aspect ratio through-wafer via and subsequently coating with an insulator and then a conducting signal layer.
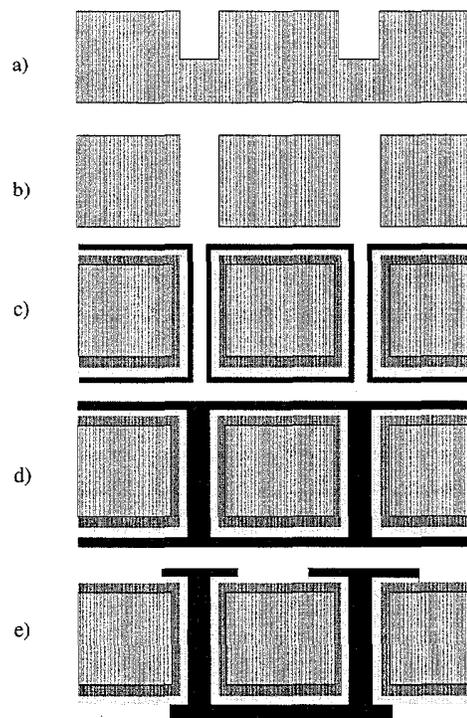
Figure 1. Fabrication process for a pre-process through-wafer interconnect. a) Holes are etched halfway through the wafer. b) Holes are etched from the backside of the wafer, meeting the frontside holes. c) After diffusion doping, a ground plane oxide and then polysilicon are deposited. d) The polysilicon is diffusion doped, and additional polysilicon is deposited to fill. More diffusion and an anneal follows. e) The conduction layer is patterned on both sides.

Etching the vias is the first, and most critical step. Ten micron thick photoresist (Shipley SPR220-7) is initially coated on one side of the wafer. After a 120 second hotplate bake at 90 °C, the other side is similarly coated, and the wafer is oven baked at 110 °C for 60 minutes. Previously patterned alignment marks on the front side silicon, combined with a backside contact mask aligner (KarlSuss), enable the exposure of aligned 20 μm circles on both sides of the wafer. The wafers then sit in air for 8 hours to absorb water. This moisture helps to ensure shorter develop times (90 seconds in Shipley LDD26W) and thus straighter photoresist sidewalls.

A high-density plasma etch (Surface Technology Systems) is performed from the front side of the wafer for 210 minutes. The etcher uses the Bosch etching process, switching between sequential passivation (85 sccm C_4F_8, 6 sec., 0 W platen, 600 W coil, 16 mtorr), and etching (80 sccm SF_6, 8 sec., 8 W platen, 600 W coil, 33 mtorr). Resist is then coated and baked on the front side for use as a seal from the wafer-chuck helium gas cooling. This is necessary because another 180 minute etch, this time from the backside, is then performed to punch through the via. The etch is very anisotropic, and alignment is as accurate as the lithography (+/- 1 μm). The diameter varies from 24 μm (top/bottom) to 18 μm (middle). By etching from both sides of the wafer, the required aspect ratio is reduced by two from 40:1 to 20:1, and it is possible to etch a 20 μm circle through a 400 μm thick wafer. The overetch time of the second etch, which punches through, should be minimized to reduce lateral etching of the front side silicon. Using an etch with increased passivation for the final 30 minutes helps to limit this effect. Figure 2 depicts an etched via.

Etches with the required aspect ratios (>20:1) have been previously reported, but they are of trenches, with one dimension very long (~1000 μm), or not as deep as through a 400 μm thick wafer [3, 14, 15]. Decreased etching rate for small areas, characteristic for flourine-based plasmas, is particularly severe for deep etches [16]. We measured the time to etch trenches (20 μm x 1000 μm) 200 μm deep to be three times less than the time to etch a circle (20 μm diameter) to the same depth. By etching from both sides of the wafer, the average etch rate is higher, which shortens the total etch time and permits the use of thinner photoresists.
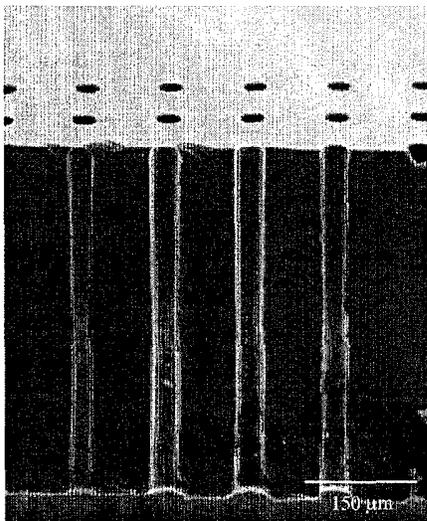
Figure 2. SEM cross-section after plasma etching vias through the wafer.

After forming the via, conformal thin films are deposited. To shield the signal, a ground plane is formed by diffusion doping the wafer and vias with phosphorous for 30 minutes at 1000 °C. This is important if the ETWI is used to conduct small signals from a sensor to an amplifier, as it shields electrical noise from the environment. The diffusion forms a glass which is removed by a brief etch in buffered hydrofluoric acid (BHF). There is a possibility the BHF does not fully wet all of the vias, even with surfactant and agitation. This could possibly leave unwanted interface charges which would compromise the capacitance stability under different biases. The ground shield masks such effects if present.

For electrical isolation, thermal silicon dioxide (2 μm) is grown at 1100 °C. For signal conduction, low pressure chemical vapor deposition (LPCVD) polysilicon (3 μm) is deposited at 600 °C, followed by phosphorous diffusion at 1000 °C for 30 minutes, and its corresponding BHF dip. To fill the via, more LPCVD polysilicon is deposited (7 μm). Another 30 minute phosphorous diffusion is performed to make the surface conducting, followed by another BHF dip, and then a 120 minute anneal at 1000 °C to fully drive dopants throughout the polysilicon. Conventional spin-on resist is then used to pattern the polysilicon with SF_6 plasma, repeating on each side to complete the ETWI (Fig. 3).
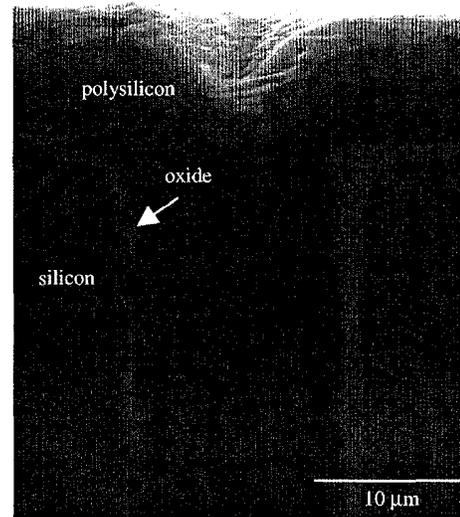
Figure 3. SEM cross-section of a completed through-wafer interconnect. Planarity over the via is adequate for conventional spin-on resist lithography. Polishing can be applied if improved planarization is required.

CHARACTERIZATION

Current versus voltage measurements were made of interconnect chains to determine the resistance of the structures and demonstrate ohmic behavior (Fig. 4). The slope of each curve is inversely proportional to the resistance sum of the ETWI and the planar lines between them (50 μm wide, 110 μm for this device). A slight hyperbolic sine form is observed for smaller resistance ETWI, characteristic of polysilicon resistors [17]. Poor probing contact or Joule heating could also cause such nonlinearites [18]. A small Schottky diode formed by the electrical probe station tip and the polysilicon resistor contact pad would have a pronounced effect on small resistance structures, where the diode effective

resistance is comparable to that of the ETWI. Joule heating effects would also be pronounced for lower resistance. Measurements were repeated multiple times and slopes (resistances) were found to vary +/- 5%.
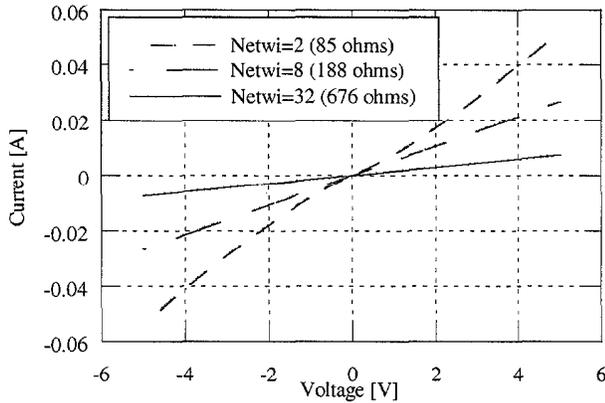
Figure 4. Current versus voltage for consecutive chains of 2, 8, and 32 ETWI.

To extract the resistance of an individual ETWI, planar test structures (no ETWI), were similarly measured. These resistances were subtracted from the chain resistances and plotted by number of ETWI in the chain (Fig. 5). The slope of the resulting line gives an average single ETWI resistance of 14 Ω. Bulk model calculations give similar results, as a cylindrical wire has a resistance of $R = \rho L / A$, where ρ is the polysilicon resistivity, L the length of the ETWI, $A = \pi r_c^2$ the cross sectional area, and r_c the radius of the conducting core. Four point probe measurements were performed to obtain a doped polysilicon sheet resistance $R_s = \rho/t$ of 1.1 Ω/sq., where t is the film thickness. Assuming the via was not significantly overfilled with polysilicon, then $t = r_c$. For L = 415 µm and an approximate r_c = 10 µm, this gives R = 14.5 Ω.

Capacitance was similarly measured on chains of interconnects. Test structures were measured so that the capacitance due to the planar lines between the ETWI could be subtracted. Capacitance was measured between the ground plane and the signal plane at 10 kHz, with the polysilicon-oxide-silicon structure biased in accumulation. Capacitance measurements for a single ETWI chain were repeatable to within 3%. The slope of a fitted line (Fig. 5) results in a measured 0.6 pF per ETWI. We model the capacitance as a coaxial transmission, where C = $2\pi \varepsilon_{ox}$
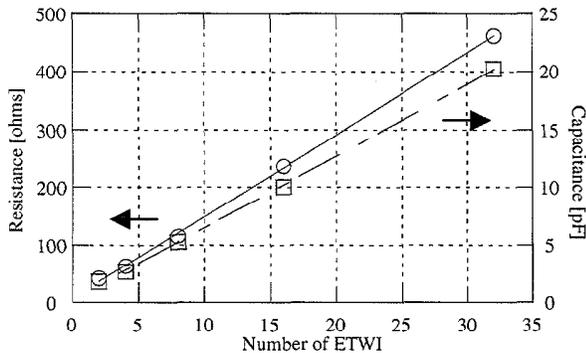
Figure 5. Resistance and capacitance of ETWI chains.

$\varepsilon_o L / ln(r_c/ r_v)$, and r_v is the outer conductor diameter (the via diameter after oxidation), ε_{ox} the relative dielectric permitivity of silicon dioxide, and ε_o the dielectric permitivity of free space. Taking r_v= 12 µm and ε_{ox} = 3.9 gives a capacitance of 0.5 pF per ETWI.

Noise characteristics were measured by placing the ETWI chains in a matched Wheatstone bridge. A chain of 128 ETWI was compared to a metal film resistor, both with a resistance of 2100 Ω. The ETWI shows increased 1/f noise compared to the reference resistors. The intersection of the 1/f noise with the Johnson noise limit is extrapolated to be ~10 kHz, similar to previous work [17, 18]. Both spectra are limited by the noise floor (8 nV/ sqrt Hz) of the amplifier (AD624 instrumentation amplifier at 100 × gain), which exceeds the Johnson noise of a 2100 Ω resistor (3 nV/ sqrt Hz). A typical spectrum is given in Figure 6.
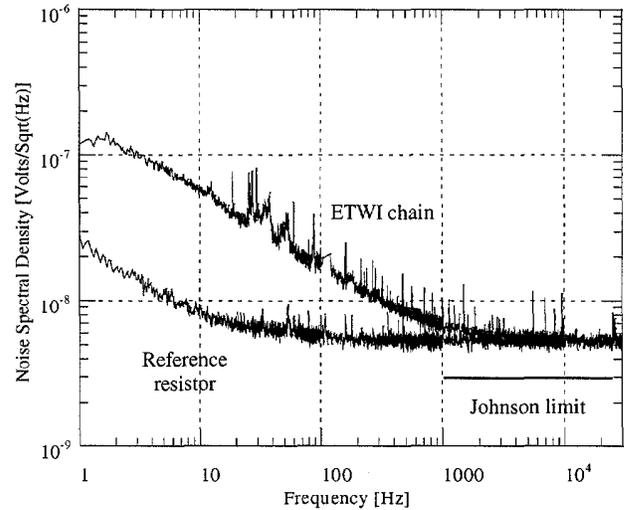
Figure 6. Noise spectrum for ETWI chains compared to a reference resistor.

DISCUSSION

The resistance and capacitance of the ETWI are satisfactory for use our standard piezoresistor sensors. The slight nonlinearity in the current-voltage curve changes the effective resistance for large voltage swings. Since piezoresistive signals are generally small (< 1 mV) and centered around a constant bias, this is acceptable for our applications. A capacitance of ~ 1 pF, with a 14 Ω resistance, gives a time constant of 10^{-11} seconds, significantly more than adequate for most piezoresistive applications. The measured low frequency noise of a ETWI chain appears to be comparable to a typical piezoresistive cantilever, which would degrade cantilever sensitivity [19]. This needs to be investigated more by measuring individual ETWI or ETWI in series with piezoresistor sensors. Aluminum contact pads could help with possible contact problems.

Thermal cycling, stress, and reliability issues have not yet been studied for this scale geometry. However, smaller polysilicon contact plugs (few microns deep) with similar or greater aspect ratios have been successfully integrated with standard IC processes [20, 21]. The thermal strain mismatch between silicon and polysilicon is small, so stresses are expected to be from oxide mismatches, and polysilicon internal stress [22]. An occasional ETWI exhibited a filling void, but because they were sealed under

vacuum we do not expect subsequent stressing in low-pressure systems. Possible absolute resistance variations because of voids are acceptable for this application.

Variations on the ground plane technique are being considered. The current technique leaves the device surface heavily doped and thus unfit for piezoresistor formation. One solution is to not use a ground plane, which might be acceptable for certain applications. While this decreases the total ETWI capacitance, because of the bulk silicon depletion, the oxide thickness still sets the maximum capacitance for the ETWI because the two capacitors are in series. Another approach is to keep the ground plane, but first deposit an isolating oxide and then doped polysilicon ground plane. After this, the insulating oxide is deposited, followed by the signal layer. This has the added advantage of reducing the overall ETWI capacitance (proportional to the signal conduction layer radius), without requiring higher aspect-ratio etching.

CONCLUSIONS

Through-wafer electrical interconnects which can be integrated with standard semiconductor processing have been demonstrated. The small size (20 μm diameter) makes the ETWI attractive for high-density applications. Unlike previous work, the ETWI are filled and uses high-temperature compatible materials, making them compatible with standard vacuum handling equipment and subsequent integrated circuit processing. The resistance (14 Ω) is adequate for piezoresistive applications, where the series sensor resistor is typically > 1kΩ. The low capacitance (< 1pF) is attractive for high-speed applications. The challenge of combining MEMS with IC electronics is one of many applications this technology helps to address.

ACKNOWLEDGMENTS

REFERENCES

1. H. T. Soh, C. P. Yue, A. McCarthy, et al., "Ultra-low resistance, through-wafer via (TWV) technology and its applications in three dimensional structures on silicon", Japanese Journal of Applied Physics, Part 1 (Regular Papers, Short Notes & Review Papers), 38(4B) 2393-6 (1999).

2. T. R. Anthony, "Forming electrical interconnections through semiconductor wafers", Journal of Applied Physics, 52(8) 5340-9 (1981).

3. C. Kenney, S. Parker, J. Segal, et al., "Silicon detectors with 3-D electrode arrays: fabrication and initial test results", IEEE Transactions on Nuclear Science, 46(4, pt.3) 1224-36 (1999).

4. M. Heschel, J. F. Kuhmann, S. Bouwstra, et al., "Stacking technology for a space constrained microsystem", Journal of Intelligent Material Systems and Structures, 9(9) 749-54 (1998).

5. X. C. Jin, S. Calmes, C. H. Cheng, et al. "Micromachined capacitive ultrasonic immersion transducer array", 10th International Conference on Solid-State Sensors and Actuators (Transducers '99), Sendai, Japan, 6/7-10/99, The Institute of Electrical Engineers of Japan (1999), pp. 1500-1503.

6. E. M. Chow, H. T. Soh, H. C. Lee, et al., "Integration of through-wafer interconnects with a two-dimensional cantilever array", Sensors and Actuators, Physical A, (Accepted for publication in Vol. 82-83).

7. E. M. Chow, H. T. Soh, H. C. Lee, et al. "Two-dimensional cantilever arrays with through-wafer interconnects", The 10th International Conference on Solid-State Sensors and Actuators, Sendai, Japan, 6/7-10/99, Institute of Electrical Engineers of Japan (1999), pp. 1886-1887.

8. C. Christensen, P. Kersten, S. Henke, et al., "Wafer through-hole interconnections with high vertical wiring densities", IEEE Transactions on Components, Packaging, and Manufacturing Technology, Part A, 19(4) 516-22 (1996).

9. S. Linder, H. Baltes, F. Gnaedinger, et al. "Fabrication technology for wafer through-hole interconnections and three-dimensional stacks of chips and wafers", Proceedings IEEE Micro Electro Mechanical Systems. An Investigation of Micro Structures, Sensors, Actuators, Machines and Robotic Systems (Cat. No.94CH3404-1), New York, NY, USA, IEEE (1994), pp. 349-354.

10. G. J. Burger, E. J. T. Smulders, J. W. Berenschot, et al., "High-resolution shadow-mask patterning in deep holes and its application to an electrical wafer feed-through", Sensors and Actuators A-Physical, 54(1-3) 669-673 (1996).

11. S. C. Minne, S. R. Manalis, and C. F. Quate, "Parallel atomic force microscopy using cantilevers with integrated piezoresistive sensors and integrated piezoelectric actuators", Applied Physics Letters, 67(26) 3918-20 (1995).

12. M. Lutwyche, C. Andreoli, G. Binnig, et al., "5x5 2D AFM cantilever arrays a first step towards a Terabit storage device", Sens. Actuators A, Phys. (Switzerland), Sensors and Actuators A (Physical), A73(1-2) 89-94 (1999).

13. K. Wilder and C. F. Quate, "Scanning probe lithography using a cantilever with integrated transistor for on-chip control of the exposing current", J. Vac. Sci. Technol. B, Microelectron. Nanometer Struct. (USA), Journal of Vacuum Science & Technology B (Microelectronics and Nanometer Structures), 17(6) 3256-61 (1999).

14. A. A. Ayon, R. Braff, C. C. Lin, et al., "Characterization of a time multiplexed inductively coupled plasma etcher", Journal of the Electrochemical Society, 146(1) 339-49 (1999).

15. J. Bhardwaj, H. Ashraf, A. McQuarrie, et al. "Dry silicon etching for MEMS", Proceedings of the Third International Symposium on Microstructures and Microfabricated Systems, Pennington, NJ, USA, Electrochem. Soc (1997), pp. 118-130.

16. R. A. Gottscho, C. W. Jurgensen, and D. J. Vitkavage, "Microscopic uniformity in plasma etching", Journal of Vacuum Science & Technology B (Microelectronics and Nanometer Structures), 10(5) 2133-2147 (1992).

17. M. Y. Luo and G. Bosman, "An analytical model for 1/f noise in polycrystalline silicon thin films", IEEE Transactions on Electron Devices, 37(3, pt.1) 768-74 (1990).

18. M. J. Deen, S. Rumyantsev, and J. Orchard-Webb, "Low frequency noise in heavily doped polysilicon thin film resistors", Journal of Vacuum Science & Technology B (Microelectronics and Nanometer Structures), 16(4) 1881-4 (1998).

19. B. W. Chui, T. D. Stowe, J. Yongho Sungtaek, et al., "Low-stiffness silicon cantilevers with integrated heaters and piezoresistive sensors for high-density AFM thermomechanical data storage", Journal of Microelectromechanical Systems, 7(1) 69-78 (1998).

20. J. M. Drynan, E. Ikawa, and T. Kikkawa. "Fabrication of polysilicon plugs for deep-submicron contact-holes", 1990 Proceedings. Seventh International IEEE VLSI Multilevel Interconnection Conference (Cat. No.90TH0325-1), New York, NY, USA, IEEE (1990), pp. 494.

21. T. Hamajima and Y. Sugano. "Low contact resistance polysilicon plug for halfmicron CMOS technology", 1989 Proceedings. Sixth International IEEE VLSI Multilevel Interconnection Conference (Cat. No.89TH0259-2), New York, NY, USA, IEEE (1989), pp. 508.

22. D. Chidambarrao, J. P. Peng, and G. R. Srinivasa, "Stresses in silicon substrates near isolation trenches", Journal of Applied Physics, 70(9) 4816-22 (1991).

HERMETIC WAFER BONDING BASED ON RAPID THERMAL PROCESSING

Mu Chiao and Liwei Lin
Department of Mechanical Engineering
University of California at Berkeley
Berkeley, CA 94720-1740, USA
E-mail: muchiao@me.berkeley.edu

ABSTRACT

Hermetic wafer bonding based on rapid thermal processing (RTP) has been demonstrated for the first time. Micro cavities are sealed by aluminum-glass bonding system under the bonding condition of 990°C for 2 seconds by using RTP. Reliability tests in both IPA and autoclave chamber show that 100 % of survival rate can be achieved when the aluminum bonding solder is 150 μm wide and 4.5 μm thick. It is found that the activation energy for RTP aluminum-to-glass bonding is 2.8 eV and the lowest successful bonding temperature is 760°C with a processing time of 30 minutes. A MEMS micro heater has been successfully packaged by the RTP bonding method and the heater is operational after the bonding process. This package demonstration shows that RTP bonding can overcome surface roughness of 2 μm as created by the interconnection line. Because RTP is a wafer-level process and is capable of providing low thermal budget, RTP bonding has promising potential for wafer-level MEMS fabrication and packaging.

INTRODUCTION

Wafer bonding is an important technology in both IC and MEMS processing. The existing bonding methods such as fusion and anodic bonding suffer from high temperature treatment, long processing time, requirement of flat surface and possible damages to the circuitry. For example, silicon-silicon direct wafer bonding requires bonding temperature over 1000°C on two flat substrates [1]. Solder bonding technology provides low processing temperature and reflow of solder such that it has potential application for MEMS packaging. However, flux used in conventional solder to improve wetting capability [2] may cause serious contamination problems for hermetic package [3]. Previously, pure aluminum has been used as solder in the aluminum-to-glass bonding system by localized heating and bonding [4]. This bonding method has the advantage of providing localized high bonding temperature for strong bonding. At the same time, it avoids possible damages due to global high temperature processing to circuitry or MEMS devices. However, it is difficult to implement localized aluminum-glass solder bonding massively and in parallel [5].

Rapid Thermal Processing is introduced as a bonding method for the first time in this paper. RTP can be easily conducted at the wafer-level and is capable of precise control of thermal budget [6]. Furthermore, RTP has been commonly used in IC post-processing because dopant redistribution in the microelectronics is minimized by using RTP as compared with regular annealing processes. Therefore, RTP bonding such as the one presented in this paper, may have the potential to solve specific bonding problems in IC or MEMS processing and provide a new class of wafer bonding technique.

RTP BONDING CHARACTERISTICS

In order to demonstrate hermetic wafer bonding by using rapid thermal processing, an experimental set up is illustrated in Fig. 1. The aluminum-to-glass bonding is used as the demonstration example for the RTP bonding experiment. A silicon substrate is oxidized to grow 0.8 μm of thermal oxide. Aluminum is evaporated and patterned to form an encapsulation area around 400 x 400 μm^2. The width of aluminum solder is designed to be 50, 100 and 150 μm with thickness varying from 2.5 to 4.5 μm to investigate the optimal RTP bonding parameters. A Pyrex glass substrate is placed on top and the whole set up is put into a RTP chamber [7]. The heat source is provided by 13 tungsten-halogen lamps, each produces 1.5 KW power. The testing specimen is heated up by thermal radiation with the temperature ramp up rate as high as 100°C per second (400-1000°C) and the cool down rate at around 50°C per second. A typical temperature profile for the RTP bonding pro-
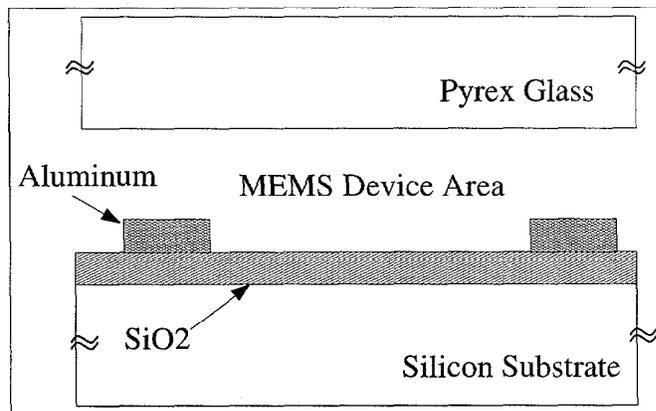
Figure 1: The set up of aluminum-glass RTP bonding experiments.

cess is shown in Fig. 2. It only takes about 60 seconds to ramp up and cool down during the RTP bonding experiment such that possible dopant redistribution can be minimized [8, 9].

In order to characterize the RTP bonding process, a series of experiments with different combinations of bonding temperature and time have been conducted. Figure 3 shows two bonding curves by RTP and localized heating and bonding [10], respectively. The symbols represent the successful bonding con-
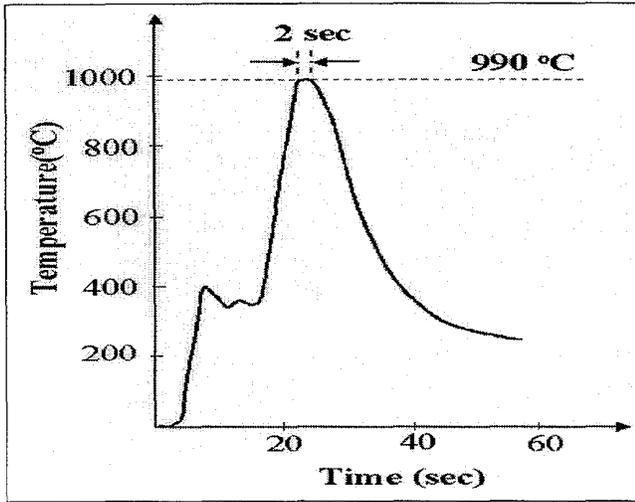
Figure 2: The temperature versus time plot in a RTP bonding experiment.
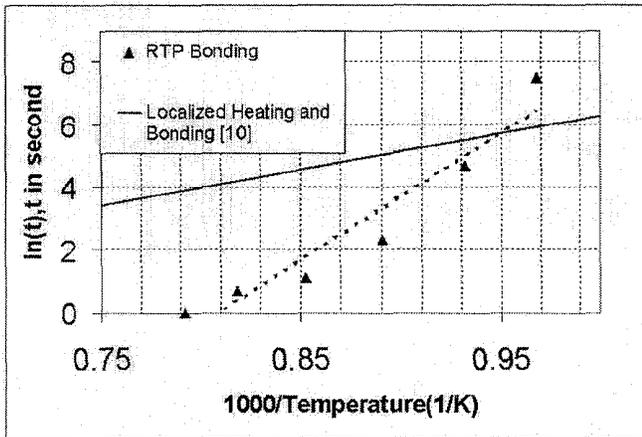
Figure 3: Aluminum-glass bonding experiments by RTP and localized heating.

Figure 4: Leak test in isopropyl alcohol(IPA). Only the right bottom cavity shows leakage.
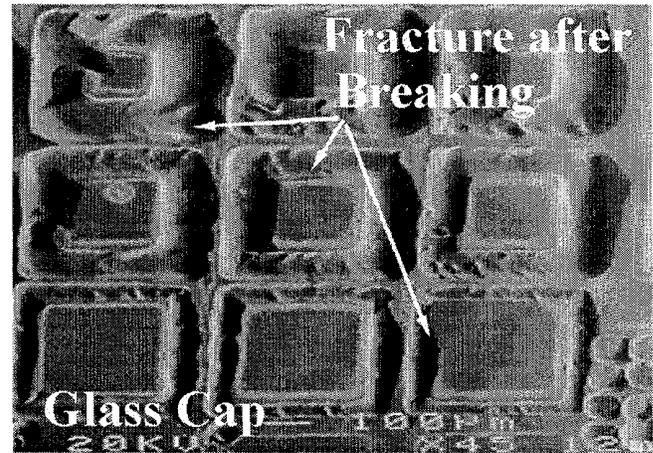
Figure 5: The SEM micrograph of glass wafer after forcefully breaking the bond.
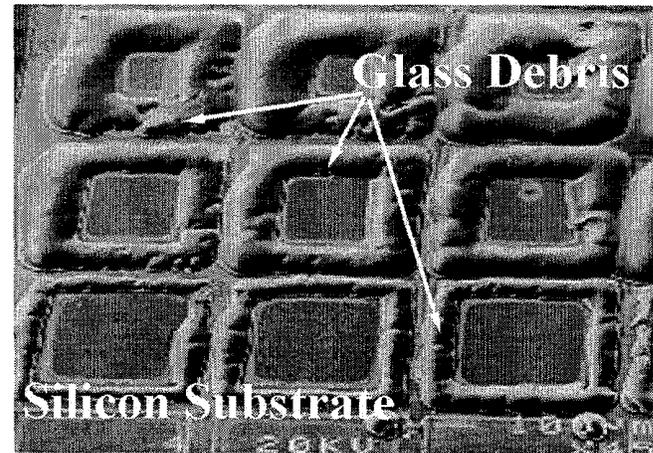
Figure 6: The SEM microphoto of silicon substrate showing glass debris after forcefully breaking the bond.

ditions by using RTP. The correlation of bonding temperature and time is governed by the Arrenius equation [10] as depicted by the straight lines. It is found that the activation energy for aluminum-glass RTP bonding is 2.8 eV as compared with 1.0 eV by using localized heating and bonding [10]. The discrepancy is probably due to different bonding conditions in establishing bonding curves such as history of applied pressure and temperature. In the RTP bonding curve in Fig. 3, only 2 seconds at 990±10°C is needed to complete the bonding. On the other hand, it takes 30 minutes to complete the bonding for the lowest bonding temperature at 760°C.

Figure 4 is the bonding result under an optical microscope. Cavities with area of 200x200, 300x300 and 400x400 μm^2 are sealed by aluminum-glass bonding and the widths of aluminum solder are 150, 100 and 50 μm, respectively. Leak tests are conducted by immersing sealed cavities into IPA(Isopropanol Alcohol). In this particular case, one of the cavities at the right bottom site is not fully sealed such that IPA went into the cavity. Other cavities have survived the leak test since no sign of IPA penetration can be identified.

The aluminum-glass bond is forcefully broken to examine the bonding interface. Figure 5 is a close view SEM microphoto of the glass wafer after the bonds are forcefully broken. The originally flat glass wafer now has fractures along the bonding

boundaries. For the samples with 100 μm wide aluminum solder, glass fracture is measured to penetrate about 25 μm deep from the flat surface and to spread about 35 μm away from the boundary of aluminum solder. These indicate that very strong bonding is achieved. Furthermore, it has been observed that parts of glass debris are now attached to the aluminum solder on the silicon substrate as shown in Fig. 6. Therefore, it can be concluded that the bonding strength by RTP is comparable to the fracture toughness of glass at around 10 MPa. Occasionally, incomplete bonding area can be identified as the leakage path as shown in Fig. 7. In this case, a small area of aluminum solder is not bonded with glass probably due to poor mechanical contact between the solder and glass cap during the RTP bonding process. This problem may be solved by applying a proper pressure on the glass cap and the device substrate to assist intimate contact during the RTP bonding process.
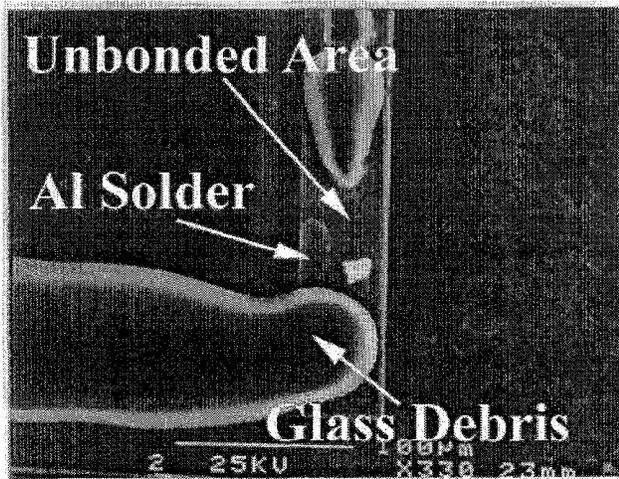
Figure 7: Non-uniform bonding is identified as the path of leakage.

Two kinds of hermeticity tests are applied: (1) IPA leak test and (2) autoclave test under 120°C with 20.7Psi of steam for 80 minutes. IPA leak test is chosen because it is easier for IPA than water to penetrate small bonding defects in detecting gross leakage. As expected, cavities with thick and wide aluminum solder have better survival rate. It is found that the survival rate reaches 100 % if the aluminum solder has width of 150 μm and thickness of 3.8 μm. However, even in the worst case when the solder width is 50 μm and the thickness is 3.2 μm, more than 40 % of cavities still passed the IPA leak test. The major cause of failure is believed to be the lack of proper bonding pressure. In the current bonding set up, the bonding pressure comes only from the gravity force of the glass cap. As a result, non-uniform pressure distribution is expected during the RTP bonding process and failure occurs when the glass cap and the aluminum solder failed to have intimate contact. A mechanical fixture is currently under design to provide better bonding pressure and to improve the RTP bonding results. After the gross IPA leak test, packaged cavities are put into autoclave chamber for tests under harsh environment. The average survival rate is 75 % and reaches 100 % for cavities with wide and thick (such as 150 μm wide and 4.5 μm thick) aluminum solder. RTP bonding may have introduced severe thermal stress and planted possible failure mechanisms to cause long term stability problems. The autoclave testing serves as

accelerated testing and the results indicate that good reliability by RTP bonding can be accomplished if design parameters are optimized and bonding process is properly executed.
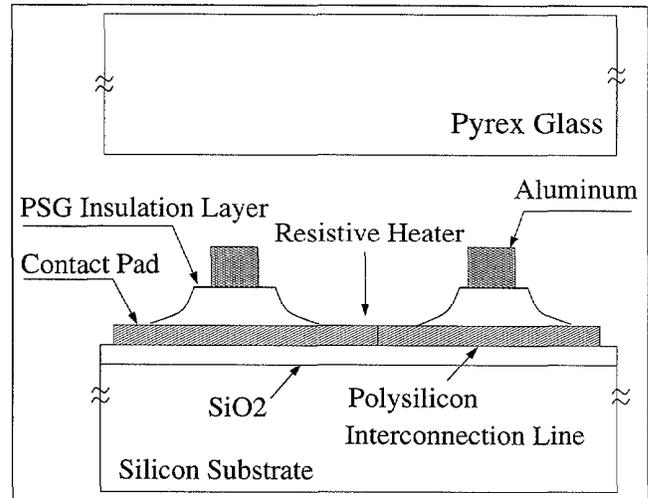
Figure 8: The cross sectional view of MEMS device packaging using aluminum-Pyrex glass bonding system by RTP.

RTP BONDING FOR MEMS PACKAGING

Figure 8 shows the cross sectional view of a micro resistive heater array with integrated aluminum solder at the boundary as a demonstration example of RTP bonding for MEMS packaging. First, a silicon substrate is grown with a 2 μm thick thermal oxide layer. It is followed by a 2 μm thick undoped LPCVD polysilicon deposition. The first mask is then used to pattern the heater array, electrical interconnection lines and contact pads. A 4.5 μm LPCVD PSG is then deposited as an electrical insulation layer. The wafer is then annealed at 1000°C for 2 hours as the drive-in process to make polysilicon conductive. Aluminum solder of 4.5 μm thick and 50 μm wide is then evaporated and patterned by the lift-off process to form the encapsulation area. The final mask opens the device and contact pad area by selective wet etching of PSG.

After going through the RTP bonding process at 990°C for 2 seconds, the packaged device is immersed in IPA as shown is Fig. 9. The device area is 400x400 μm^2. From the contrast color in the figure, IPA stays outside and does not go inside the cavity for more than 20 days and this long term testing is still going. Shown in Fig. 10 is one of the micro heaters working under the glass package. It is found that the resistance of the micro heater changes from 0.8 KΩ to 0.7 KΩ before and after RTP bonding process probably due to the activation of dopant. Although further characterizations are required, no major damages to the microelectronics are expected for this RTP bonding process.

After forcefully breaking the bond, the silicon substrate is shown in Fig. 11. The bond appears to be strong and uniform such that breakages occur on glass cap and some glass debris are left on the silicon substrate. A portion of PSG on the top surface of polysilicon interconnection lines seems to react partially with glass substrate as shown in the same figure. This implies that glass-glass fusion bonding may also be achieved by RTP processing. Furthermore, the polysilicon interconnection line creates a 2 μm step higher than the surrounding areas. The fact that hermetic bonding was achieved proves that this

aluminum-glass RTP bonding process can overcome the surface roughness of 2 μm or less during the bonding process.
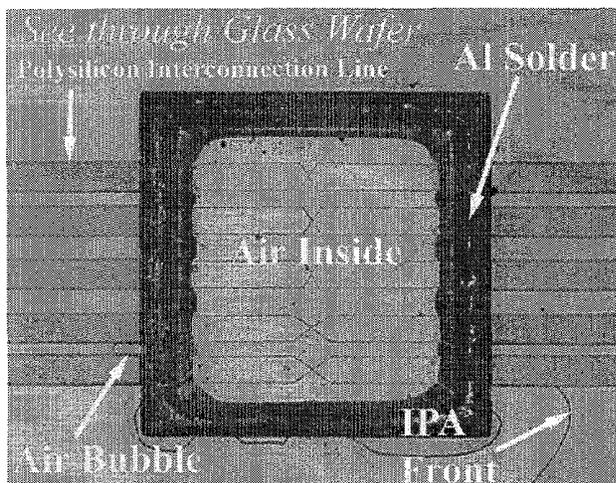
Figure 9: Hermetically sealed MEMS resistive heaters by RTP bonding under IPA leak test.
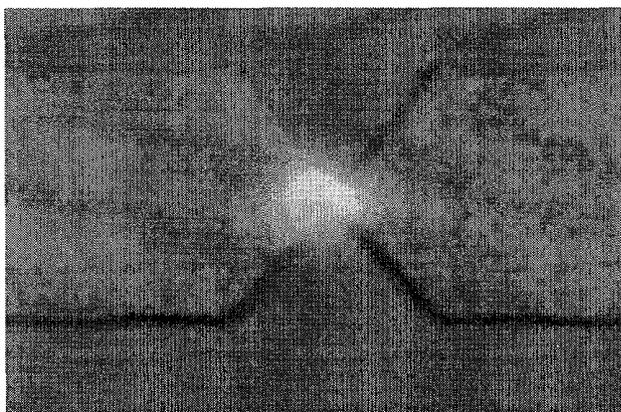
Figure 10: Glowing heater under glass package.

CONCLUSION

Hermetic wafer bonding by Rapid Thermal Processing on aluminum-glass bonding system has been demonstrated. RTP bonding characteristics, including width and thickness of aluminum solder, process time versus temperature and bonding interface have been investigated. It is found that the activation energy of RTP bonding for aluminum-to-glass system is 2.8 eV. Packaged dies have been put into both IPA and autoclave tests for reliability evaluation. It is found that if bonding parameters and process are carefully chosen, 100 % survival rate can be achieved. A micro heater array has been hermetically packaged by using the RTP bonding process and three important conclusions have been drawn. First, RTP bonding successfully overcome the 2 μm step caused by interconnection line. Second, hermetic sealing is accomplished because the packaged cavity has passed IPA test. Third, the device is operative after the RTP bonding.

ACKNOWLEDGEMENT
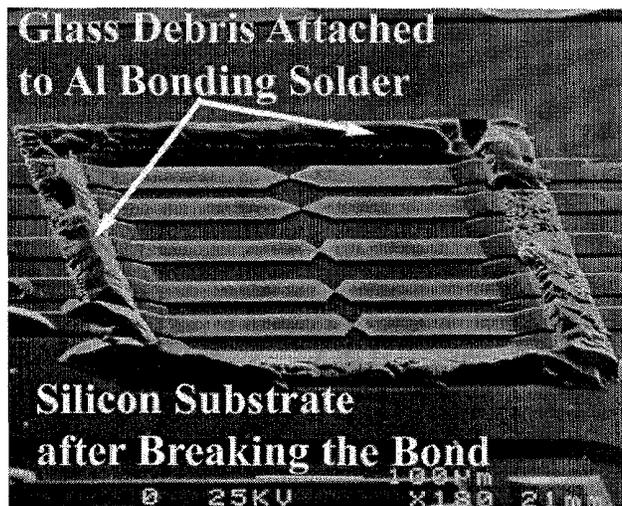
Figure 11: Silicon substrate after breaking the bond.

REFERENCES

[1] C. Harendt, H.G. Graf, B Hofflinger, and J. Penteker. Silicon fusion bonding and its characterization. *J. of Micromech. and Microeng.*, 2:113–116, 1992.

[2] M. B. Cohn. *Assembly Techniques for Microelectromechanical Systems*. PhD thesis, University of California at Berkeley, 1997.

[3] J. S. Hwang. *Solder Paste in Electronics Packaging*. Van Nostrand Reinhold, 1989.

[4] Y.T. Cheng, L Lin, and K Najafi. Localized silicon fusion and eutetic bonding for mems fabrication and packaging. *Journal of Microelectromechanical Systems*, 9(1):3–9, March 2000.

[5] Y.T. Cheng, L Lin, and K Najafi. Reliability of hermetic encapsulation by localized aluminum/silicon-to-glass bonding. In *Proceedings of IEEE Micro Electro Mechanical Systems*, pages 757–762, Jan 2000.

[6] P. J. Timans. Rapid thermal processing technology for the 21st century. *Materials Science in Semiconductor Processing*, 1:169–179, 1998.

[7] *Heatpulse 210T, STEAG RTP Systems Inc.* 4425 Fortran Drive San Jose, CA 95134-2300 USA.

[8] C. B. Cooper III and R. A. Powell. The use of rapid thermal processing to control dopant redistribution during formation of tantalum and molybdenum silicide/n$^+$ polysilicon bilayers. *IEEE Electron Device Letters*, EDL-6(5):234–236, May 1985.

[9] A. A. Pasa, J. P. de Souza, I. J. R. Baumvol, and F. L. Freire Jr. Dopants redistribution during titanium-disilicide formation by rapid thermal processing. *Journal of Applied Physics*, 61(3):1228–1230, Feb. 1987.

[10] Y.T. Cheng, L Lin, and K Najafi. Formation and characterization of the localized aluminum-glass and aluminum/silicon-glass bond. Submitted to J. App. Phy.

MESHLESS TECHNIQUES FOR EFFICIENT SIMULATION OF NON-LINEAR BEHAVIOR IN ELECTROSTATIC MEMS

Gang Li and N. R. Aluru*
Beckman Institute for Advanced Science and Technology
University of Illinois at Urbana-Champaign
Urbana, IL 61801

ABSTRACT

Accurate analysis of electrostatic MEMS and the development of reliable reduced-order (or macro) models for rapid system-level analysis requires proper understanding and the characterization of the non-linearity in electrostatic devices. The use of linear theories instead of non-linear theories can produce inaccurate results and the use of non-linear theories, in situations where linear theories can produce accurate results, can be very expensive and time consuming. Hence, it is critical to identify the regimes where linear and non-linear theories are valid. Simulation of non-linear behavior in electrostatic devices using classical finite-difference and finite-element methods can be very inefficient because of complex meshing requirements. We propose new non-linear meshless techniques for efficient simulation of electrostatic MEMS. The key results of this paper are two-fold: (i) the development of new meshless techniques for radically simpler analysis of electrostatic MEMS and (ii) the identification of the regimes where linear and non-linear theories are accurate and the various parameters that contribute to the non-linearity in electrostatic devices.

INTRODUCTION

Meshless techniques [1-3] are extremely powerful numerical methods as they do not require the generation of a mesh for complex three-dimensional microelectromechanical devices and systems (MEMS). Meshless techniques radically simplify the computer-aided design (CAD) for MEMS and significantly reduce the design cycle time for innovative MEMS. In this paper, we present a new finite cloud meshless method for geometrically non-linear analysis of electrostatic devices. The key features of the method are (i) it is a true meshless technique -- which means that only points need to be sprinkled and no connectivity information among the points is necessary (ii) a fixed reproducing kernel approximation is employed for constructing the interpolation functions and (iii) a collocation method is used for discretizing the governing equations.

By using the finite-cloud meshless method we investigate the regimes where linear and non-linear theories are suitable for electrostatic MEMS. The peak deflection and pull-in voltage obtained with linear and non-linear theories for a cantilever beam and a fixed-fixed beam with various lengths and gaps are studied. It is shown that for a small gap, both linear and non-linear theories give identical results, which are in excellent agreement with published data for pull-in voltages [5]. However, as the gap/length ratio increases, with all other variables held fixed, significant

deviations are observed between linear and nonlinear theories. Specifically, we have observed that, for larger gaps in comparison with the length, the pull-in voltages obtained with non-linear theories are accurate, and the linear theories are in significant error. The comparison between linear and non-linear theories for the fixed-fixed beam is much more pronounced. As the gap to length ratio increases, the error in pull-in voltage with linear theory increases significantly. From the results presented in this paper, it is evident that the gap/length ratio is one of the key variables in determining whether a linear or a non-linear theory should be employed for electrostatic MEMS design.

GOVERNING EQUATIONS

We will be primarily concerned with two-dimensional devices in this paper. For small micro-mechanical deformations, the linear theory of elasticity as defined by equations (1)-(3) can be used.

$$\nabla \cdot \tau + b = 0 \qquad in \ \Omega \qquad (1)$$

$$u = g \qquad on \ \Gamma_g \qquad (2)$$

$$\tau \cdot n = h \qquad on \ \Gamma_h \qquad (3)$$

In the above equations, τ is the Cauchy stress, b is the body force, g is the specified displacement on the boundary portion Γ_g and h is the surface traction on the boundary Γ_h, n is the unit outward normal. In coupled electro-mechanical analysis, the surface tractions are nothing but the electrostatic forces. When the electrostatic forces are dominant, the micro-mechanical structure can undergo large deformations, in which case the linear theory described by equations (1)-(3) may produce inaccurate results.

When an elastic body undergoes large deformation, the difference between the initial configuration and the deformed configuration of the body can not be neglected as was done for the case of linear elasticity. Instead, Eq. (1) is only valid for the deformed body whose configuration is unknown. While Cauchy stress and engineering strain are used as stress and strain measures for linear theory, the second Piola-Kirchhoff stress and the Green-Lagrangian strain are used as the stress and strain measures, respectively, for the geometrically non-linear analysis. The governing equations for large deformation analysis can be described by using either a total Lagrangian or an updated Lagrangian description (see [3] for differences between the two descriptions). In this paper we use a total Lagrangian description, in which all the variables reside in the initial configuration. The governing equations for a large deformation case are summarized below:

$$\nabla \cdot (FS) + B = 0 \qquad in \ \Omega \qquad (4)$$

$$u = G \qquad on \ \Gamma_g \qquad (5)$$

$$P \cdot N = H \qquad on \ \Gamma_h \qquad (6)$$

where F is the deformation gradient

$$F = I + \nabla u \qquad (7)$$

u is the displacement from the initial configuration X to the deformed configuration x, $x = X + u$, S is the second Piola-Kirchhoff stress given by:

$$S = CE \tag{8}$$

where C is the material tensor and E is the Green-Lagrangian strain:

$$E = \frac{1}{2}(F^T F - I) \tag{9}$$

B is the body force per unit undeformed volume and H is the surface traction per unit undeformed area. P is the first Piola-Kirchhoff stress tensor given by

$$P = FS \tag{10}$$

FINITE CLOUD MESHLESS METHOD

The finite cloud meshless (FCM) method uses a fixed reproducing kernel approximation to construct the interpolation functions and a point collocation technique for discretizing the governing partial differential equations [1]. In a fixed reproducing kernel technique, an approximation $u^a(x,y)$ to an unknown $u(x,y)$ is given by

$$u^a(x,y) = \int_\Omega C(x,y,s,t)\varphi(x_K - s, y_K - t)u(s,t)dsdt \tag{11}$$

Where the correction function $C(x,y,s,t)$ is given by

$$C(x,y,s,t) = P^T(s,t)C(x,y) \tag{12}$$

φ is the kernel function which is usually taken as a cubic spline or a Gaussian function. P is the vector of basis functions

$$P^T = \begin{cases} [1,s,t] & m=3 \\ [1,s,t,s^2,st,t^2] & m=6 \end{cases} \tag{13}$$

$C(x,y)$ are the unknown correction function coefficients computed by satisfying the reproducing conditions, i.e.

$$\int_\Omega P^T(s,t)C(x,y)\varphi(x_K - s, y_k - t)p_i(s,t)dsdt = p_i(x,y) \tag{14}$$
$$i = 1,2,...,m$$

In discrete form, Eq. (14) can be written as

$$\sum_{I=1}^{NP} P^T(x_I,y_I)C(x,y)\varphi(x_K-x_I, y_K-y_I)p_i(x_I,y_I)\Delta V_I = p_i(x,y) \tag{15}$$
$$i = 1,2,...,m$$

Eq. (15) can be rewritten in a matrix form as

$$MC(x,y) = P(x,y) \tag{16}$$

$$M_{ij} = \sum_{I=1}^{NP} p_i(x_I,y_I)\varphi(x_K - x_I, y_K - y_I)p_j(x_I,y_I)\Delta V_I$$
$$i,j = 1,2,...,m \tag{17}$$

From Eq. (16), the unknown correction function coefficients are computed as

$$C(x,y) = M^{-1}P(x,y) \tag{18}$$

Substituting the correction function coefficients into Eq. (11) and employing a discrete approximation for Eq. (11), we obtain

$$u^a(x,y) = \sum_{I=1}^{NP} N_I(x,y)\hat{u}_I \tag{19}$$

where $\hat{u}_I$ is the nodal parameter for node I, and $N_I(x,y)$ is the fixed reproducing kernel interpolation function defined as

$$N_I(x,y) = P^T(x,y)M^{-1}P(x_I,y_I)\varphi(x_K-x_I, y_K-y_I)\Delta V_I \tag{20}$$

The derivatives of the unknown u are approximated by

$$\frac{\partial u^a}{\partial x} = \sum_{I=1}^{NP} \frac{\partial N_I(x,y)}{\partial x}\hat{u}_I \tag{21}$$

$$\frac{\partial^2 u^a}{\partial x^2} = \sum_{I=1}^{NP} \frac{\partial^2 N_I(x,y)}{\partial x^2}\hat{u}_I \tag{22}$$

FCM FOR LARGE DEFORMATION ANALYSIS

For large deformation analysis the displacements u, v and their derivatives are approximated by using Eq. (20)-(22). Substituting the approximations into Eq. (4), one gets a non-linear system of equations

$$f_j(\hat{u}_1, \hat{u}_2,...,\hat{u}_{NP}, \hat{v}_1, \hat{v}_2,...,\hat{v}_{NP}) = 0 \qquad j = 1,2 \tag{23}$$

This non-linear system can be solved by using a Newton's method. Boundary conditions are enforced on the boundary points where Dirichlet or Neumann boundary conditions are specified. The governing equations for such boundary points are

$$u - G = 0 \qquad on \ \Gamma_g$$
$$P \cdot N - H = 0 \qquad on \ \Gamma_h \tag{24}$$

The accuracy of FCM is demonstrated by computing the large deformation of a cantilever beam. In Figure 1, we consider the large deformation of a 10μm x 1μm x 1μm cantilever beam. In Figure 2, the results obtained with the meshless method are compared with ANSYS results. The good match between the two results establishes the accuracy of the non-linear meshless technique. In Figure 3, we simulate the same beam by using a random distribution of 455 points. A comparison with the analytical solution is shown in Figure 4. This result clearly establishes the significance of the method – by simply sprinkling points randomly, we can obtain accurate results without having to go through the complex process of generating meshes for micro-machined devices.

MESHLESS ANALYSIS OF ELECTROSTATIC MEMS

The behavior of many microswitches can be characterized by studying the deformation of cantilever and fixed-fixed beams subjected to electrostatic potentials. The linear and non-linear models presented in Eq. (1)-(3) and Eq. (4)-(6), respectively, are solved when the surface tractions are the electrostatic forces.

In Figures 5-10, we investigate the regimes where linear and non-linear theories are appropriate. Cantilever and fixed-fixed beams with a width of 10 μm and a thickness of 0.5 μm are considered. Figure 5 and 6 compare the peak deflection and pull-in voltage obtained with linear and non-linear theories for a fixed-fixed beam with a length of 70 μm and two gap lengths. For a small gap length of 0.5μm (shown in Figure 5), we observe that both linear and non-linear theories give identical results. However, for a large gap length of 2μm (shown in Figure 6), we observe that the pull-in voltage is significantly different. More extensive studies for the cantilever beam with lengths varying from 10μm to 100μm and gap/length ratios varying from 1% to 60% are shown in Figures 7 and 8. For small gaps in comparison with the length of the beam, we observe that the pull-in voltages obtained with both linear and non-linear theories are very close. However, for larger gaps in comparison with the length, we observe that the pull-in voltages obtained with linear theories are in significant error compared to non-linear theories. The comparison between linear and non-linear theories for the fixed-fixed beam, shown in Figures 9 and 10, is more pronounced. As the gap to length ratio increases, the error in pull-in voltage with the linear theory

increases significantly. Furthermore, for the fixed-fixed beam example, for a fixed gap/length ratio, the change of length has a significant effect on the pull-in voltage error while this effect is not very pronounced in the cantilever beam case.

CONCLUSIONS

The primary contributions of the paper are summarized as follows: (a) For electrostatic structures, linear theories hold over a very small regime (gap/length ratio of less than 10% for cantilever beams and less than 1% for fixed-fixed beams). For accurate simulations or for the development of reduced-order models, designers need to use non-linear theories appropriately (b) We have developed new non-linear meshless techniques to efficiently and accurately simulate the non-linear behavior of electrostatic devices. These approaches offer radically simpler and easy-to-use CAD tools for MEMS designers.

REFERENCES

1. N. R. Aluru and Gang Li, "Finite Cloud Method: A True Meshless Technique Based On A Fixed Reproducing Kernel Approximation", *Submitted for publication*(1999).

2. N. R. Aluru, "A Point Collocation Method Based On Reproducing Kernel Approximations", *International Journal For Numerical Methods In Engineering*, Vol. 47, pp. 1083 – 1121(2000).

3. N. R. Aluru, "A Reproducing Kernel Particle Method for meshless analysis of microelectromechanical systems", *Computational Mechanics*, Vol. 23, pp. 324-338(1999).

4. K. J. Bathe, E. Ramm, and E. L. Wilson, "Finite Element Formulations For Large Deformation Dynamic Analysis", *International Journal For Numerical Methods In Engineering*, Vol. 9, pp. 353 – 386(1975).

5. P. M. Osterberg and S. D. Senturia, "M-Test: A Test Chip for MEMS Material Property Measurement Using Electrostatically Actuated Test Structures", *Journal of Microelectromechanical Systems*, Vol. 6, No. 2, pp. 107 – 118(1997).

Figure 2. Point refinement to improve the accuracy of the solution.
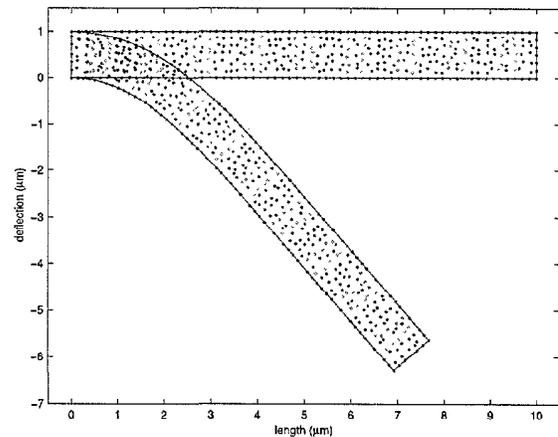
Figure 3. Deformed cantilever beam using a random distribution of points.
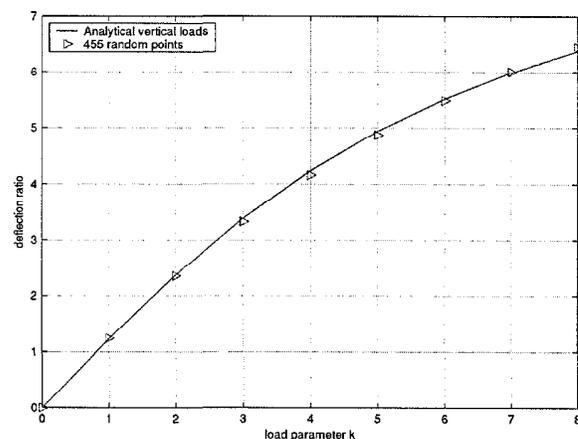
Figure 4. Comparison of the meshless solution (obtained with a random distribution of points) with the analytical solution.

Figure 1. Deformed cantilever beam using a uniform point distribution.
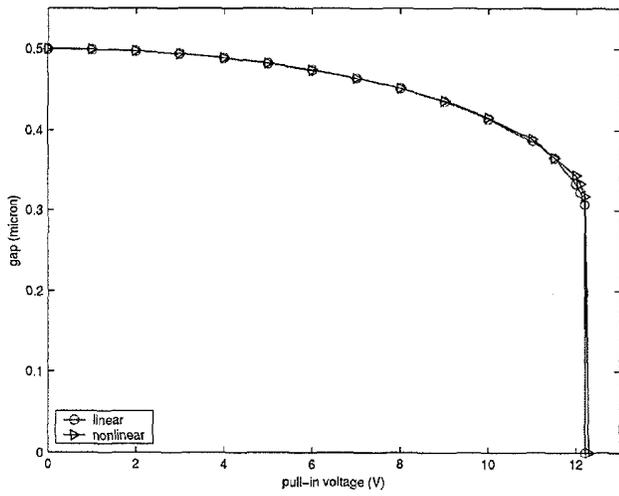
Figure 5. Comparison of Linear and non-linear results for a fixed-fixed beam with a length of 70μm and a gap of 0.5 μm.
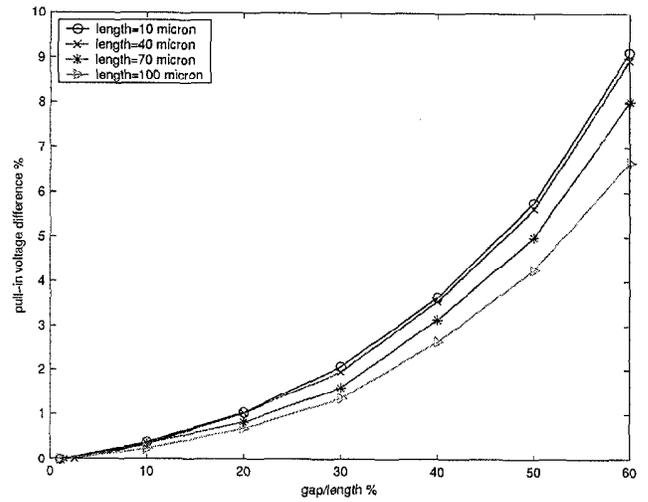
Figure 8. Gap/length ratio vs. pull-in voltage error between linear and non-linear results for cantilever beams.
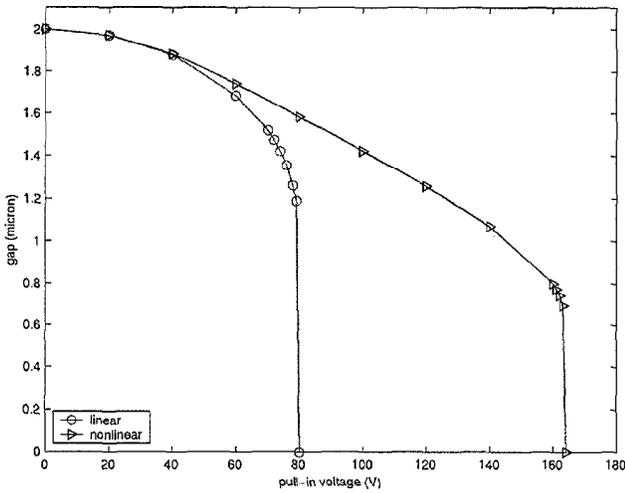
Figure 6. Comparison of Linear and non-linear results for a fixed-fixed beam with a length of 70μm and a gap of 2 μm.
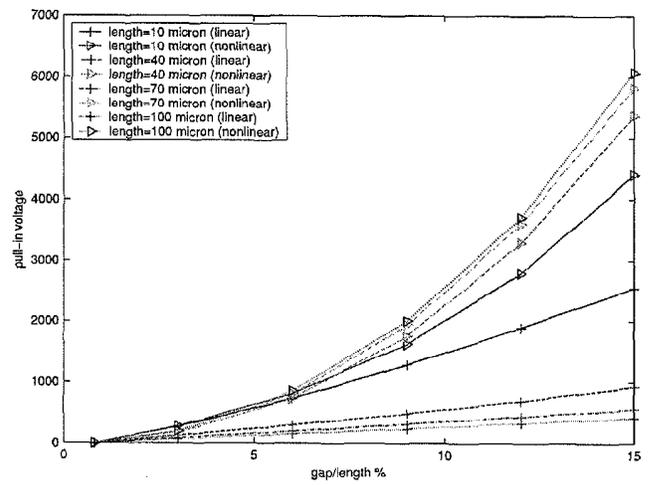
Figure 9. Gap/length ratio vs. pull-in voltage for fixed-fixed beams.
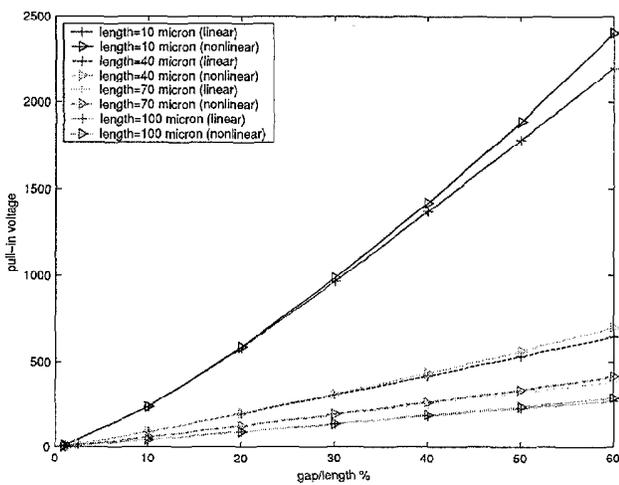
Figure 7. Gap/length ratio vs. pull-in voltage for cantilever beams.
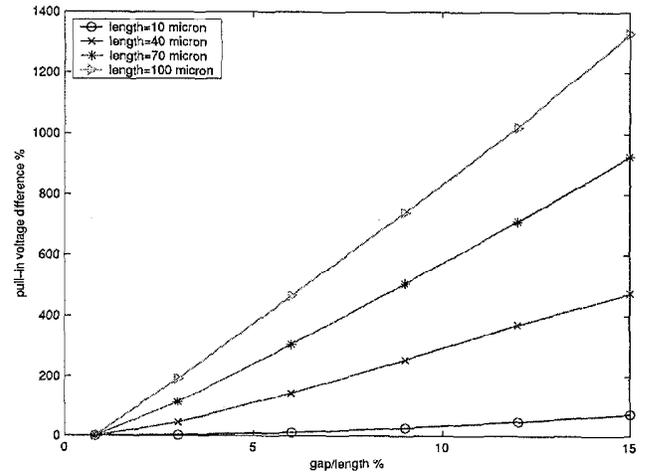
Figure 10. Gap/length ratio vs. pull-in voltage error between linear and non-linear results for fixed-fixed beams.

EXTRACTION OF FREQUENCY DEPENDENT MACRO-MODELS FOR GAS DAMPING AND SPRING EFFECTS FOR MEMS DEVICES

Yao-Joe Yang, Mattan Kamon, Vladimir L. Rabinovich, Chahid Ghaddar, Manish Deshpande, Ken Greiner, and John R. Gilbert

Microcosm Technologies Inc.,
Cambridge, MA 02142, USA

ABSTRACT

In this paper, we present an efficient macromodel extraction technique for gas damping and spring effects for arbitrarily shaped geometry. The technique applies an Arnoldi-based model-order-reduction algorithm to generate low-order models from a FEM approximation of the linearized Reynolds equation. We demonstrate that this approach for generating the frequency-dependent gas-damping model is more than 100 times faster than previous approaches, which solves the linearized Reynolds equation using a transient finite-element method. The low-order gas-damping model can be easily inserted into a system-level modeling package for transient and frequency analysis. The simulated results are in good agreement with experimental results for four different devices.

INTRODUCTION

A large class of MEMS devices must operate at significant gas pressures. To understand the dynamic behavior of these devices the effects of the gas surrounding the movable component can be critical. Devices such as accelerometers, gyroscopes, switches, optical modulators, micromirrors, and resonant sensors all require an understanding of gas damping and spring effects for accurate modeling.

Most work in damping has focused on getting more accurate simulations of the small or large signal 3D gas damping and spring effects [1-7]. The typical results of such analysis are similar to Figure 1, which shows the gas damping and spring components vs. frequency for the microrelay [8] shown in Figure 2(a). This paper is about the development of efficient techniques to take the solid models of full 3D structures and automatically extract spice-like time domain system models to accurately represent the frequency dependent damping and spring forces.
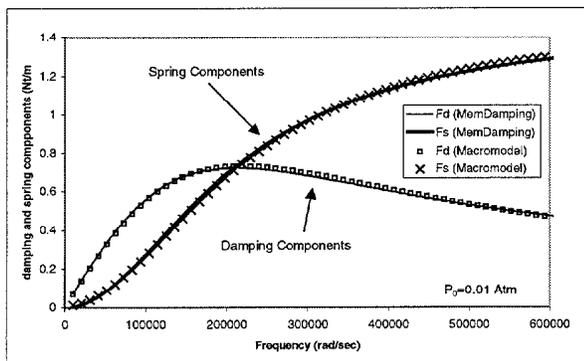
Figure 1 *Results of 3D physics analysis of gas spring and damping for beam switch of Figure 2(a), operating in its first vibration mode. The solid lines are obtained using MemDamping [9]. The points show the overlap of the macromodel obtained from the Arnoldi-based model-order-reduction technique.*

The extraction procedure is shown in Figure 3. A 3D model of a MEMS device is built using MEMCAD [9]. The finite-element (FE) matrix system for solving transient small signal analysis such as NSR or linearized Reynolds is constructed and then reduced to a low order system using the Arnoldi-based PRIMA algorithm [10]. After the reduced system is scaled with proper gap-dependent coefficients, it is then converted to a time domain representation written in an HDL, in this case MAST for use in SABER.

(a)   (b)
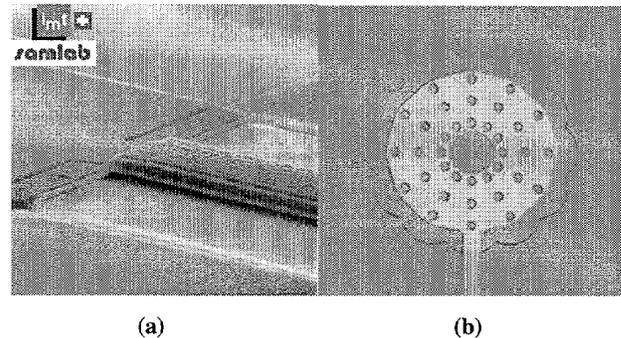
Figure 2 (a) *SEM picture of a microrelay from IMT.* (b) *An optical micrograph of a high frequency optical modulator from Lucent Technology.*
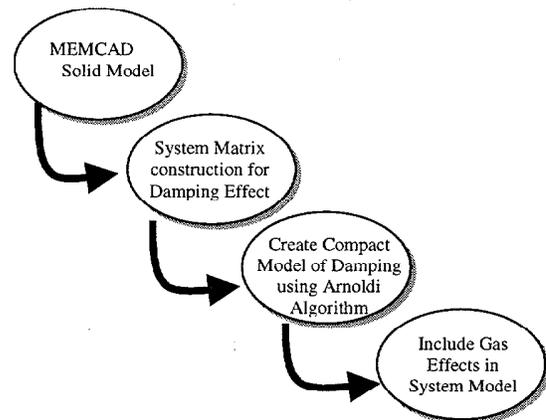
Figure 3 *Procedure of the efficient and accurate air damping modeling described in this paper*

Note that this new approach does not require the meshing of air gaps that contribute to gas-damping effects. The effect of perforations and the effect of viscous damping on plate edges [7] are also incorporated into the FE matrix system in the model construction step. Because the model is generated directly from the FE system, the transient calculations of the full FEM model for a wide frequency range, as reported in [5], is not performed. This brings the total computation time to obtain a frequency-

dependent gas damping model to a few minutes vs. a few hours/days in previous approaches [3,5,7].

This work complements prior work in macro-model extraction for electromechanics that has succeeded in extracting 6 DOF models for a general electromechanical plate on tether systems [11]. We can now extract full systems such as the electromechanical switch schematic of Figure 4.

THEORY

The models are extracted using the Arnoldi-based model reduction algorithm known as PRIMA [10], which is commonly used for model reduction of electrical interconnects. Our approach is similar to the model reduction approach used in [12]. In [12], model reduction was applied to a linearized form of the fully coupled electro-mechanical-fluid damped system. Here, we treat the squeeze film damping separately and build models in the mechanical mode shape basis. Such an approach would allow these models to be readily combined with the mode shape based models in the low order model of the entire system [13]. Additionally, we show that the models can be used for large-signal motion with certain restrictions.

To begin, the linearized Reynolds equation for squeeze film damping from [1] is

$$\frac{h_0}{P_0} \cdot \frac{\partial p}{\partial t} = \frac{h_0^3}{12\mu} \nabla^2 p - \frac{\partial e}{\partial t} \quad (1)$$

where the variation in plate spacing h is assumed to be small compared to the mean spacing, h_0, given by

$$h = h_0 + e(x,t)$$

with $x \in \mathfrak{R}^2$ and $e \ll h_0$. The variation in pressure, P, will thus be small compared to ambient,

$$P = P_0 + p(x,t)$$

Equation (1) can be solved by finite element analysis for a given $e(x,t)$. Let $f(x)$ be the shape (perhaps mode shape) of the displacement so that $e(x,t) = f(x) \cdot u(t)$. The dynamic system from discretizing (1) by finite elements can be written in state space form as

$$\frac{h_0}{P_0} B \frac{d\mathbf{p}}{dt} = \frac{h_0^3}{12\mu} A\mathbf{p} + Bfu(t) \quad (2)$$

$$y = (Bf)^T \mathbf{p}$$

where $A, B \in \mathfrak{R}^{n \times n}$, where n is the number of nodal degrees of freedom, $\mathbf{p}$ is the pressure at the nodes, and f is $f(x)$ evaluated at the node points. y is then the net force projected into the shape defined by $f(x)$ in a finite element sense.

The dynamic system of (2) is too large to insert directly into a system simulator such as SPICE or SABER. We thus apply PRIMA to generate a low order representation of (2) which still accurately captures the dynamic behavior. To apply the PRIMA algorithm to generate a k-th order model, k orthogonal vectors $\{v_i\} \in \mathfrak{R}^n$ are computed which span the vector space known as Krylov subspace:

$$K_k = \left\{ (Bf), A^{-1}B(Bf),..., (A^{-1}B)^{k-1}(Bf) \right\}$$

These vectors can be stably computed via the Arnoldi algorithm [10]. Given the matrix V whose columns are $\{v_i\}$, the reduced order model is

$$\frac{h_0}{P_0} \tilde{B} \frac{d\tilde{\mathbf{p}}}{dt} = \frac{h_0^3}{12\mu} \tilde{A}\tilde{\mathbf{p}} + \tilde{\mathbf{f}}u(t) \quad (3)$$

$$y = \tilde{\mathbf{f}}^T \tilde{\mathbf{p}}$$

where $\tilde{B} = V^T BV$, $\tilde{A} = V^T AV$, and $\tilde{\mathbf{f}} = V^T Bf$. The attractive properties of such an approach are that the first k Taylor series coefficients of the transfer function of (3) match those of the original model in (1). In addition, the model is guaranteed to be passive. Finally, the method easily extends to a single model with multiple inputs $\{u_1(x), u_2(x),....\}$ corresponding to multiple mode shapes, $\{\mathbf{f}_1(x), \mathbf{f}_2(x),....\}$.

As will be seen in the next section, k = 5 is generally adequate for an accurate damping model. This low order model can be inserted directly into a system simulator such as SPICE or SABER. Note that since the vector space spanned by $\{v_i\}$ does not depend on the mean gap, ambient pressure or viscosity, the above model is valid for any choice of those parameters. Going one step further, we can model large signal behavior by letting the mean gap vary with time, $h_0 = h_0(t)$. Such an approach would be valid if $h_0(t)$ varies slowly compared to $u(t)$. In fact, from numerical experiment, we find that replacing h_0 in (2) with $h_0(t) = h_0 + u(t)$ for even large $u(t)$ gives good results.

COMPARISON TO EXPERIMENTAL RESULTS

Once we obtain a reduced-order model (macromodel) for gas damping from the PRIMA algorithm, an HDL template of the gas model written in MAST is automatically generated. A Saber system-level model of a whole MEMS device with this frequency-dependant gas-damping macromodel can be easily built using Saber, and small signal as well as transient analyses can be simulated. Figure 4 is an example of a Saber system-level model of a microswitch [14], which includes a lumped inertia, two mechanical springs, two automatically-generated frequency-dependent gas-damping models for two gaps, a electrostatic force actuator with a voltage source, and two mechanical stoppers.
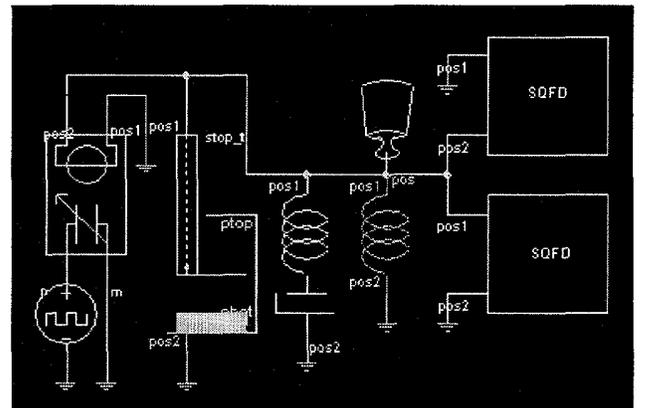
Figure 4 *Schematic of a full MEMS switch [13] model including components for an inertia, two mechanical springs, an electrostatic actuator, a voltage source, two mechanical stoppers and two extracted macromodels for gas induced forces.*

A. Small signal oscillation with or without mode shapes

Three experimental data sets are used to verify the gas-damping models generated by the Arnoldi-based algorithm: a microrelay from IMT [8], optical modulators from Lucent Technology [15] (shown in Figure 2(b)), and a low-frequency accelerometer from Motorola (shown in Figure 5). Figure 6 shows the frequency response (small AC analysis by Saber) of an IMT microrelay for different ambient pressures. The fixed-fixed-beam microrelay operates in its first oscillation mode. The changes in quality factors as well as resonance shifting effects can be easily observed in the figure as pressure varies.
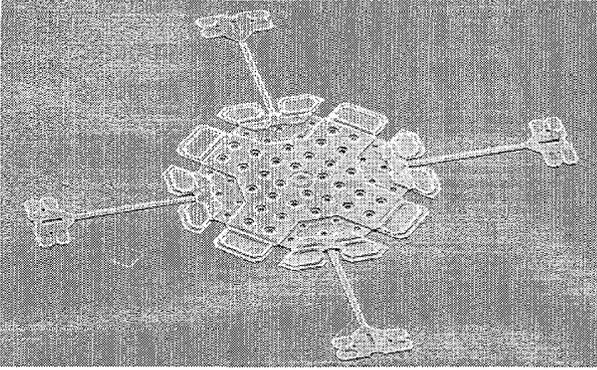
Figure 5 *A SEM of the Motorola accelerometer*
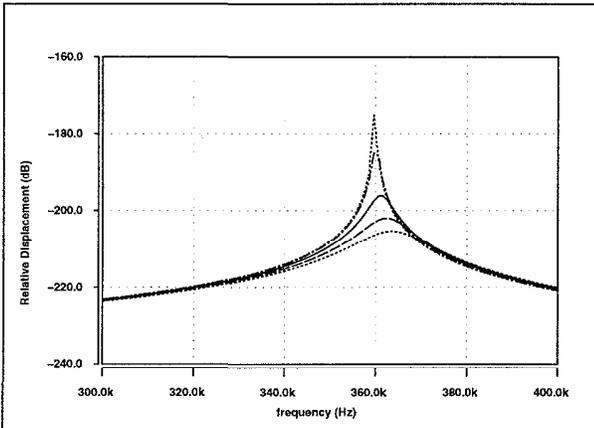
Figure 6 *Frequency response of an IMT microrelay for different pressures of 0.1, 1, 5, 10 and 15 mbar (curves from top to bottom). The quality factors and resonance shifting effects can be easily extracted from the curves.*

Figure 7 shows the gas damping coefficients vs. frequency for these three devices. The gas spring constant vs. frequency is shown in Figure 8. The simulated results in these figures are generated by reduced models of order 5. Note that the experimental and simulated results are in good agreement spanning about 6 orders of magnitude in frequency.

The comparison of computational times for calculating the frequency-dependent damping the spring components (e.g., Figure 1) is shown in Table 1. In the previous approaches [3,5,7], at least 20 transient simulations with different small-amplitude-oscillation frequencies were needed to generate the damping and spring components spanning across a desired frequency range. The new approach is a least 100 times faster than the previous approaches. It also automatically builds a time-domain model for system-level analysis.
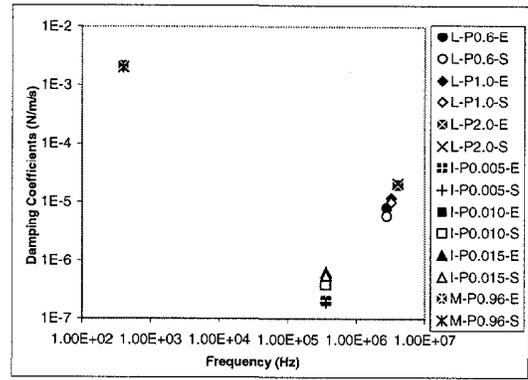
Figure 7 *Gas damping coefficients vs. Frequency for three devices. (L: Lucent, M: Motorola, I: IMT; P#.#: pressure at #.# bar; E: experiment data, S: simulation data)*
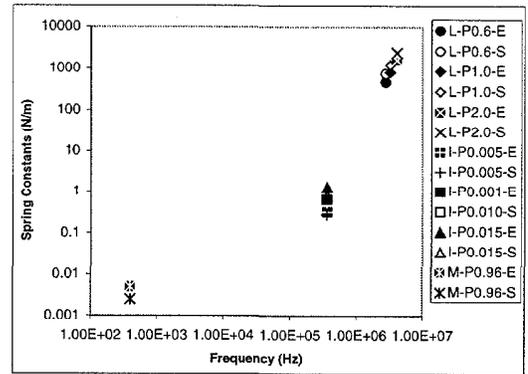
Figure 8 *Gas spring constants vs. Frequency for three devices. (L: Lucent, M: Motorola, I: IMT; P#.#: pressure at #.# bar; E: experiment data, S: simulation data)*

	Total Computation Times	
	Previous Approach	New Approach
Lucent	948 min	1.5 min
Motorola	431 min	1.1 min
IMT	200 min	0.4 min

Table 1 *Comparison of computation times between previous and new approaches*

B. Transient Analysis

Figure 9 shows experimental and simulated transient responses of a microswitch which schematic is shown in Figure 4. A step voltage applied between the plate and substrate for 1 ms. The plate collapses on the stopper, then is released from the substrate (after turning off the applied voltage) and oscillates around its neutral position.

The large amplitude macromodel transient results are also compared with the full non-linear Reynolds simulations [3] using finite-difference method (FDM). Figure 10 and Figure 11 show the transient results for a square plate on tethers over a counter electrode with applying step voltages of 2 Volts (below pull-in voltage) and 3 Volts (above pull-in voltage). The ambient pressure is 100 N/m^2, the gap is 2 μm, and the plate is 500x500 μm^2. Below pull-in voltage, the macromodel and the FDM results match very well, as shown in Figure 10. However, the results of the macromodel with fixed h_0 (see equation 2) underestimates the damping as the plate moves away from its original gap.

When applied voltage is larger than pull-in voltage, the plate velocity increases significantly when plate is close to touch down, which in turn significantly increases the pressure back force against the plate. This large pressure increase breaks the linearized Reynolds equation assumption that the pressure variation is much less than ambient pressure. The macromodel result diverges from the FDM result after the plate travels across half gap, as shown in Figure 11, due to this highly non-linear pressure force. The result of the macromodel with fixed h_0 starts to diverge from the FDM result after the plate travels about a quarter of the gap.
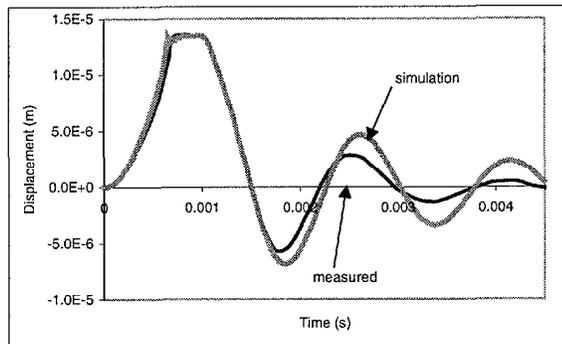
Figure 9 *Comparison between Experimental transient in MEMS switch and simulation using extracted macromodels for gas damping and spring effects. Experimental data is from Minami, et. al [14].*
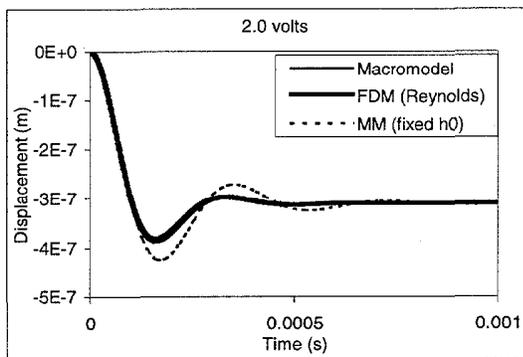
Figure 10 *A comparison among macromodel, macromodel with fixed h_0, and full non-linear Reynolds for large displacement transient behavior. The applying voltage is 2 V.*
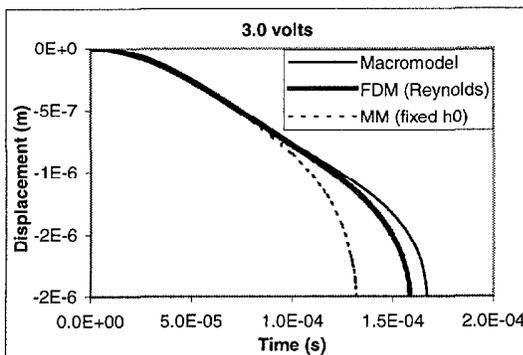
Figure 11 *A comparison among macromodel, macromodel with fixed h_0, and full non-linear Reynolds for pull-in transient behavior. The applying voltage is 3 V.*

CONCLUSION

A new approach to extract frequency-dependent gas damping models for arbitrary geometries is demonstrated. The Arnoldi-based algorithm is applied for creating a low-order model from the transient FE system matrices of NSR analysis. The frequency-dependant gas damping and spring effects can be obtained using the low-order models without any computationally intensive transient simulation for wide frequency range. After constructing 3D solid models, more than two order-of-magnitude reduction in computational time has been demonstrated. The numerical results of small signal gas damping and spring effects have shown in good agreement with experimental results. Examples of large-amplitude transient analysis and comparison to experimental results are also provided.

ACKNOWLEDGEMENTS

REFERENCES

1. W. E. Langlois, "Isothermal Squeeze Films," Quar. Applied Mathematics, Vol. XX, No. 2, 1962, pp. 131-150.
2. J.J. Blech, "On Isothermal Squeeze Films", Journal of Lubrication Technology, Vol. 105, 1983, 615-620.
3. Y.-J. Yang and S. D. Senturia, "Numerical Simulation of Compressible Squeezed-Film Damping," Tech. Digest, Solid State Sensor and Actuator Workshop Hilton Head Island, SC, June 1996,
4. T. Veijola, et. al., "Model for Gas Film Damping in a Silicon Accelerometer", Proceedings Transducers 97 Vol. II, Chicago, IL, USA June 16-19, 1997, pp. 1097-1100
5. Y-J. Yang, et. al., "Effects of Air Damping on the Dynamics of Nonuniform Deformations of Microstructures", Transducers 97 Vol. II, Chicago, IL, USA June 16-19 1997, 1093-1096
6. T. Veijola, "Finite-Difference Large Displacement Gas-Film Model", Transducers 99, Sendai, Japan, Jun 7-10, 1999, pp1152-1155.
7. M. da Silva, et. al., "Gas Damping and Sprng Effects on MEMS Devices with Multiple Perforations and Multiple Gaps", Transducers '99, Sendai, Japan, Jun 7-10, 1999, pp1148-1151.
8. M-A. Gretillat, "Electrostatic Polysilicon Micro-relays" Ph.D. Thesis, IMT, University of Neuchâtel, Switzerland 1997.
9. MEMCAD 4 Module Guide, Microcosm Technology Inc. 1999.
10. A. Odabasioglu, el. al., "PRIMA," IEEE Transaction on Computer-Aided Dedign of Integrated Circuits and Systems, Vol. 17, No. 8, August 1998.
11. M.H. Zaman, et. al. "A Technique for Extraction of Macro-Models in System level Simulation of Inertial Electro-Mechanical Micro Systems", MSM '99 pp 163-167. April 19-21 San Juan PR.
12. F. Wang and J. White, "Automatic Model Order Reduction fo a Microdevice using the Arnoldi Approach," ASME IMECE 98, DSC-Vol. 66, pp527-530.
13. G. K. Ananthasuresh, et. al., "An Approach to Macromodeling of MEMS for Nonlinear Dynamic Simulation," ASME International Mechanical Engineering Congress and Exposition, Symposium on MEMS, Atlanta, GA, Nov. 1996, pp. 18-22.
14. K. Minami, et. al. "Simple Modeling and Simulation of the Squeeze Film Effect and Transient Response of The MEMS Device," MEMS 99, pp. 338-343.
15. D.S. Greywall, el. al., "Phenomenological Model for Gas Damping of Micromechanical Structures", Private Communication, Lucent Technologies, NJ 1998

UNDERSTANDING PARAMETRIC RESONANCE EFFECTS IN COMMON MEM ACTUATORS

Kimberly L. Turner and Noel C. MacDonald
Department of Mechanical & Environmental Engineering
University of California, 2355 Engineering Building II, Santa Barbara, CA 93106
Phone: (805) 893-5106, Fax: (805) 893-8651, email: turner@engineering.ucsb.edu

ABSTRACT

Parametric resonance has been previously observed in electrostatically actuated torsional microelectromechanical systems (MEMS) [1], but it can occur in other common MEMS devices, such as in-plane resonators which use non-overlapping comb drives or parallel-plate drives. Parametric resonance can occur whenever the force generated by a periodically-driven actuator changes with position, characteristically creating very large amplitude responses at specific frequencies. In this paper, two distinct actuator types are analyzed, and experimental results are shown. Parametric resonance has been verified for a non-overlapping comb drive actuator, and the fundamental differences between a parametrically excited state and a directly forced state are shown. We present a straightforward way of determining if parametric resonance is present in a MEM system.

INTRODUCTION

Deviations from 'expected' or 'designed' behavior is one mechanism of failure in MEMS devices. Failure can come from material fatigue, defects, or incorrect operation but also from system dynamics which are not accounted for in many simple models. There are certain regimes where the device can strongly deviate from linear approximations often used for design. Parametric resonance is just one such event which can occur, and cause device behavior to strongly deviate from the equations often used in design. Parametric resonance has been theoretically studied [2,3] for many years, and has been observed in the forced motion of a swing, stability of ships, Faraday crispations, along with others. Theory predicts that parametric resonances occur near drive frequencies which are certain fractions of the drive frequency. Thus the drive frequency can be a value far removed from resonance, yet still lead to a large amplitude, 'resonant' type of response from the device. Thus the phenomenon must be clearly understood, so that it can be utilized or avoided as necessary to maintain desired performance.

THEORY

In the systems discussed here, parametric resonances occur when [2]:

$$\omega = \omega_o/n \ (n=1,2,3...). \qquad [1]$$

When designing actuators to operate at or near resonance, it is important to understand how parametric resonance will affect the device operation, as it can occur at the same drive frequency as a standard resonant response.

PARALLEL PLATE ACTUATOR

Parametric resonance has been observed and studied in a variety of structures. One such structure is the well-known parallel-plate actuator (see Figure 1).
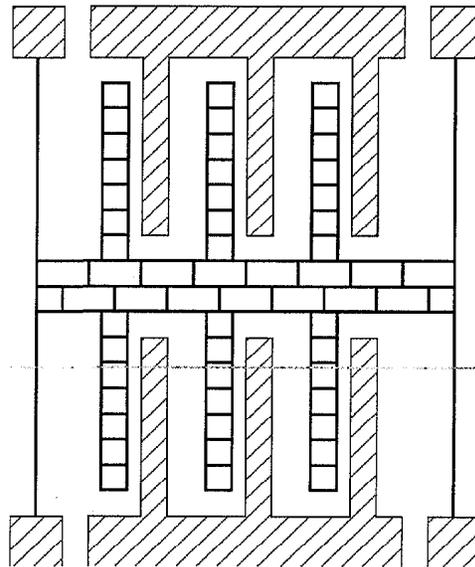
Figure 1. *Schematic of a MEMS oscillator with parallel-plate capacitors for actuation. The equation of motion for such an oscillator is given by Equation 1, where m is the mass, c is a damping coefficient, k is a linear stiffness term, A is the area of overlapping plates, d_o is the initial gap between movable and fixed plate, and ω is the drive frequency.*

The equation of motion of a parallel-plate capacitor when sinusoidally driven is (assuming linear springs and linearizing the parallel plate force):

$$m\ddot{x} + c\dot{x} + \left(\left(k - \frac{\varepsilon_o A V_{ac}^2}{2d_o^3} \right) + \frac{\varepsilon_o A V_{ac}^2}{2d_o^3} \cos(2\omega t) \right) x = \frac{\varepsilon_o A V_{ac}^2}{4d_o^2}(1 - \cos 2\omega t) \quad [2]$$

(variables defined in Figure 1, and V_{ac}=AC drive amplitude, ε_0=permittivity constant). Through numerical simulation, it is determined that the parametric forcing term has a large effect on this system. Figure 2 shows the normalized ($\tau=\omega t$) response of an equation driven at $\omega=0.5\omega_o$ ($n=2$, see Equation 1) with the parametric forcing term set equal to zero. Figure 3 shows the response of the normalized equation with a nonzero parametric forcing term. Note the differences between the two responses. Figure 2 displays a linear growth, which is expected for an externally forced system, whereas Figure 3 displays an exponential

growth, common to parametric resonance [2,3]. This behavior can be easily measured using motion characterization techniques available for MEMS.

Damping and applied signal strength have a significant effect on parametric resonant response, although not through amplitude. Primary *(n=1)* parametric resonance can only occur if the drive strength term, δ_0 (in this case, $\delta = \frac{\varepsilon_o A V_{ac}^2}{4\omega^2 d_o^3}$) is greater than 2μ, where $\mu = c/2\omega$. For $n=2$, $\delta^2 > 16\mu$. Thus, if parametric resonance is not desired, it can be avoided by adjusting the drive strength and/or operating environment according to the above condition.

Parametric resonance has been observed in parallel-plate actuators, novel comb actuators, and torsional resonators. Using precise experimental testing techniques, it is now straightforward to distinguish between parametric resonance and directly forced resonance in MEMS devices. Understanding the dynamics of resonant MEMS leads to more intelligent, robust designs and a greater range of applications which can be developed.
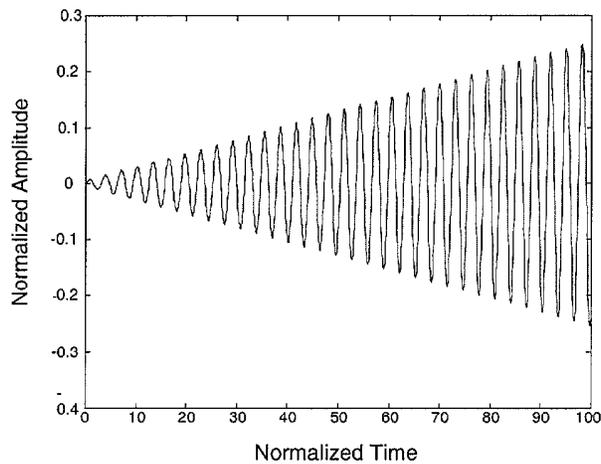
Figure 2. *Normalized resonant response of a non-parametric resonator. Note the linear growth in amplitude inherent in undamped second-order ODE's. The normalized equation solved was:* $\ddot{x} + \beta x = e\cos 2\tau$, *where* $\beta = 4$, $e = 0.01$
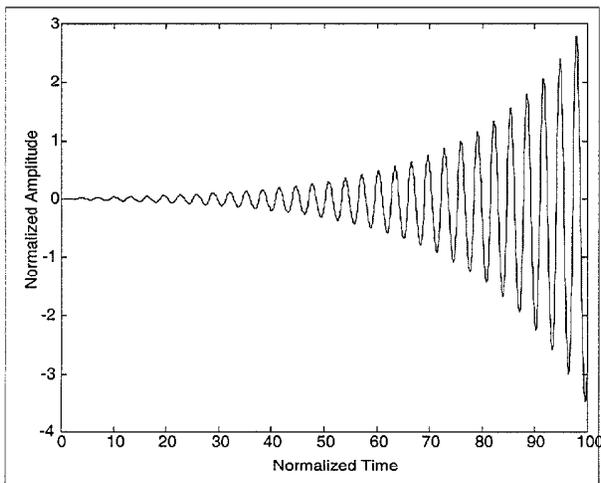
Figure 3. *Normalized response of above system, with parametric driving term:* $\ddot{x} + (\beta + 2\delta\cos 2\tau) = e\cos 2\tau$, *where* $\beta = 4$, $\delta = 0.5$, *and*

$e=0.01$. *Note the exponential growth present due to the parametric term. Also note the larger amplitude obtained within the same time.*

NON-OVERLAPPING COMB DRIVE ACTUATORS

Other novel actuator designs have been developed which display parametric resonance behaviors. Some important actuators which display this behavior are the novel fringing-field actuators which were designed by Adams, et. Al [4] for the independent tuning of linear and cubic stiffness terms [4,5]. The actuators resemble comb drive actuators, except that the fingers are not interdigitated. Figure 5 shows a diagram of the actuator configurations. The top of the figure shows the reduction configuration (if the actuator moves to the left, the net force toward the left is increased), while the bottom describes the augmentation configuration (if the actuator moves to the left, the net electrostatic force is toward the right). These actuators, unlike standard comb-drive actuators, move perpendicular to the direction of the fingers, as shown schematically in the figure. They move due to fringing electric fields between the movable and fixed fingers.

Depending on the original position of the movable fingers with respect to the fixed fingers, the electrostatic spring constant acts negative or positive. A schematic showing the force vs. displacement is shown in Figure 5. This graph shows the general trend of this type of actuator. Specific simulations and experiments of force vs. displacement relations for this type of actuator have been performed by Adams et. Al. [5]. If you examine Figure 5, it is easy to identify two distinct linear regions. In the region surrounding points marked by the open squares, the curve has a negative slope (augmentation), and for the region surrounding points marked by open circles (reduction), the curve has a positive slope. Thus, depending on the initial position of the actuator, the behavior of the device can be quite different.
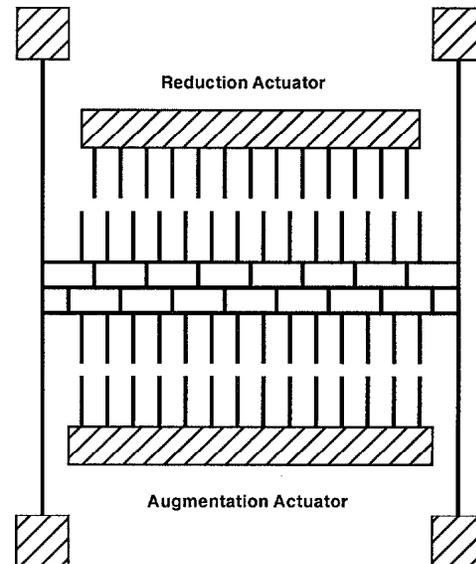
Figure 4. *Schematic of a non-overlapping comb drive actuator, suspended from fixed-fixed springs.*

360
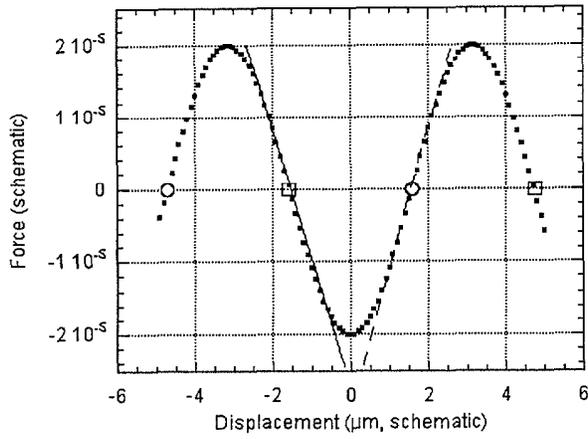
Figure 5. A schematic of the force curve showing different regions of the force relation for reduction and augmentation actuators. The open squares represent centers of the augmentation regions, while the open circles represent centers of the reduction regions.

An approximate equation of motion for each of these regions can be developed. Beginning with the standard equation of motion, and assuming that the devices have a fixed-fixed spring configuration, with mechanical stiffnesses which are linear with a small cubic term, we get:

$$m\ddot{x} + c\dot{x} + kx + hx^3 = F(x,t) \qquad (3)$$

where F depends on whether a reduction or augmentation configuration is used. First, consider the augmentation alignment, where the slope of F vs. position is negative. A sinusoidal driving term is applied, although a square-rooted sinusoid as developed by Turner et. Al. [1] can also be applied. For a sinusoidal driving term, the equation of motion becomes:

$$m\ddot{x} + c\dot{x} + kx + hx^3 = -\xi x(V\sin\omega t)^2 \qquad (4)$$

and if $V^2\sin^2\omega t$ is represented by trigonometric substitution as $V^2(1-\cos2\omega t)/2$, the equation becomes:

$$\ddot{x} + \frac{c}{m}\dot{x} + \frac{1}{m}\left[\left(k + \frac{V^2}{2}\xi\right) - \frac{V^2}{2}\xi\cos2\omega t\right]x + \frac{h}{m}x^3 = 0 \qquad (5)$$

which is a parametrically forced equation of the Mathieu type. When applying the same theory as presented in Chapter 2, instabilities are predicted where $\omega = \omega_0/n$, where $\omega_0^2 = (1/m)(k + V^2\xi/2))$. A small DC offset may be present in the input signal which can lead to a small term on the right hand side of the equation. Thus as in the parallel plate equation, there can be a small direct driving term, which can dominate at times in the response.

The reduction actuator has the same form as Equation 4, except the sign is reversed on the term proceeding x due to electrostatic forcing. Thus, the equation of motion appears as:

$$\ddot{x} + \frac{c}{m}\dot{x} + \frac{1}{m}\left[\left(k - \frac{V^2}{2}\xi\right) + \frac{V^2}{2}\xi\cos2\omega t\right]x = 0 \qquad (6)$$

which is a parametrically forced differential equation of Mathieu type, and respond as previously discussed. This equation also predicts instabilities at frequencies at $\omega = \omega_0/n$.

EXPERIMENT

The device shown in Figure 6 is a single-crystal Silicon MEMS actuator fabricated using the SCREAM process. It has comb drives on either end, and along the backbone, it has non-

overlapping comb drives, as shown schematically in Figure 2. The non-overlapping comb drives are used for this experiment, thus the equation of motion is similar to Equation 6. As seen in theory, the growth rate and envelope are distinctly different depending on whether parametric resonance is the primary reason for the amplitude growth. Thus by capturing this growth region, it can be determined if parametric resonance is present, and if it has a strong effect on the device behavior.

Using an oscilloscope, the time-series behavior can be captured. The motion is measured using a laser doppler vibrometer coupled through an optical microscope [6]. The device is in a pressure-controlled chamber. Figure 7 shows the input signal which produced the response in Figure 8. This input signal was at ~13kHz. The resonant frequency of the device is at ~26kHz. Figure 8 shows an experimental response typical of non-parametric oscillation. The response has a growth rate typical of an underdamped oscillator. The amplitude is limited in this case by both damping and nonlinearity [7].
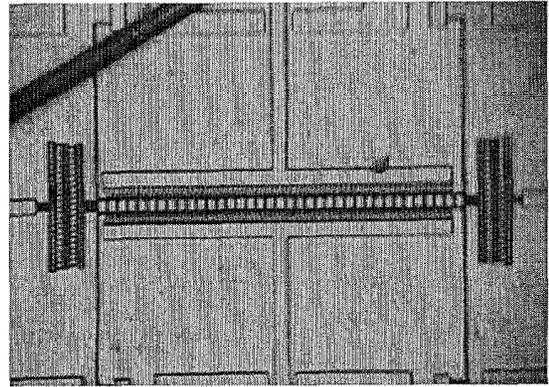
Figure 6. Device with novel comb actuators used for the experiments in this device. This device is being driven by actuators as shown in Figure 2, which are non-overlapping comb drives.
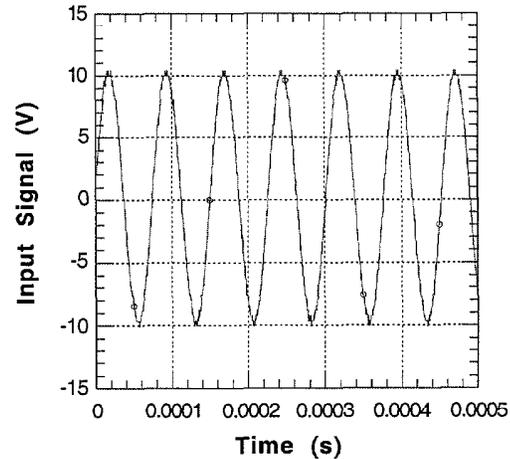
Figure 7. Input signal to the device at 13kHz. This response is dominated by direct forcing, and does not display parametric resonance behavior, as shown below. Parametric effects only occur at higher orders (n>1) in the presence of very low damping. Thus at this frequency,(corresponding to n=2) the parametric resonance response is not present.
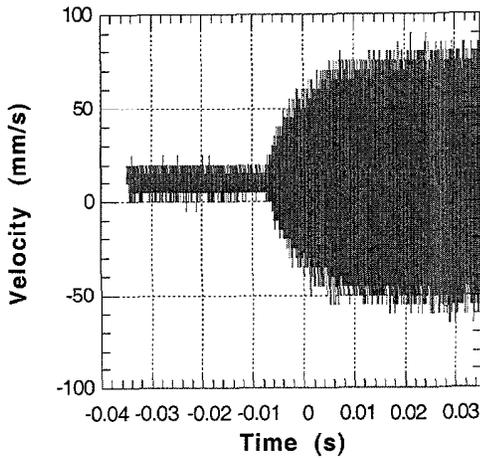
Figure 8. Turn on data when driving at a frequency which does not have a large parametric frequency response. As seen, this device has a response typical of a damped, driven oscillator. The limiting amplitude is primarily due to the damping present. The device was at ~200mTorr under test.

Directly forced oscillations are primarily limited by damping, which is not the case for parametric resonance. Thus, in some cases the parametric oscillations can lead to larger amplitude motions than direct resonance.

Figure 9 shows the input signal used to obtain the response curve shown in Figure 10. In this case, the device was driven at $\omega=\omega_0$, corresponding to the n=1 parametric instability region. This is the strongest parametric oscillation, and it is apparent from the exponential growth behavior that the response is largely due to parametric resonance.
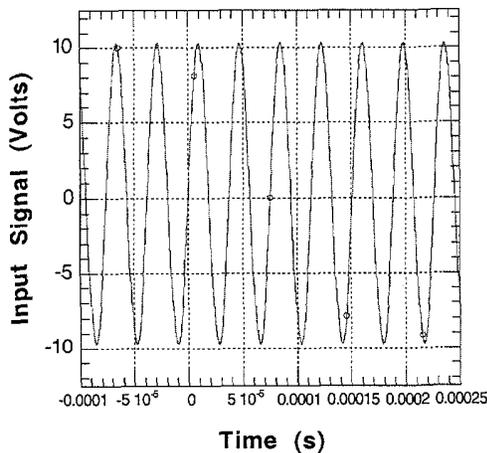
Figure 9. Input signal which produced the parametric resonance curve shown below. As seen in the simulations, the behavior strongly suggests a parametric resonant type of behavior.

The amplitude reaches a *maximum* which is not predicted by the linear theory. This maximum value is caused by a nonlinearity, not damping as in the previous case [7]. The nonlinearity can come from electrostatics as well as geometric nonlinearity, and is currently being investigated.
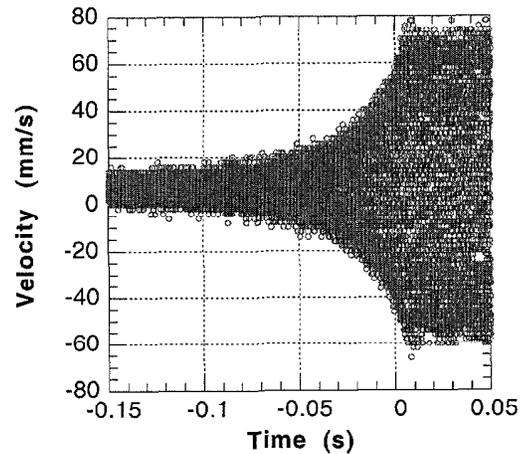
Figure 10. Turn-on data when driving in Parametric resonance n=1 instability region. Note the exponential growth behavior as predicted in numerical simulation, and also by perturbation theory. The pressure was ~100mTorr for this experiment.

CONCLUSION

As shown by theory as well as simulated and experimental results above, large amplitude responses can appear at frequencies other than the first natural frequency of common MEM devices. These behaviors are present in many types of MEM actuators, including non-overlapping comb drives, parallel plate actuators, and torsional actuators. Understanding and having simple tests to determine which type of resonance is present is key to understanding and designing robust and predictable MEMS for many applications.

ACKNOWLEGDMENTS

REFERENCES

[1] Turner, K. L. et. Al., " Five parametric resonances in a microelectromechanical system," *Nature, 396,* 149-153 (1998).

[2] Cunningham, W. J ., *Introduction to Nonlinear Analysis* , McGraw Hill, Inc., New York, 1958.

[3] Bolotin, V. V., *The Dynamic Stability of Elastic Systems.* Holden-Day, San Francisco. (1964).

[4] Adams, S. G., et. Al., "Independent Tuning of Linear and Nonlinear Stiffness Coefficients," *Journal of Microelectromechanical Systems,* 7, 2 (1998).

[5] Adams, S. G., "Capacitance based tunable resonators," *Journal of Micromechanics and Microengineering,* 8, 15-23 (1998).

[6] Turner, K. L., P. G. Hartwell, and N. C. Macdonald, "Three dimensional motion characterization using Laser Vibrometry," *Transducers '99,* Sendai, Japan, 7-10 June 1999, pp. 1144-1147.

[7] Nayfeh, A. H. and D. T. Mook, *Nonlinear Oscillations,* John Wiley and Sons, New York (1979).

ACTIVE FREQUENCY TUNING FOR MICRORESONATORS BY LOCALIZED THERMAL STRESSING EFFECTS

TODD REMTEMA[1] and LIWEI LIN
Department of Mechanical Engineering University of California, Berkeley
5126 Etcheverry Hall, MC 1740, CA 94720-1740
TEL: 510-643-5495, FAX: 510-643-5599, lwlin@me.berkeley.edu

ABSTRACT

Active frequency tuning for comb-shape micro resonators has been successfully demonstrated by means of localized stressing effects. A mechanical beam structure that can be resistively heated to generate thermal stress is integrated as part of the micro resonator for frequency tuning. Experimentally, frequency change up to 6.5% is measured for resonators with central frequency around 31kHz. The required tuning power is 25mW in the form of localized joule heating. Analytically, both a one-dimensional electrothermal model and a dynamic model are established to characterize electrical, thermal and frequency responses of active frequency tuning. The simulation results of frequency spectrum are consistent with experimental measurements. A reliability test has been conducted for more than 300 million cycles under 6.5% of frequency tuning range and reveals no material damages on the micro resonator. This scheme enables active frequency tuning under low power consumption, independent of the input/output function of a micro resonator. As such, it has potential applications to resonator based MEMS devices, such as rate gyroscopes and microelectromechanical filters.

INTRODUCTION

Micro resonators have been used in various resonator-based microsystems, such as resonant accelerometers [1,2], microelectromechanical filters [3-5], and gyroscopes [6]. One common roadblock in the manufacturing process comes from local process variations. As a result, even though the micro fabrication process is well-controlled, structural deviations exist between designed and fabricated devices and between same devices adjacent to each other on the same wafer. Therefore, it is common that the frequency responses of micro resonators may be off from the desired design value and should be eradicated by post-tuning processes. For example, up to several percentage variations of resonant frequency can be found in the widely used comb-shape micro resonators on the same chip [7].

Frequency tuning methods can be split into two major categories, those that make one time permanent changes and those that make active adjustments without permanent damages to microstructures. The permanent changes can be accomplished by annealing the folded springs of a comb structure using voltage pulses to change the stiffness and quality factor [8] or by a post-deposition process [9]. On the active frequency tuning side, three basic methods have been investigated. The first of these is tuning via an "electrostatic stiffness" change, where the stiffness can be adjusted electrostatically with capacitive structures [6,10-13]. The second method tunes by making changes in the internal stresses of the structure to induce changes in resonance [14,15]. The third method takes advantage of the inherent temperature dependence of Young's modulus [16]. The method presented in this work takes advantage of the temperature dependence of Young's modulus exhibited by most materials as well as the thermally induced internal stress effects. That is, as the temperature of polysilicon goes up its Young's modulus drops and this will cause a minor reduction in the resonant frequency. The localized thermal stress generated by joule heating plays a major role in the proposed frequency tuning scheme. If a compressive stress is generated, the structure becomes more compliant resulting in a lower resonant frequency. If a tensile stress is generated, the structure becomes stiffer resulting in a higher resonant frequency. Through the use of filament like heating as implemented in this work, the device heats up resulting in a decrease in Young's modulus and an increase in compressive stresses. Both effects help to actively lower the resonant frequency.

A TUNABLE MICRO RESONATOR

A tunable micro resonator by means of localized thermal stressing effects has been designed as illustrated in Fig. 1. The comb drive structure is modified by replacing one folded-beam suspension with a straight suspension as the tuning beam. The folded beam suspension helps minimizing the non-linear effect that may arise from the straight-beam structure design [17,18]. One drawback of this folded-beam/straight-beam design as opposed to a straight-beam/straight-beam design is that the tuning range of the device is reduced. However, most non-linear effects are avoided. The straight-beam suspension is 4μm wide and 165μm long. On the folded-beam suspension, the four beams are 2μm wide and 150μm long and the truss is 5μm wide and 36μm long. The anchors on the two sides of the suspension system are electrically isolated such that a voltage can be applied across the structure. This voltage heats the structure through filament like heating that causes the structure itself to heat up and to expand. During the heating process the substrate stays relatively cool due to the rapid heat transfer capabilities of silicon.
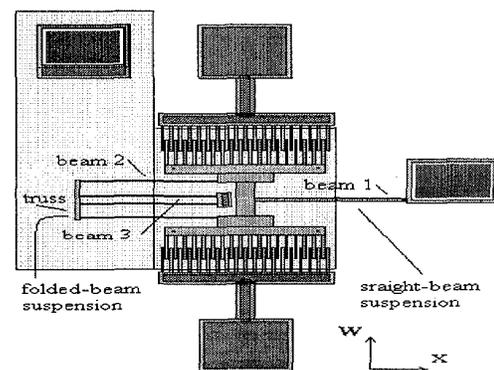
Fig. 1, Schematic diagram of a comb-shape micro resonator with a straight-beam for active frequency tuning via localized stressing effects.

---

[1] Mechanical Engineering and Applied Mechanics Department, the University of Michigan

A standard surface-micromachining process provided by MCNC [19] is used to fabricate these tunable micro resonators and all the simulation and experimental measurements presented here are based on the MUMPS-22 run. A sheet resistance of 6.4 Ω/square and a thermal coefficient of resistivity of 1.7 x 10^{-3}ohm/°K are measured for the structural polysilicon layer. Young's modulus and density of the polysilicon layer used in this work are 160 GPa and 2300 kg/m^3. The thermal coefficient of expansion and specific heat of polysilicon are adopted from silicon as 2.6 x 10^{-6}/°K and 712 J/(kg°K), respectively. The values of thermal conductivity of the silicon, polysilicon, and nitride layers used in this paper are 148, 34, and 1.6 W/(m°K).

The frequency responses of the device are measured using a CCD camera built into a probe station. Pictures are first taken and then opened in Canvas™ [20] and the amplitudes of vibration for every 100 Hz of driving frequency are measured and recorded. Measurements of resonance are obtained by adjusting the driving frequency until the peak driving amplitude of the device could be observed. This method could locate the resonant frequency within 100 Hz when the resonant frequency is typically around 30 kHz. This gives an accuracy of roughly 0.33%.

EXPERIMENTAL RESULTS AND DISCUSIONS

A dynamic model and a one-dimensional electro-thermal model have been established [21]. With this model, the transient responses of the temperature versus time can be graphed as seen in Fig. 2. The rising and falling times for the system to reach steady state under heating and cooling are found to be around 140µsec. Using the model developed above, a temperature profile of the system at a time significantly larger than the rise time can be attained as shown in Fig. 3. Nodes 1 to 25 are the inner portion of the folded-beam suspension, nodes 26 to 32 are the truss portion of the folded-beam suspension, nodes 33 to 58 are the outer portion of the folded-beam suspension, node 59 is the rigid central mass, and nodes 60 to 100 are the straight-beam suspension.
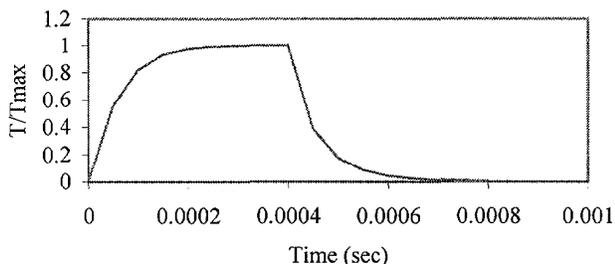
Fig. 2, Transient temperature simulation.
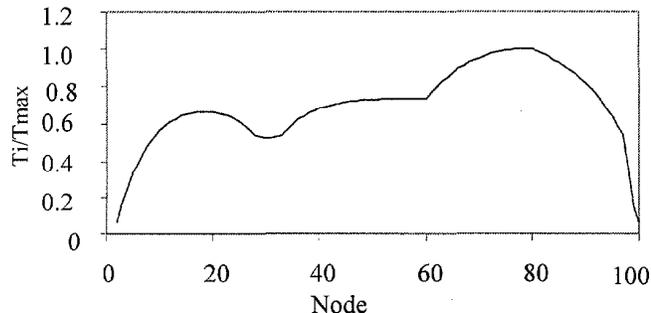
Fig. 3, Temperature profile simulated by the electro-thermal model.

Figure 4 is an optical photo showing the tunable micro resonator under a microscope. When a high tuning power is supplied for resistive heating of the system and the background light is turned off, Fig. 5 shows the hot spots. When Fig. 5 is compared with Fig. 3, the highest temperature region agrees well with what is seen experimentally. The temperature simulations under various input power are used to predict the thermal expansion of comb-drive to derive the tuning force P and frequency changes.
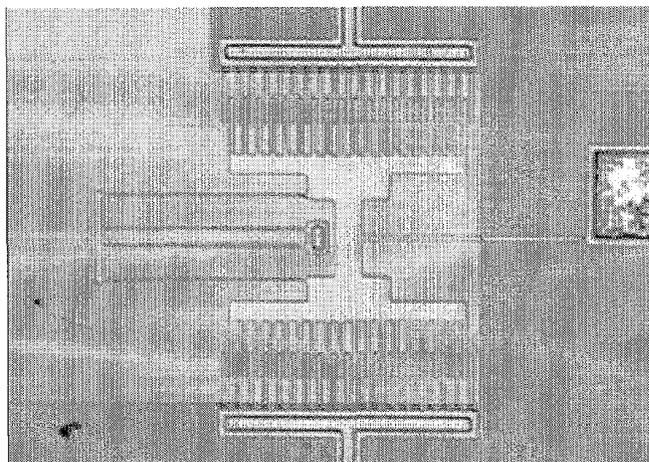
Fig. 4, A tunable resonator under an optical microscope.
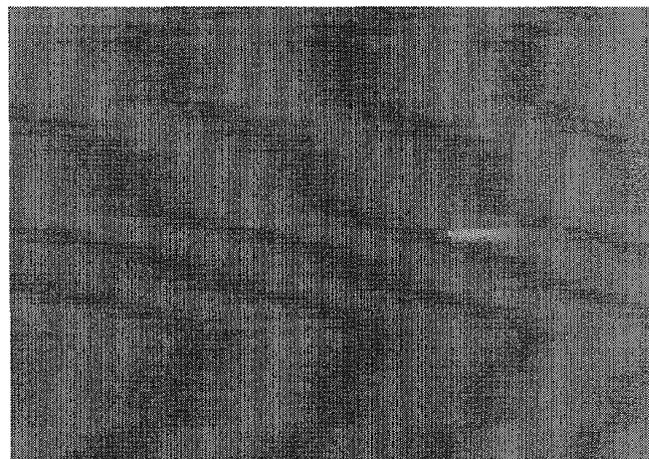
Fig. 5, Hot spot is shown under the joule heating effect when the illumination of the optical microscope is turned off.

Using the experimental setup described earlier, four basic tests have been conducted on the tunable resonator. The first test is a comparative test where the resonator is heated by external means and by localized joule heating, respectively. The external heating is accomplished by placing the device on a hot plate. In both tests, the resistance of the suspended beams is measured at various temperatures when resonant frequency change is measured. The resistance is used as the guidance to estimate the temperature of the device. Figure 6 shows the resonant frequency changes with respect to the resistance changes (temperature changes). It is found that when the resistance is at 940 ohm it corresponds to an average temperature of 150°C. At this temperature, the external heating causes a frequency change of about 1% and the internal joule heating makes the frequency change of 6.5%. Under this joule heating condition, the maximum temperature at the tuning beam is 255°C and the average temperature of the resonator is 150°C.
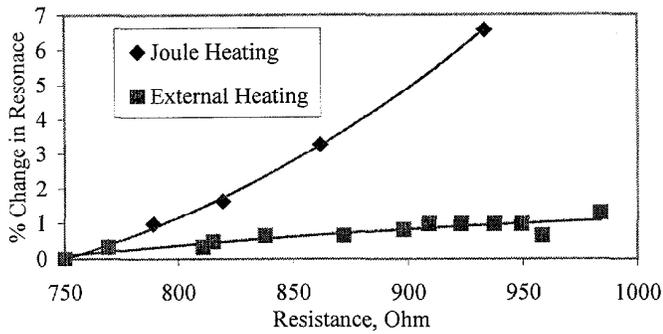
Fig. 6, Frequency changes by means of external (global) heating and joule (localized) heating.

It is concluded that under the same average temperature, the joule heating causes the resonance to change much more than the external heating. The external heating makes both the substrate and the device to expand such that this keeps the thermal stressing effects to a minimum. The change in resonance by external heating can be attributed mainly to the change in Young's modulus and partially to the change in density and thermal mismatch. The temperature coefficient of Young's modulus can be found using the following equation [3]:

$$TC_f = \frac{1}{2}(TC_E - TC_h) \qquad (1)$$

Where TC_f is the temperature coefficient of the resonant frequency, TC_E is the temperature coefficient of Young's modulus and TC_h is the temperature coefficient of thermal expansion. In this work, TC_E is found to be –40ppm/°C by the external heating experiments, which is close to previously reported values [3-5]. In the joule heating case the device itself heats up where as the substrate stays relatively cool. This creates a large internal stressing effect as well as a change in Young's modulus.

The second experiment is to check for repeatability and compliance of the dynamic and electrothermal models. In this test, five different structures from four different dice are tested over a range of tuning powers. The change in resonance versus the applied tuning power is plotted for 5 different devices along with the predicted theoretical response as shown in Fig. 7. As seen in the plot, the theory predicts a quasi-linear change in resonance versus input power. The experimental data points are close to the theoretical prediction when the input power is lower than 30mW. However, the model is only valid for a first-order approximation because it does not take into account several secondary effects, including non-linearity due to the softening of the spring as the structure deflects, the change in specific heat, thermal conductivity, and the thermal coefficient of expansion as the temperature increases.
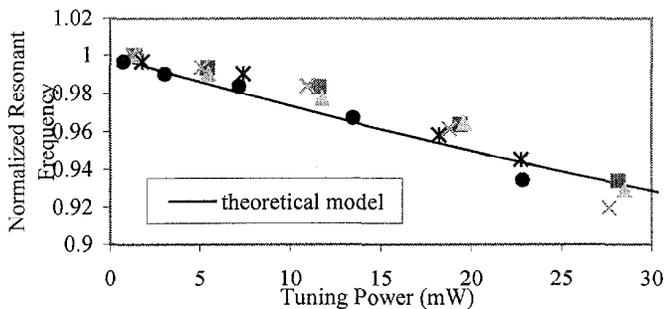
Fig. 7, Measured frequency changes vs. tuning power for 5 different devices compared to the theoretical model.

The third test is to measure the spectrum response of the resonator. One device is tested in this manner. The driving frequency is changed every 100 Hz and the amplitude is recorded. This is done by scanning both into the higher frequencies and into the lower frequencies until the amplitude is relatively low. Several tuning currents are conducted and recorded as shown in Fig. 8 from 0 to 5mA. As seen in the plot, the system starts out with a fairly linear response. As the tuning current is increased the response begins to go more and more non-linear. This effect is to be expected. Non-linearity arises from the stiffening of the straight-beam suspension as the device deflects. The stiffening comes from the stretching of the beam as it deflects. This stretching increases the restoring force of the spring. It is well documented that as the magnitude of the intrinsic stress goes up the non-linearity also goes up [12]. Also seen in the plot is that the quality factor of the system stays nearly the same at about 50. This is because most of the dampening comes from the friction with the air that remains relatively unchanged during the frequency tuning process.
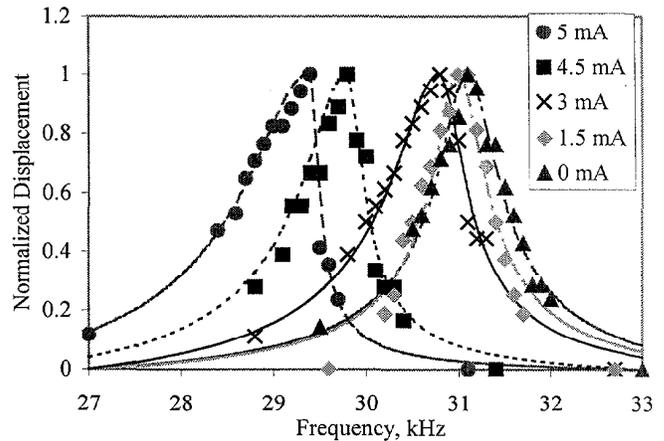
Fig. 8, Frequency spectrum of one device at 5 different tuning amperages from 0 to 5 mA.

The final test is the reliability experiment for the tuning method. A resonator is tuned to 1% and again to 6.5% for approximate 300 million cycles, respectively. At the end of these tests, no visual damages can be found on the device. Most importantly, it returns completely to its original natural frequency and the measured resistance does not change before and after the tests. This indicates that there is no structural degradation and the active tuning method via thermal stressing effects as presented is feasible for repeated, long-term applications. Although structural damage is not expected at the tuning power of 6.5% with a highest temperature at 255°C, further investigations should be conducted for the fatigue and creep [22] phenomena under high temperature and high stress operations.

CONCLUSION

An active frequency tuning mechanism has been proposed and has been proven to adjust the natural frequency of comb-shape micro resonators up to 6.5% with minimal power requirements at 25mW. This method can be applied at the component level after the completion of the packaging process and it can actively tune the frequency of micro resonators with a response time of 140μsec. It is important for microstructures to have frequency tuning capabilities in order to adjust for manufacturing variations in microsystems. Both dynamic and electrothermal models are established in this work to provide the design guidelines for active frequency tuning. Four

different kinds of experiments have been conducted. The experiment of external versus joule heating demonstrates that the thermal stressing effect is the dominant factor in the proposed tuning mechanism. The second experiment measured 5 different devices for frequency changes at various tuning powers. It is concluded that the analytical model is consistent with the experimental results. The third experiment measured the spectrum of comb-resonator at various tuning power. It is found that nonlinear effects begin at high tuning power probably due to the spring stiffening effect. Finally, long-term tests for 300 million cycles demonstrate the good reliability of the proposed tuning mechanism.

ACKNOWLEDGEMENT

REFERENCES

1. T.A. Roessig, R.T. Howe, A.P. Pisano and J.H. Smith, "Surface-micromachined resonant accelerometer," Intl. Conf. On Solid-State Sensors and Actuators (Transducers 97), Chicago, 1997, pp. 859-862.
2. Y. Omura, Y. Nonomura and O. Tabata, "New resonant accelerometer based on rigidity change," Intl. Conf. On Solid-State Sensors and Actuators (Transducers 97), Chicago, 1997, pp. 855-858.
3. L. Lin, R.T. Howe and A.P. Pisano, "Microelectromechanical filters for signal processing," Journal of Microelectromechanical systems. Vol. 7. No. 3. September 1998, pp. 286-294.
4. K. Wang and C. T.-C. Nguyen, "High-order micromechanical electronic filters," Proceedings of 1997 IEEE International Micro Electro Mechanical Systems Workshop, Nagoya, Japan, 1997, pp. 25-30.
5. C. T.-C. Nguyen, "Frequency-selective MEMS for miniaturized low-power communication devices," IEEE Transactions on Microwave Theory and Techniques, Vol. 47, No. 8, pp. 1486-1503.
6. Y. OH, B. Lee, S. Baek, H. Kim, J. Kim, S. Kang and C. Song, "Surface-micromachined tunable vibratory gyroscope," Proceedings, 1997 IEEE International Micro Electro Mechanical Systems Workshop, Nagoya, Japan, 1997, pp. 272-277.
7. W. C. Tang, T.C. H. Nguyen and R. T. Howe, "Laterally driven polysilicon resonant microstructures," Sensors and Actuators, Vol.20, 1989. pp.25-32.
8. K. Wang, A.-C. Wong, W.-T. Hsu and C. T.-C. Nguyen, "Frequency trimming and Q-factor enhancement of micromechanical resonators via localized filament anealing," Intl. Conf. On Solid-State Sensors and Actuators (Transducers 97), Chicago, 1997, pp. 109-112.
9. D. Joachim and L. Lin, "Localized deposition of polysilicon for MEMS post-fabrication processing," 1999 ASME International Mechanical Engineering Congress and Exposition, Proceedings of Microelectromechanical Systems, MEMS-Vol. 1, pp. 37-42, Nashville, Tennessee, 1999.
10. K. Lee and Y. Cho, "Frequency tuning of laterally driven microresonator using an electrostatic comb array of linearly varied length," Intl. Conf. On Solid-State Sensors and Actuators (Transducers 97), Chicago, 1997, pp. 113-116.

11. S.G. Adams, F.M. Bertsch, K.A. Shaw, P.G. Hartwell, F.C. Moon and N.C. Macdonald, "Capacitance based tunable resonators," J. micromech. Microeng. Vol. 8, 1998, pp. 15-23.
12. C. Gui, R. Legtenberg, H. Tilmans, J. Fluitman and M. Elwenspoek, "Nonlinearity and hysteresis of resonant strain gauges," Journal of Microelectromechanical systems. Vol. 7. No. 1. March 1998, pp. 122-127.
13. S. G. Adams, F. M. Bertsch, K. A. Shaw and N. C. MacDonald "Independent tuning of lineat and nonlinear stiffness coefficients," Journal of Microelectromechanical systems. Vol. 7. No. 2. June 1998, pp. 172-180.
14. J.J. Yao and N.C. MacDonald, "A micromachined, single-crystal, silicon, tunable resonator," J. micromech. Microeng. Vol. 6, 1996, pp. 257-264.
15. R. R. A. Syms, "Electrothermal frequency tuning of folded and coupled vibrating micromechanical resonators," Journal of Microelectromechanical systems. Vol. 7. No. 2. June 1998, pp. 164-171.
16. H. Kahn, M.A. Huff and A.H. Heuer, "Heating effects on the Young's modulus of films sputtered onto micromachined resonators," Mat. Res. Soc. Proc. Vol. 518, 1998, pp. 33-38.
17. R.I. Pratt, G.C. Johnson, R.T. Howe and J.C. Chang, "Mechanical structures for thin film characterization," in Dig. Transducers '91, Int. Conf. Solid-State Sensors and Actuators, pp. 205-208.
18. H. Tilmans, M. Elwenspoek and J. Fluitman, "Micro resonant force gauges," Sensors and Actuators, Vol. A30, 1992, pp. 35-53.
19. MCNC, MEMS Technology Applications Center, Research Triangle Park, NC, 27709.
20. Canvas. Graphics Software. Deneba Software, Inc., 7400 S.W. 87th Avenue, Miami, Florida 33173, USA, 1986.
21. Todd Remthma, "Active Frequency Tuning Via Localized Thermal Stressing Effects," Master Thesis, University of Michigan, 1999.
22. K.S. Teh, Liwei Lin and M. Chiao, "The creep behaviour of polysilicon microstructure," 10th Int. Conference on Solid State Sensors and Actuators, Transducer's 99, Technical Digest, Sendai, Japan, June 1999, pp. 508-511.

NANOELECTROMECHANICAL SYSTEMS

M. L. Roukes

Condensed Matter Physics 114-36, California Institute of Technology
Pasadena, CA 91125
e-mail: roukes@caltech.edu

ABSTRACT

Nanoelectromechanical systems, or *NEMS*, are MEMS scaled to submicron dimensions [1]. In this size regime, it is possible to attain extremely high fundamental frequencies while simultaneously preserving very high mechanical responsivity (small force constants). This powerful combination of attributes translates directly into high force sensitivity, operability at ultralow power, and the ability to induce usable nonlinearity with quite modest control forces. In this overview I shall provide an introduction to NEMS and will outline several of their exciting initial applications. However, a stiff entry fee exists at the threshold to this new domain: new engineering is *crucial* to realizing the full potential of NEMS. Certain mainstays in the methodology of MEMS will, simply, *not* scale usefully into the regime of NEMS. The most problematic of issues are the size of the devices compared to their embedding circuitry, their *extreme* surface-to-volume ratios, and their unconventional "characteristic range of operation". These give rise to some of the principal current challenges in developing NEMS. Most prominent among these are the need for: ultrasensitive, very high bandwidth displacement transducers; an unprecedented control of surface quality and adsorbates; novel modes of efficient actuation at the nanoscale, and precise, robust, and *routinely reproducible* new approaches to surface and bulk nanomachining. In what follows I shall attempt to survey each of these aspects in turn, but will conclude by describing some exciting prospects in this new field.

INTRODUCTION

NEMS have a host of intriguing attributes. They offer access to fundamental frequencies in the microwave range; *Q*'s, *i.e.* mechanical quality factors, in the tens of thousands (and quite possibly much higher); active masses in the femtogram range; force sensitivities at the attonewton level; mass sensitivity at the level of *individual* molecules, heat capacities far below a "yoctocalorie" [2] — this list goes on. These attributes spark the imagination, and a flood of ideas for new experiments and applications ensues. Of course, in time this initial enthusiasm

Figure 1. Schematic representation of a three-terminal electromechanical device.
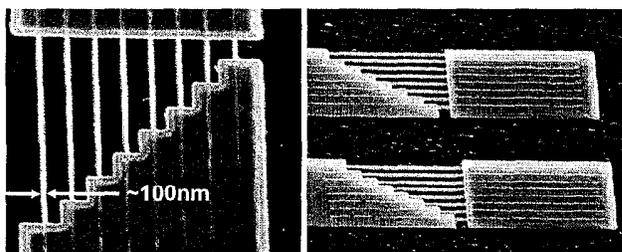
Figure 1. SiC NEMS. This first family of submicron doubly-clamped SiC beams exhibit fundamental resonant frequencies from 2 to 91 MHz. They were patterned at Caltech from 3C-SiC epilayers grown at Case Western Reserve University. After Yang, Ekinci, Zorman, Mehregany and Roukes, (Ref. 3).

gives way to deeper reflection, and a multitude of new questions and concerns emerge. Prominent among these is what will be our ultimate ability for optimally controlling and engineering these miniature systems. Clearly, the characteristic parameters of NEMS are extreme by all current measures. This paper is in three main sections: in the first, with this latter point in mind, I will attempt to convey a balanced introduction to the attributes of NEMS.

In the second section of the paper I offer my thoughts and projections regarding the most crucial aspects of NEMS engineering. How shall transducers and actuators be realized at the nanoscale? How shall surface properties be controlled? How can reproducible fabrication be attained?

The final section of the paper concerns ultimate limits. As we move forward in the development of NEMS it will become increasingly apparent what aspects will be susceptible to improvement through systematic engineering, and what hard, immutable limits are imposed by their fundamental physics. However, certain issues are already clear at the outset. I will try to summarize those that seem most apparent at this juncture.

MULTITERMINAL MECHANICAL DEVICES

The attributes of NEMS described in the next section make clear that we should be envisioning applications for electromechanical devices with response times and operating frequencies that are as fast as most of today's electron devices. Furthermore, *multiterminal* electromechanical devices are possible — *i.e.* two-, three-, four-ports, etc. — in which electromechanical transducers provide input stimuli (*i.e.* signal forces), and read out a mechanical response (*i.e.* output displacement). At additional *control* terminals, electrical signals— either quasi-static or time-varying— can be applied, and subsequently converted by the control transducers into quasi-static or time-varying forces to perturb the properties of the mechanical element in a controlled, useful manner. The generic picture of this scheme is shown in Figure 1.

There is an important point to be made regarding the "orthogonality" attainable between the input, output and (the
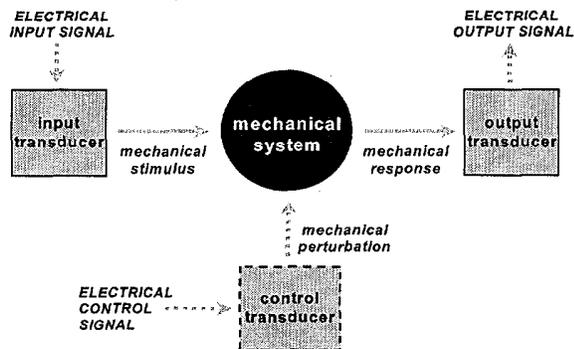
Table 1: Fundamental Frequency vs. Geometry for *SiC*, [Si], and (GaAs) Mechanical Resonators

Boundary Conditions	Resonator Dimensions ($L \times w \times t$, in μm)			
	$100 \times 3 \times 0.1$	$10 \times 0.2 \times 0.1$	$1 \times 0.05 \times 0.05$	$0.1 \times 0.01 \times 0.01$
Both Ends Clamped or Free	*120 KHz* [77] (42)	*12 MHz* [7.7] (4.2)	*590 MHz* [380] (205)	*12 GHz* [7.7] (4.2)
Both Ends Pinned	*53 KHz* [34] (18)	*5.3 MHz* [3.4] (1.8)	*260 MHz* [170] (92)	*5.3 GHz* [3.4] (1.8)
Cantilever	*19 KHz* [12] (6.5)	*1.9 MHz* [1.2] (0.65)	*93 MHz* [60] (32)	*1.9 GHz* [1.2] (0.65)

possibly multiple) control port(s). Different physical processes of electromechanical transduction available make it conceivable to achieve highly independent interaction between these ports, *i.e.* to have each of these strongly interacting with the mechanical element, but with only weak direct couplings to each other. For time-varying stimuli when frequency conversion is the goal, this orthogonality can be provided by tuned (narrowband) transducer response to select input and output signals from control (*e.g.* pump) signals. I shall discuss transduction mechanisms in a bit more detail below.

NEMS ATTRIBUTES

Frequency. Table 1 displays attainable frequencies for the *fundamental* flexural modes of thin beams, for dimensions spanning the domain from MEMS (leftmost entries) to deep within NEMS. The mode shapes, and hence the force constants and resulting frequencies, depend upon the way the beams are clamped; Table 1 lists the results for the simplest, representative boundary conditions along three separate rows. The last column represents dimensions currently attainable with advanced electron beam lithography. Of course, even smaller sizes than this will ultimately become feasible; clearly the ultimate limits are reached only at the molecular scale. Nanodevices in this ultimate limit will have resonant frequencies in the *THz* range, *i.e.* that characteristic of molecular vibrations.

Each entry is in three parts, corresponding to structures made from silicon carbide, silicon, and gallium arsenide. These materials are of particular interest to my group, and are among the "standards" within MEMS. They are materials available with extremely high purity, as monocrystalline layers in epitaxially grown heterostructures. This latter aspect yields dimensional control in the "vertical" (out of plane) dimension at the monolayer level. This is nicely compatible with the lateral dimensional precision of electron beam lithography that approaches the atomic scale. The numbers should be considered loosely as "typical"; they represent rough averages for the various commonly used crystallographic orientations.

It is particularly notable that for structures of the same dimensions, Si yields frequencies a factor of two, and SiC a factor of three, *higher* than that obtained with GaAs devices [3]. This increase reflects the increased phase velocity, $\sqrt{(E/\rho)}$, in the stiffer materials. E is Young's modulus, and ρ is the mass density.

One might ask at what size scale does continuum mechanics break down and corrections from atomistic behavior emerge. Molecular dynamics simulations for ideal structures appear to indicate that this becomes manifested only at the truly molecular scale, of order tens of lattice constants in cross section [4]. Hence, for most initial work in NEMS, it appears that continuum approximation will be adequate. However a very important caveat must be kept in mind. The frequencies in Table 1 are for structures with *zero* internal strain. In bi- or multi- layered structures (common for devices that include transducers) this may

actually be the exception rather than the rule [5]. Even for homogenous mechanical devices, *e.g.* those patterned from doped semiconductor materials, surface nonidealities in nanoscale devices may impart significant corrections to this simple picture.

Quality Factor. The Q's attained to date for NEMS in moderate vacuum, are in the range from 10^3 to 10^5. This greatly exceeds those typically available from electrical resonators. This small degree of internal dissipation ($D=1/Q$) impart to NEMS their low operating power levels and high attainable force sensitivity. For signal processing devices, high Q directly translates into low insertion loss [6].

One might expect nanomechanical resonators fabricated from ultrapure, single crystal semiconductor materials to have extremely high quality factors. But in our group similar Q's have been obtained for NEMS with resonant frequencies in the 20 MHz range from polycrystalline silicon. This trend holds at lower frequencies for very thin quasi-amorphous, low-strain silicon nitride devices. Figure 2 displays a rough trend that seems to be manifested in mechanical resonators in general – from those that are truly macroscopic in size, to those well within the domain of NEMS. As seen the maximum attainable Q's seems to scale downward with linear dimension.

It is important to note that large Q *does* imply a reduction of bandwidth, yet this need not be deleterious to performance for two reasons. First, feedback damping, which can be applied without introduction of significant additional noise, may be useful to increase bandwidth as desired. Second, for resonators operating at 1GHz, even with extremely high Q's of order 100,000, bandwidths
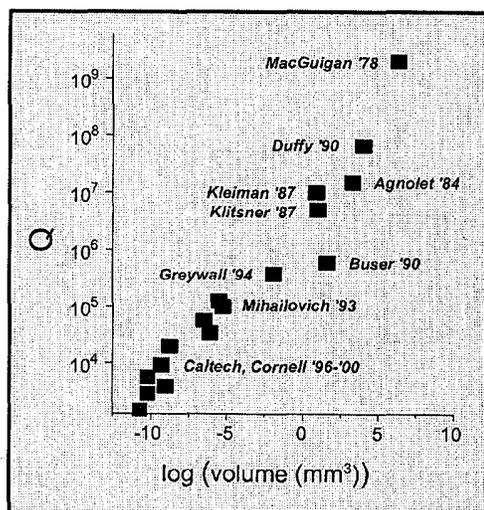
Figure 2. Q's in mechanical resonators varying in size from the macroscale to the nanoscale. The data follow a trend showing a decrease in Q that occurs, roughly, with linear dimension — i.e. with the increasing surface-to-volume ratio of small structures, (After D. Harrington, unpublished.)

Table 2: Representative operating power levels for NEMS.

f_0	Q	P_{min}	$10^6 \cdot P_{min}$
100 MHz	10,000	40 aW	40 pW
"	100,000	4 aW	4 pW
1 GHz	10,000	0.4 fW	0.4 nW
"	100,000	40 aW	40 pW

of order 10KHz are obtained, already sufficient for various narrow band applications.

Characteristic Operating Power Level. Applications of NEMS resonators will typically involve the use of a specific mode. A rough understanding of the minimum operating power levels using this mode can be obtained by dividing the thermal energy, $k_B T$, by the characteristic time scale for energy exchange between the mode, at frequency ω_0, and its surroundings (*i.e.* "the environment"). The time scale is set, roughly, by the "ring-up" or "ring-down" time of the resonator, $\tau = Q/\omega_0$. This simple estimate for the minimum power is then given by the ratio,

$$P_{min} \sim k_B T \omega_0 / Q. \tag{1}$$

It represents the signal power that must be fed to the system to drive it to an amplitude equal to the thermal fluctuations.

As displayed in Table 2, this minimum power is remarkably small for NEMS. For device dimensions accessible today via electron beam lithography, the characteristic level is of order *ten attowatts* (10^{-17}W). Even if we multiply this by a factor of a million, to achieve robust signal-to-noise ratios, and then further envision a million such devices acting in concert to realize some sort of future NEMS-based mechanical signal processing or computation system — the total system power levels still are only of order 1 µW. This is six orders of magnitude smaller than power dissipation in current systems of similar complexity based upon digital devices that work solely in the electrical domain.

Responsivity and Aspect Ratios. To reach *VHF*, *UHF* and, ultimately, microwave bands, the mechanical elements must be scaled downward from the current size domain of *MEMS*. It is possible to employ existing micron-scale *MEMS* technology to attain high frequencies, but this approach has serious disadvantages which preclude realization of the full scope of potentialities offered by NEMS technology. Attainment of high frequencies with *micron-scale* structures can only occur with extremely foreshortened aspect ratios, of order unity. (In the present context, aspect ratio corresponds to L/w or L/t.) Such geometries yield extremely high force constants— hence large excitation signals must be applied to yield appreciable mechanical response (*i.e.* deflection). This lack of response is deleterious for many device characteristics that are crucial for ultralow power electroacoustic signal processing. In effect these structures operate more as miniature bulk wave devices than as flexural or torsional resonators, which are more the focus of MEMS.

Large force constants adversely affect: *a)* the power level required for operation, *b)* the attainable dynamic range, *c)* the ability to tune the devices using "control" signals (applied mechanical forces), *d)* the attainment of maximum Q (through minimization of acoustic radiation to the support structures, *i.e.* clamping losses), and *e)* the excitation levels required to induce nonlinear response. All of these characteristics are optimized in large aspect ratio structures, *i.e.* structures with geometries

currently used in MEMS, but with all dimensions reduced to nanoscale dimensions. On the other hand, large force constants do lead to a scaling upward of the topmost portion of the dynamic range; such devices allow larger mechanical signal power levels. The cost, however, is that the thermomechanical fluctuation level is significantly suppressed and providing a noise-matched transducer is likely to become impossible.

Estimation of Available Dynamic Range. From Table 2 it is clear that NEMS have the potential to provide new types of ultralow power electromechanical signal processing and computation. However realizing them will be non-trivial; this potential can only be harnessed by employing them optimally, *i.e.* within their characteristic operating range. To utilize the full potential of NEMS, displacement transduction schemes are required that can provide resolution at the level of the thermomechanical fluctuations. The bottom segment of a mechanical system's available dynamic range will be forfeited unless such optimal, *i.e.* noise matched, transducers are employed to read out its motion. Specifically, the transducer's internally generated noise output, referred back to the input (RTI), must have magnitude comparable or smaller than the thermomechanical fluctuations. This will be discussed in more detail below. The challenge is that the r.m.s. amplitude of vibration for a mechanical device (operating within its linear range) scales downward in direct proportion to its size.

These considerations indicate two crucial areas for NEMS engineering: (a) Development of ultrasensitive transducers that are capable of both enhanced displacement resolution and increasingly higher frequencies as device sizes are progressively scaled downward, and (b) Development of techniques and system architectures tailored to operate within the characteristic dynamic range NEMS, *i.e.* in femto- to picowatt regime.

To estimate the characteristic dynamic range for a linear NEMS device requires knowledge of its displacement noise floor, and of the vibration amplitude at the onset of nonlinearity.

The former is set by thermally driven displacement fluctuations, *i.e* the thermomechanical noise, the mechanical analog of Johnson/Nyquist noise. For a damped simple harmonic oscillator, it is characterized by the expression

$$S_x(\omega) = \frac{4 k_B T}{\kappa \omega_0 Q} \left[\frac{\omega_0^4}{\left(\omega^2 - \omega_0^2\right)^2 + \left(\omega \omega_0 / Q\right)^2} \right]. \tag{2}$$

$S_x(\omega)$ is the spectral density of displacement noise (with units m^2/Hz), and κ is the force constant. Spectral densities are tabulated for the four devices of Table 3. To put these numbers in perspective, state-of-the-art transduction schemes at low frequencies currently provide displacement sensitivities typically of order 10^{-13} to 10^{-15} m/$\sqrt{\text{Hz}}$. I shall discuss this in a bit more detail below.

This establishes the r.m.s. displacement amplitude at the displacement noise floor, *i.e.* the *bottom* end of the dynamic range. To establish where the *top* end resides requires specific knowledge of the device geometry. First, one must establish a consistent criterion to define $\langle x_N(\omega) \rangle$, the r.m.s. amplitude at the onset of nonlinearity. The one I shall employ is the following: in a power series expansion for elastic potential energy, $\langle x_N(\omega) \rangle$ is the level at which the largest term beyond the quadratic (Hooke's law), grows to become 10% the size of the second order term. For both

uniform and point loading of a doubly clamped beam, this condition translates into the relation $\langle x_N(\omega)\rangle \sim 0.53\,t$, which depends only solely upon the beam thickness in the direction of vibration [7]. Given this definition, the onset of nonlinearity is displayed in Table 3 for a representative family of devices.

With these limits, the dynamic range is then

$$DR = 10\log\left[\langle x_N(\omega)\rangle^2 / \left(S_x(\omega)\,\Delta\omega\right)\right], \qquad (3)$$

where $\Delta\omega$ is the measurement bandwidth. Usually the entire resonant response is used; for this case $\Delta\omega \sim \omega_0/Q$. The DR displayed in Table 3 assumes a bandwidth equal to the linewidth.

It is relevant to recast this expression in terms of signal forces, i.e. the "input", since the displacements are, in essence, the "output" of the resonant mechanical system. For a resonant, linear mechanical system, the force-to-position transduction can be expressed as $x(\omega) = F(\omega)\,A(\omega)/\kappa$. Here $A(\omega)$ is the normalized amplitude response function that appeared in Eq. 2,

$$A(\omega) = \frac{\omega_0^2}{\sqrt{\left(\omega^2 - \omega_0^2\right)^2 + \left(\omega\omega_0/Q\right)^2}}. \qquad (4)$$

In direct analogy with $\langle x_N(\omega)\rangle$, I define the characteristic r.m.s. input force that drives the system to the threshold of nonlinearity as $\langle F_N(\omega)\rangle$. This leads to an expression for the DR in terms of the signal force,

$$DR = 10\log\left[\frac{\langle F_N(\omega)\rangle^2}{S_F\,\Delta\omega}\right] \qquad (5)$$

Here $S_F = 4k_B T \kappa/(\omega_0 Q)$ is the force spectral density, discussed in more detail below.

Active Mass. Only a fraction of the total resonator's mass is involved in its motion. For beams or cantilevers, multiplying the total mass by the integral of a normalized function describing the mode shape yields a measure of the active mass. For a doubly-clamped beam operating in fundamental mode, this turns out to be

about half (actually, ~ 0.523) of the total mass of the beam.

Mass Sensitivity. For NEMS resonators, their minuscule active masses compounded with their high Q yields an extreme sensitivity to added mass. We can make a simple estimate of the added mass required to shift the resonant frequency by its halfwidth, i.e. $\omega_0 \rightarrow (\omega_0 + \partial\omega_0)$ where $\partial\omega_0 \sim \omega_0/(2Q)$. This is simply given by multiplying $\partial\omega_0$ by the inverse of what one can call the frequency "responsivity" to added mass,

$$\partial M|_{min} \sim \left(\partial\omega_0/\partial m\right)^{-1}\partial\omega_0. \qquad (6)$$

Approximating the resonator mode as a simple harmonic oscillator one finds $\partial M|_{min} \sim (2m/\omega_0)(\omega_0/2Q) \sim m/Q$. Here m is the "active mass" of the resonator. As shown in Figure 3, this leads to remarkably high mass sensitivities. For the smallest beams envisioned in the near term, mass resolution at the level of a few hundred Daltons ($1D = 1.7 \times 10^{-24}$ g) is obtained. Clearly it is possible to resolve shifts of the centroid of the line shape to a small fraction of the linewidth. Hence, with these smallest NEMS, it seems completely feasible to resolve frequency shifts for added masses as small as 1D.

This mass sensitivity is a double-edged sword; on the one hand it offers unprecedented sensitivity for mass sensing, but this sensitivity can also make device reproducibility challenging, even elusive. We have found that it places quite stringent requirements on the cleanliness of fabrication techniques with regard to processing residues, etc.

It is also clear that adsorbates on NEMS will play an important role in their properties. We are embarking upon a program to investigate NEMS placed in UHV at room temperature, to allow studies that commence with careful surface preparation cycles. Recent work at Stanford has shown improvement of resonator properties after a high temperature, annealing step *in-vacuo* [8].

Phase Noise. A corollary of extreme mass sensitivity of NEMS is the expectation of large phase noise for NEMS resonators. This expectation is predicated upon the fact that at finite pressures and temperatures, there will be thermodynamically driven fluctuations in the total number of adsorbed species on a device. This will give rise to a fluctuating total mass, and hence a

Table 3: Important attributes for a family of doubly-clamped Si beams. The effective force constant, k_{eff}, is defined for point loading at the beam's center. S_X and S_F are the spectral densities for displacement and force noise from thermomechanical fluctuations; a noisy readout will degrade these ideal values. Nonlinear onset has been characterized using the criterion described in the text. The linear dynamic range (DR) is defined as the ratio of this onset to $\sqrt{(S_X\,\omega_0/Q)}$. Noise-matched cryogenic operation at 4K adds ~19dB to the linear DR values shown. Active mass is ~half the beam's mass for the fundamental mode. Mass sensitivity, measured in Daltons ($1D=1.7\times10^{-24}$ g) is for a half linewidth shift.

Dimensions $L \times w \times t$, (all in μm)	Resonator Attributes assuming $Q = 10,000$ (100,000)							
	Frequency $\omega_0/2\pi$	κ_{eff} (N/m)	$S_x^{1/2}(\omega_0)$ at 300K (m/√Hz)	Nonlinear Onset, $\langle x_N\rangle$,(m)	Linear DR (dB)	$S_F^{1/2}(\omega_0)$ at 300K (N/√Hz)	Active Mass	Approx. Mass Sens., (D)
$100 \times 3 \times 0.1$	77 KHz	.007	2×10^{-10} (7×10^{-10})	5×10^{-7}	51 (51)	3×10^{-16} (5×10^{-17})	40 pg	10^9 (10^8)
$10 \times 0.2 \times 0.1$	7.7 MHz	0.5	3×10^{-12} (8×10^{-12})	5×10^{-7}	68 (68)	1×10^{-16} (4×10^{-17})	0.3 pg	10^7 (10^6)
$1 \times 0.05 \times 0.05$	380 MHz	16	7×10^{-14} (2×10^{-13})	3×10^{-8}	59 (59)	1×10^{-16} (3×10^{-17})	3 fg	10^5 (10^4)
$0.1 \times 0.01 \times 0.01$	7.7 GHz	25	4×10^{-14} (1×10^{-13})	5×10^{-9}	35 (35)	3×10^{-17} (9×10^{-18})	10 ag	10^3 (10^2)

fluctuating resonant frequency. Knowledge of the spectral density for this process is necessary to establish its importance as a phase noise process. Clearly important are the ambient pressure, the exposed surface area of the beam (the normalized mode shape will play a role here), the sticking probability for the gaseous species, and the average dwell time of the species once adsorbed upon the beam. Some preliminary efforts to understand this process have been offered [9], based on simple constitutive models. These must be re-examined for the smallest size domain of NEMS, where such relations quickly become inapplicable.

Force Sensitivity. The spectral density of force noise, driven by thermodynamic fluctuations, is white. For a simple harmonic oscillator it can be expressed through the fluctuation-dissipation theorem as

$$S_F(\omega) = \frac{4\kappa k_B T}{\omega_0 Q}. \tag{7}$$

This represents the ideal case; a mechanical system can act as a resonant force detector with this noise floor only if its readout is noiseless. Representative values for the square root of this quantity, commonly called the "force noise" (with units N/√Hz) are displayed in Table 3.

In general, it will not be possible to read out the response of the mechanical system with arbitrary precision. To estimate the force sensitivity attained with a readout chain (displacement transducer + subsequent amplifiers) having a finite noise contribution we must make use of the mechanical resonator's amplitude response function.

Using this, we can write the effective force sensitivity for the total system, mechanical resonator plus readout, as

$$S_F^{eff}(\omega) = S_F(\omega) + S_x^{eff}(\omega)\left[\kappa/A(\omega)\right]^2 \tag{8}$$

Here $S_F(\omega)$ and $S_F^{eff}(\omega)$ are the thermodynamic and *effective* (coupled) force spectral densities respectively. The noise floor of the displacement sensing system is given by the square root of its effective spectral density (RTI), $S_x^{eff}(\omega)$.

As seen, the force spectral density is highest when the resonant response function, which is located in the denominator of the second term, peaks at ω_0 and thereby suppresses the noise of the readout chain. The magnitude of the force constant and the effective displacement spectral density of the readout chain govern the efficacy of such suppression. This expression also shows that extremely stiff resonators, such as those obtained in structures with foreshortened aspect ratios, can easily become impossible to read out effectively. In this case, as previously mentioned, the bottom region of the dynamic range will be forfeited.

Attainable Nonlinearity and Tunability. The onset of *nonlinearity* – crucial for frequency conversion, demodulation, and parametric processes – occurs for smaller applied force (hence lower input power) in large aspect ratio structures.

Intimately related to this is the *tunability of frequency and non-linearity*, important for frequency-agile applications. With smaller force constants in large aspect ratio structures, the range of tunability is greater, and can be achieved with lower power. It is bounded between the threshold of nonlinearity, and the force where internal stresses exceed the yield modulus. The rate at which the properties of mechanical element may be perturbed (*i.e.* altered in a useful way) is limited only by the response time

($\tau = Q/\omega_0$) of the element itself, and by the bandwidth of the control transducer. In principle, shifts on sub-nanosecond time scales should be achievable.

PRINCIPAL NEMS ENGINEERING CHALLENGES

Pursuit of Ultrahigh Q. Central to attaining the ultimate limits of *VHF/UHF* NEMS performance is the pursuit of ultrahigh Q. This overarching theme underlies all research in NEMS, with exception of nonresonant and fluidic applications. Dissipation (~1/Q) within a resonant mechanical element limits its sensitivity to externally applied forces (signals), and sets the level of fluctuations that degrade its spectral purity (i.e. broaden its natural linewidth), and determine the minimum intrinsic power levels at which the device must operate. Hence ultrahigh Q is extremely desirable for low phase noise oscillators and highly selective filters; it also makes external *tuning of dissipation* easier.

Extrinsic and *intrinsic* mechanisms are operative to limit Q in real devices. Many extrinsic mechanisms can be suppressed by careful engineering; these include air damping, "clamping losses" at supports, and "coupling losses" mediated through the transducers. Some of the intrinsic mechanisms may be suppressed by careful choice of materials, processes, and handling. These include anelastic losses involving: *a)* defects in the bulk, *b)* the interfaces, *c)* fabrication-induced surface damage, and *d)* adsorbates on the surfaces. Certain anelastic loss mechanisms are, however, fundamental; these impose the ultimate upper bounds to attainable Q's; such processes include thermoelastic damping arising from anharmonic coupling between mechanical modes and the phonon reservoir [10].

Surfaces. NEMS devices patterned from single crystal, ultrapure heterostructures can contain very few (even *zero*) crystallographic defects and impurities. Hence, the initial hope was that within small enough structures bulk acoustic energy loss processes should be suppressed and ultrahigh Q-factors thereby attained. In this size regime one might even expect bulk dissipation to become sample-specific – *i.e.* dependent upon the precise configuration and number of defects present.

Figure 2 illustrates what happens as we scale downward. The dependence on dimension, which is inversely proportional to surface to volume ratio, clearly seems to implicate the role of surfaces.

It is worthwhile to ponder this size regime in a bit more detail. The smallest-scale entries in Tables 1 and 3 involve a beam with 10nm cross-section and 100nm length. Its corresponding volume is approximately 10^{-23} m^{-3}; hence, given the 0.543 nm lattice constant of Si and its 8 atoms per unit cell, this resonant structure contains only about 5×10^5 atoms. The surface area of this beam is approximately 4×10^{-15} m^{-2}; hence, with two Si atoms per unit cell face, there are roughly 3×10^4 atoms at the surface of this mechanical structure. Hence, more than *ten percent* of the constituents of this structure are surface, or near surface atoms. It is clear that they will play a central role.

Transducers. Electrostatic transduction, the staple of MEMS, does not scale well into the domain of NEMS. Electrode capacitances of order 10^{-18} F and smaller are to be expected for electrodes at the nanoscale, hence parasitic capacitance will dominate the dynamical capacitance of interest. In effect, as device size shrinks and the frequency of operation increases, the motional modulation of the impedance becomes progressively

371

smaller while the static parasitic and embedding impedances continue to grow.

Optical detection, including both simple beam deflection schemes and more sophisticated optical and fiber-optic interferometry also do not scale well into the domain of NEMS. Contemporary scanned probe microscopies, such as atomic force microscopy (AFM) and magnetic force microscopy (MFM), make extensive use of optical techniques to measure minute displacements of MEMS-fabricated cantilevers having lateral dimensions in the few-to-hundred micron range. Sensitivities as small as a few $\times$ 10^{-4} $Å/\sqrt{Hz}$ are attainable with these methods [11]. Unfortunately, this conventional approach fails for objects with cross sections much smaller than the diameter of an optical fiber, such as NEMS. (Single mode fibers for 833nm have core diameters of order a few μm.) The lateral spot size of radiation emanating from the end of even the narrowest single-mode fiber, or the diffraction limited spot from high numerical aperture optics, are both still at the μm scale. This, of course, can be much larger than an entire NEMS device. Accordingly, conventional optical approaches appear to hold little promise for high-efficiency displacement transduction with the smallest of NEMS devices.

Reproducible Nanofabrication. The mass sensitivity displayed in Table 3 makes clear that fabrication reproducibility is key for NEMS. Device trimming is ubiquitous in quartz frequency control technology. It is clear that such techniques will also be required for NEMS. However, optimal fabrication techniques will reign in the device-to-device spreads arising from mass variations.

POTENTIAL NEMS ENGINEERING SOLUTIONS

NEMS Surfaces. Given what is known from electronic and photonic device physics regarding oxidation and reconstruction of the Si surface, it seems clear that the mechanical properties of the smallest NEMS devices deviate greatly from those in bulk. It may prove quite difficult to achieve ultrahigh Q with such extreme surface-to-volume ratios, if only conventional patterning approaches are utilized. Surface passivation will undoubtedly become imperative for nanometer scale MEMS devices.

One might project that structures such as nanotubes may well represent the ideal for NEMS, given their perfectly terminated surfaces. So far, however, the existing technology of manipulating, anchoring, and measuring the mechanical properties is still quite primitive. Hence there is currently insufficient information available even to permit a crude extrapolation of the Q's that might ultimately be attainable at high frequencies with nanotube-based NEMS.

Novel Displacement Transducers for NEMS.

For the *electrostatic* case it appears there may be one apparent escape from the spiral of deceasing motional impedance accompanied by increasing parasitics that occur when device size is scaled downward. The solution would appear to be simply to *eliminate* the large embedding and parasitic impedances. This could be achieved by locating a subsequent amplification stage (which would, in effect, acting as an impedance transformer/line driver) directly at the capacitive transducer. In effect, this would make the NEMS electrode serve dual purposes, as both motion sensor for the NEMS, and as gate electrode of an FET readout.

Although conventional optics quickly become of limited used for sizes below the diffraction limit, possibilities do exist for *integrated* and *near-field optical* displacement sensors. Noteworthy is the fact that, for example, in the GaAs/AlGaAs

materials system, a single-mode waveguide has cross sectional dimensions well below one micron. With such on-chip optical waveguide technology, evanescent wave radiation, such as that emanating from the end of a waveguide beyond cut-off, can provide displacement sensing with large bandwidth and extremely high spatial resolution [12]. The rather high optical power levels involved, however, may preclude the most sensitive cryogenic applications.

A variety of alternate techniques also appear to hold promise for NEMS. My group has made extensive use of *magnetomotive (i.e. magnetogalvanic)* detection [13]. It is based upon the presence of a static field, either uniform or spatially inhomogeneous, through which a conductor (actually a conducting *loop*) is moved. The time-varying flux generates an induced e.m.f. in the loop. This same principle can apply to a fixed, motionless loop that is placed near a moving nanomagnet; in this way a time varying flux can also be coupled to the loop, and an e.m.f. thereby generated.

I make a distinction between magnetomotive and *direct magnetic* detection. For the latter case we consider additional means, beyond simple conducting loops, by which a time-varying magnetic field arising from a moving nanomagnet (attached to the mechanical resonator) may be detected. We have developed a high sensitivity detection scheme in which a nanomagnet, affixed to a torsional resonator and thereby moving in concert with it, couples a time-varying fringe field to a low-electron-density, high Hall coefficient field sensor [14]. Magnetic sensing can be realized using SQUIDs (superconducting quantum interference devices), and flux-gate magnetometers, however both of these suffer from limited bandwidth and cannot provide the frequency response to access the upper range of NEMS via direct sensing. Nonetheless, they may ultimately prove quite useful as i.f. (intermediate frequency) or output amplifiers in mechanically-realized heterodyne or homodyne frequency conversions schemes, respectively.

(Inverse) magnetostrictive transduction may also prove useful with using these sensitive post-detectors; in this case magnetic materials, whose centers of mass are static, respond to strain by changing their internal magnetization. These materials may be employed to transduce time-varying strain (e.g. near a point of support in a mechanical resonator) into a time-varying local field outside the material.

Piezoelectric detection can be realized in two principal forms. In the first, local time-varying strain fields within a piezoelectric medium (e.g. at points of high strain within a mechanical resonator) create corresponding time-varying polarization fields. These can be detected by placing the low density electron gas of a field effect transistor channel where the time-varying electric polarization is largest [15]. We have been working to optimize this technique using suspended, high mobility HEMT structures [16]. *Piezoresistive detection* is closely related to this scheme.

The second possible form of piezoelectric detection is *nonlocal*, and based upon the coupling of local mechanical modes of a resonator to *surface acoustic waves*, which are subsequently detected piezoelectrically, downstream by means of interdigitated transducers.

Electron tunneling is a technique that scales well, even down to extremely small dimensions. However, because the impedance of a tunnel junction is quite high, the bandwidth of such detectors is minimal (in the presence of unavoidable, uncontrolled parasitic capacitance). Despite this problem, high frequency detection via electron tunneling may prove extremely useful when used as the i.f. readout for nonlinear mechanical downconversion scheme.

Finally, *thermal transduction* is possible since irreversible heat flow is always associated with flexure of a beam. With sensitive local thermal transducers it may be possible to detect the time-varying temperature field associated with this process. This would seem quite difficult to realize in nanometer scale structures.

PROMISING APPLICATIONS OF NEMS

Ultimately, NEMS will undoubtedly be employed in a broad range of applications. Even at this early stage of development it seems clear that one of the principal areas will be signal processing in the VHF, UHF and microwave bands [17]. Among my own group's explorations with NEMS have been their uses for metrology and fundamental science: i.e. for mechanical charge detection [18], and for thermal transport studies at the nanoscale [19, 20]. We are also pursuing a number of NEMS applications that we believe will hold immense technological promise. In the remainder of this section, I shall briefly describe two examples that go by the acronyms MRFM and BioNEMS.

Mechanically-detected Magnetic Resonance Imaging. It has been more than fifty years since nuclear magnetic resonance (NMR) was first observed [21], yet still takes about $10^{14} - 10^{16}$ nuclei to generate a measurable signal (via conventional inductive detection techniques). This means that state-of-the-art magnetic resonance imaging (MRI) in research laboratories attains, *at best*, minimum resolution (voxel size) of order 1µm. More typically, the resolution yielded by standard clinical MRI using commercially available instrumentation is much poorer, only of order 1mm. While scanned probe techniques such as AFM now routinely give *atomic* resolution for studies of surfaces, attaining MRI with resolution at the atomic scale would appear to be only a distant dream — the conventional approach is still ~14 orders of magnitude away from single spin detection. In 1991, however, Sidles proposed that nuclear magnetic resonance (NMR) spectrometry with sensitivity at the level of a *single proton* might be achievable through mechanical detection [22]. Achieving this degree of sensitivity would constitute a truly revolutionary advance; it would permit three-dimensional atomic-scale imaging, with chemical specificity. It is hard to overemphasize the impact this would have upon many fields.

Recent experimental work first at IBM Almaden, then in our laboratory, and subsequently several others worldwide, has demonstrated that the mechanical (*i.e.*, force-) detection principle for magnetic resonance is sound. In fact in 1994, Rugar and co-workers at IBM detected a signal from ~10^{13} protons [23]. Even in this *first* experimental demonstration of the mechanical detection of NMR, the sensitivity attained exceeded the state-of-the-art using conventional inductive methods. In 1995, in our close collaboration with the group of P. Chris Hammel at Los Alamos National Laboratory, we independently confirmed Sidles' concept [24]. Late last year, our collaborative work culminated in NMR detection on what is, effectively, equivalent to of order ~10^7–10^9 fully polarized Co nuclei [25].

Mechanically-detected MRI, now commonly called *magnetic resonance force microscopy* (MRFM) is now significantly more sensitive than conventional MRI. Comparing the time that elapsed since the first detection of NMR signals and the present, with the elapsed time since the first MRFM signals were obtained, the mechanical detection technique has clearly provided staggering advances. Extrapolating this rate of development, *vis-à-vis* the tasks ahead and the tools in place, it seems clear that significant *additional* gains are on the horizon for MRFM. Our recent advances make us optimistic that MRI with atomic resolution will be attained in several years.

There are four principal components to an MRFM instrument (Fig. 1). An "antenna structure", *e.g.* a coil or microstripline, generates a roughly uniform r.f. *excitation field* (frequency ω_0) that, in concert with the static magnetic field emanating from a miniature *gradient magnet*, induces *local* spin resonance within the sample. (A static homogeneous magnetic field may also be applied to enhance the sample's spin polarization.) The gradient magnet is affixed to a mechanical resonator, *e.g.* a cantilever. The interaction of the resonant spins with the gradient magnet results in a time-varying, back-action force upon this cantilever. The mechanical system, read out by a high resolution displacement sensor, such as a fiber-optic interferometer, constitutes a resonant force sensor. In effect, it detects the extremely weak forces exerted by the resonant spins upon the compliant mechanical system. Microscopy is realized by scanning the resonant volume (i.e., scanning the cantilever with its attached gradient magnet) over the sample, and then correlating the resonant mechanical response with position. This response is then deconvolved to obtain spatial imaging of spin density.

The inhomogeneous magnetic field from the nanomagnet plays two roles. First, it causes spins only within a very small region of the sample to become excited — it defines a surface on which the Larmor condition, $\omega_0 = \gamma |\vec{B}(\vec{r})|$, is locally satisfied. This "resonant volume" comprises spins of gyromagnetic ratio, γ, approximately located between surfaces determined by the linewidth, $\delta\omega$ (Fig. 1). Roughly speaking, this volume is bounded by surfaces defined by the relation $|\vec{B}(\vec{r})| = (\omega_0 \pm \delta\omega) / \gamma$. With the extremely large gradient field emanating from a nanomagnet this shell thickness can be reduced to atomic dimensions. Second, the nanomagnet acts as a spin-to-force transducer — the magnetic interaction between the precessing spins and the nanomagnet's inhomogeneous field results in a time-varying "back-action" force $\vec{F}(t) = V[\vec{M}(t) \cdot \vec{\nabla}]\vec{B}$ imposed upon the mechanical resonator.

To date, the mechanical device acting as the resonant force detector has typically operated at a frequency (~10 kHz) far below the spin precession frequency (~100 MHz and higher). To couple these disparate systems, it is necessary to orchestrate a low-frequency modulation (at the cantilever frequency, ω_0) of the rapidly time-varying magnetization. This can impose serious constraints upon the process. Also, the relatively slow response time of a low frequency mechanical system limits the rate of image acquisition; without mechanical feedback one must integrate on a time scale $\tau_{mech} \sim Q/\omega_0$ per voxel. Furthermore, the spin resonance must remain coherent on a time scale commensurate
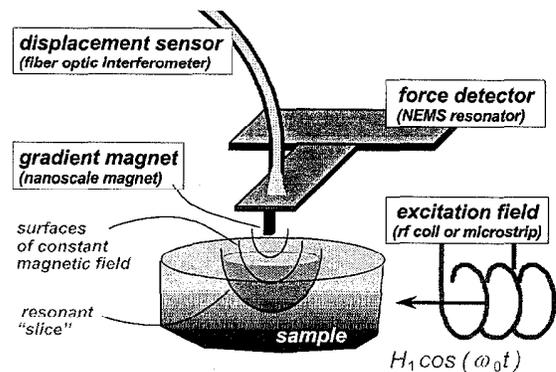
Figure 3. Schematic of the force detection approach to MRI.

373

with this long mechanical "ring-up" time to achieve a transfer of energy. These constraints associated with a low-frequency mechanical force detector can reduce the applicability of the technique solely to rather select compounds or materials.

NEMS will provide key advances for MRFM. Our current work is focused upon utilization of a nanometer-scale, VHF mechanical resonator to allow coupling directly at the spin precession frequency. This has the advantage of being able acquire data on a much shorter time scale, and can also help by moving the measurement away from the frequency regime of environmental noise and 1/f fluctuations. Equally important is that this new approach may ultimately provide access to the whole range of techniques from conventional *pulsed* magnetic resonance.

The current first-generation instruments used, worldwide, by the small group of current MRFM practitioners are all, in essence, *completely hand assembled* research prototypes. Recent work from Stanford/IBM collaboration, in which techniques from MEMS have been utilized, is a prominent exception

A crucial requirement for advances in MRFM is ultrasensitive displacement detection. Off-chip, fiber-optic interferometers used in previous and current generations of MRFM are too large and spatially insensitive for this task with NEMS resonant force detectors. These issues are the same as have been discussed above in a broader context.

The field of MRFM is still very much in its infancy. Sustained effort will be required to take it from the realm of being an interesting scientific demonstration to that of being a useful research technique for high resolution MRI. With its potential for atomic resolution, such efforts would seem certain to be of great potential importance, especially for biochemical applications.

BioNEMS. Mechanical devices have recently yielded impressive demonstrations of single molecule sensitivity for biochemical research. Exciting advances have been made using the separate techniques of atomic force microscopy (AFM) and "optical tweezers".

AFM, which was first developed in 1988 [26], has proven increasingly useful for probing extremely weak forces, including chemical forces involving *individual* molecules [27, 28]. The growing literature of chemical force microscopy (CFM) has shown that it enables investigations of the binding force of interactions ranging from single hydrogen bonds, to single receptor-ligand interactions, to single covalent bonds. One of the earliest papers in this field demonstrated sufficient sensitivity to detect the force required to break an individual hydrogen bond, estimated to be of order 10 *pN* [29]. As an example of recent work, delicate forces involved in the unfolding of a protein have been observed [30].

On the other hand, "optical tweezers" have also recently led to some quite spectacular measurements of weak forces in biological systems. In this technique an optical beam, focused to the diffraction limit, is employed. This yields optical gradient forces are much too spatially extended to permit direct manipulation of *individual* biomolecules under study. So, instead, functionalized dielectric beads, typically having diameter ~1um, are attached to the analytes to provide a dielectric "handle". In this way direct measurements of *pN*-scale biological forces from molecular "motors" have been obtained [31]. In another recent study, internal dynamics of DNA, yielding forces in the *fN* range, have been observed via two-point correlation techniques [32].

On a completely separate (and non-mechanical) front, developments in microfabricated biochemical microarray technologies have provided significant recent advances in analyzing protein receptors and their ligands, as well as in
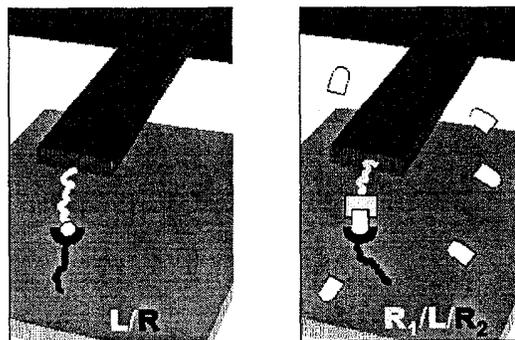
Figure 4. Two realizations of BioNEMS for chemical force sensing. A nanoscale cantilever with biofunctionalized surfaces senses forces between bound ligand/receptor (L/R) pairs (left), and the presence of a ligand in solution (right). After **[35]** .

analyzing gene expression profiles. Microarrays of a few thousand targets have become a major technique; they are now available commercially and widely used by the drug discovery industry. Although these methods are becoming increasingly widespread, the large size of the reader instrumentation and the intrinsic limitations of the fluorescence analysis employed make the technique ill suited for applications in which both portability and robust performance are required. Furthermore, this is a single-use methodology; hence the devices cannot easily accommodate applications that require continuous monitoring. Finally, the devices rely on significant volumes of analyte (by synthetic biochemistry or cellular standards), making them ill-suited to the recent advances in drug discovery provided by combinatorial chemistry or for exploring gene expression at the cellular scale.

All of these impressive advances pose an interesting challenge: *can one realize a nanoscale assay for a single cell?* The assay must be capable of responding to the ~10-1000 copies of a given molecular species in the volume of a single cell (~1pL), and must have the temporal resolution to follow the binding kinetics of single biomolecules (on the ~1μs scale).

Biochips involving nanoscale mechanical systems (BioNEMS) appear capable of providing important advances. Figure 4 is a "cartoon" depicting several possibilities. The recent advances in NEMS described earlier in this paper, and recent developments in the theory of fluid-loaded nanomechanical systems [33], suggest that the challenge posed above can be met.

The *pN* range of biological forces is well within the detection capability of CFM. The advances and attributes described above have led to proposals for "force-detected" chemical assays, as well as a concrete realization for a complete analysis system [34]. However, the MEMS-scale devices at the heart of these approaches cannot provide μS-scale temporal response in solution that is essential for following the binding and unbinding of biological ligands and their receptors. In addition, the surface areas involved, despite being at the micron scale, are still quite large for single molecule assays.

The BioNEMS approach is uniquely suited to these tasks. They offer a number of crucial advantages unattainable from micron scale mechanical devices (such as those conventionally used in CFM) or from the technology of "optical tweezers". BioNEMS are scaleable; can interact with highly controlled, extremely reduced population of analytes – *i.e.* down to the level of a single adsorbate per active device; their small active mass promises sensitivity that is crucial for single-molecule work; and their fluid-loaded response can easily provide response times down to the μs-scale. BioNEMS force measurements offer an attractive

alternative to the fluorescent labeling and optical detection that are principal protocols for biochemical microarray assays. Given the size of NEMS, force measurements can be performed extremely locally, providing the ability to provide assays on samples smaller than the diffraction limit. Furthermore, when pushed to single molecule sensitivity, optical detection can involve light intensity levels sufficient to induce photochemical damage to the analytes. By contrast, mechanical force measurements involve interactions solely at very low energy. We are currently pursuing the interesting possibilities offered by BioNEMS at Caltech [35].

THE QUANTUM LIMIT FOR MECHANICAL DEVICES

The ultimate limit for nanomechanical devices is operation at, or even beyond, the quantum limit. In this regime the individual mechanical quanta, $\hbar\omega_0$, are of order, or greater than the thermal energy, $k_B T$, and one must appeal to the quantum theory of measurement to understand and optimize force and displacement measurements. Below, I shall try to elucidate what I see as three phases, or "doors", which we must pass through along the path toward realizing mechanical systems at the quantum limit.

One of the most intriguing aspects of today's nanomechanical devices is that they *already* verge on the quantum limit. The fundamental flexural mode for the even the first VHF resonators made in our group back in 1994 operate at sufficiently high frequency that thermal excitation at a temperature of 100mK involves, on average, only about 20 vibrational quanta.

Such temperatures are readily attained using a ^{3}He/^{4}He dilution refrigerator. The question that immediately comes to mind is *"can quantized amplitude jumps be observed in such a device?"* If so, one could observe discrete transitions as the system exchanges quanta with the outside world and jumps, *e.g.* downward from 20 to 19 quanta, or upward from 20 to 21 quanta, etc. At this point, the answer to the question seems to be affirmative, but there are two important criteria that must be met. First, the device must be in a discrete quantum ("number" or "Fock") state. To accomplish this one must insure that the transducers that couple to the resonator measure only the *mean-squared* position, without coupling linearly to the resonator's position itself. Such transducers were discussed in a pioneering paper on mechanical devices at the quantum limit [36]. Second, the transducer used to measure mean-squared position clearly must have the sensitivity to resolve an *individual* quantum jump. The need for ultrasensitive nanoscale transducers again appears; this criterion is the crucial key needed to unlock this first door to the quantum domain. Simple estimates lead one to the conclusion that sensitivities of order 10^{-30} m^2/Hz (or better!) are necessary. With magnetomotive detection we are currently only about a factor of ten away from this.

Beyond the second "door" one truly enters the quantum limit for mechanical systems. Here frequencies are sufficiently high that, at the temperature of operation, thermal occupation factors fall below one for the lowest-lying modes. In this situation, device noise is governed by zero-point mechanical fluctuations. Accessing this domain requires high frequencies and very low temperatures; for example, a 1 GHz mechanical resonator enters this regime only when cooled below about 50 mK. Although we can indeed make nanoscale devices today with resonant frequencies within this realm, cooling them to this regime (and truly insuring that they *really are* cold), and then measuring their response with sensitivity at the single quantum level pose formidable challenges. The payoff for efforts in these directions, however, will be truly significant. Force and displacement detection at the quantum limit may open new horizons in measurement sciences at the molecular level, new device possibilities for phase coherent measurements and quantum computation, and intriguing experiments and control of thermal transport involving the exchange of *individual* phonons between nanomechanical systems, or between a nanosystem and its environment.

Once we have passed through this second door, the division between quantum optics and solid state physics becomes increasingly blurred. Many of the same physical principles governing the manipulation of light at the level of individual photons will come into play both for the mechanical and thermal properties of nanoscale systems.

In fact, this will become most evident upon passing through a third figurative "door" that will takes us beyond the standard quantum limit for force and displacement measurements. As first pointed out by some time ago [37] the principles of "squeezed" states can be applied to bosons, whether they are photons, phonons, or the mechanical quanta in a moving device. Hence in this domain it should become possible in this domain to "squeeze" the quantum-limited mechanical states to achieve "quantum non-demolition" (QND) force and displacement measurements that exceed the standard quantum limit imposed by the uncertainty principle. This realm is off in the future, but passing through the first "door" appears imminent.

CONCLUSIONS

NEMS offer access to a parameter space for sensing and fundamental measurements that is unprecedented and intriguing. Taking full advantage of it will stretch our collective imagination, as well as our current methods and "mindsets" in micro- and nanodevice science and technology. It seems certain that many new applications will emerge from this new field. Ultimately, the nanomechanical systems outlined here will yield to *true* nanotechnology. By the latter I envisage reproducible techniques allowing mass-production of devices of arbitrary complexity, that comprise, say, a few million atoms − *each of which is placed with atomic precision* [38]. Clearly, realizing the "Feynmanesque" dream will take much sustained effort in a host of laboratories. Meanwhile, NEMS, as outlined here, can *today* provide the crucial scientific and engineering foundation that will underlie this future nanotechnology.

ACKNOWLEDGEMENTS

REFERENCES

1. A.N. Cleland and M.L. Roukes, "Fabrication of high frequency nanometer scale mechanical resonators from bulk Si substrates", *Appl. Phys. Lett.*, **69**, 2653 (1996).

2. M.L. Roukes, "Yoctocalorimetry: Phonon Counting in Nanostructures", *Physica B: Condensed Matter* **263-264**, 1 (1999).

3. Y.T. Yang, K. Ekinci, C. Zorman, M. Mehregany, and M.L. Roukes, "Silicon carbide NEMS resonators", *to be published*.

4. R.E. Rudd and J.Q. Broughton, "Coarse-grained molecular dynamics and the atomistic limit of finite elements", *Phys.Rev.* B **58**, R5893 (1998).

5. K.L. Ekinci, X.-M. Huang, and M.L. Roukes, "Frequency tuning and internal strain in NEMS and MEMS devices", *to be published*.

6. C. T.-C. Nguyen, L.P.B Katehi, G.M. Rebeiz, "Micromachined devices for wireless communications", *Proc. IEEE* **86**, 1756 (1998).

7. D.A. Harrington and M.L. Roukes, "Electrical tuning of the frequency, nonlinearity, and dissipation factor of NEMS resonators", *Caltech Technical Report (Dec. 1994), unpublished*.

8. A. Tewary, K.Y. Yasumura, T.D. Stowe, T.W. Kenny, B.C. Stipe and D. Rugar, "Low temperature mechanical dissipation in ultrathin single-crystal silicon cantilevers", Bull. Am. Phys. Soc. **45** (1), 600 (2000).

9. J.R. Vig and Y. Kim, "Noise in microelectomechanical system resonators", *IEEE Transactions on UFFC*, **46**, 1558 (1999).

10. Ron Lifshitz and M.L. Roukes, "Thermoelastic damping in micro- and nanomechanical systems", *Phys. Rev.* B **61**, 5600 (2000).

11. T.R. Albrecht, P. Grütter, D. Rugar, D.P.E. Smith, "Low temperature force microscope with all-fiber interferometer", *Ultramicroscopy* **42**, 1638 (1992).

12. L.C. Gunn, E. Buks, and M.L.Roukes, "Integrated near-field optics for NEMS motion detection", *to be published*.

13. See Ref. 1; see also A.N. Cleland and M.L. Roukes, "External control of dissipation in a nanometer-scale radio frequency mechanical resonator", *Sensors and Actuators*, **72**, 256 (1999); D.A. Harrington and M.L. Roukes, "Ultimate limits of magnetomotive displacement sensing for NEMS", *to be published*.

14. M.J. Murphy, F.G. Monzon, and M.L. Roukes, "NEMS motion detection by the local Hall effect", *to be published*; Also: Caltech Patent Disclosure No. 3025, July 1999.

15. R.G. Beck, M.A. Eriksson, and R.M. Westervelt, "GaAs/AlGaAs self-sensing cantilevers for low temperature scanning probe microscopy", *Appl. Phys. Lett.* **73**, 1149 (1998).

16. R.H. Blick, M.L. Roukes, W. Wegscheider, and M. Bichler, "Freely-suspended two dimensional electron gases", *Physica B*, **251**, 784 (1998);

17. See, *e.g.*, C.T.-C. Nguyen, "Micromechanical components for miniaturized low-power communications", *Proc. of the IEEE MTT-S Symposium on RF MEMS (Anaheim, CA)*, 48 (1999).

18. A.N. Cleland and M.L. Roukes, "A nanometer scale mechanical electrometer", *Nature* **392**, 160 (1998).

19. T.S. Tighe, J.M. Worlock, and M.L. Roukes, "Direct thermal conductance measurements on suspended monocrystalline nanostructures", *Appl. Phys. Lett.*, **70**, 2687 (1997).

20. K. Schwab, E.A. Henriksen, J.M. Worlock, and M.L. Roukes, "Measurement of the quantum of thermal conductance", *Nature* **404**, 974 (2000).

21. E.M. Purcell, H.C. Torrey, and R.V. Pound, "Resonance absorption by nuclear magnetic moments in a solid", *Phys Rev.* **69**, 37 (1946); F. Bloch, W.W. Hansen, and M. Packard, "The nuclear induction experiment", *Phys. Rev.* **70**, 474 (1946).

22. J.A. Sidles, "Noninductive detection of single-proton magnetic resonance", *Appl. Phys. Lett.* **58**, 2854 (1991).

23. D. Rugar, O. Züger, S Hoen, C.S. Yannoni, H.-M. Vieth, and R.D. Kendrick, "Force detection of nuclear magnetic resonance", *Science* **264**, 1560 (1994).

24. P.C. Hammel, Z. Zhang, G.J. Moore, and M.L. Roukes, "Subsurface imaging with the magnetic resonance force microscope" *J. Low Temp. Phys.* **101**, 59 (1995).

25. D. Pelekhov, A. Suter, M.L. Roukes, and P.C. Hammel , *to be published*.

26. G. Binnig, C.F. Quate, and Ch. Gerber, "Atomic force microscope," *Phys. Rev. Lett.* **56**, 930–933, (1986).

27. G.U. Lee, D.A. Kidwell, and R.J. Colton, "Sensing discrete streptavidin–biotin interactions with atomic force microscopy," *Langmuir* **10**, 354–357, (1994).

28. E.-L. Florin, V.T. Moy, and H.E. Gaub, "Adhesion forces between individual ligand-receptor pairs," *Science* **264**, 415–417, (1994); V.T. Moy, E.-L. Florin, and H.E. Gaub, "Intermolecular forces and energies between ligands and receptors", *Science* **266**, 257–259, (1994).

29. J.H. Hoh, J.P. Cleveland, C.B. Prater, J.-P. Revel, and P.K. Hansma, "Quantized adhesion detected with the atomic force microscope," *J. Am. Chem. Soc.* **114**, 4917–4918, (1992).

30. B.L. Smith, *et al.*, "Molecular mechanistic origin of the toughness of natural adhesives, fibres, and composites, *Nature* **399**, 761 (1999).

31. See, *e.g.*, K. Visscher, M.J. Schnitzer, and S.M. Block, "Kinesin motors studied with an optical force clamp", *Biophysical Journ.* **74**, A49 (1998).

32. J.C. Meiners and S.R. Quake, "A direct measurement of the hydrodynamic interaction between two particles", *Phys. Rev. Lett.* **82**, 2211 (1999).

33. M.L. Roukes and M.C. Cross; "BioNEMS: nanomechanical devices in solution.", *to be published*.

34. See, *e.g.*, D.R. Baselt, G.U. Lee, K.M. Hansen, L.A. Chrisey, and R.J. Colton, "A high-sensitivity micromachined biosensor", *Proc. IEEE* **85**, 672-680 (1997), and references contained therein.

35. M.L. Roukes, S.E. Fraser, M.C. Cross and J.B. Solomon, Caltech Patent Disclosure, April 2000.

36. C.M. Caves, K.S. Thorne, R.W.P. Drever, V.D. Sandberg, and M. Zimmerman, "On the measurement of a weak classical force coupled to a quantum mechanical oscillator. I. Issues of principle" *Rev. Mod. Phys.* **52**, 341 (1980).

37. See, *e.g.*, B. Yurke, "Detecting squeezed boson fields via particle scattering", *Phys. Rev. Lett.* **60**, 2476 (1988).

38. R.P. Feynman, "There's plenty of room at the bottom", *American Physical Society Meeting*, Pasadena, CA, 1960.

AUTHOR INDEX

KEYWORD INDEX

Tunable capacitor - 246
Tunable devices - 324
Tuning – 363

Ultrasonic – 206

Vacuum microelectronics - 324
Variable capacitor - 255
Vibratory gyroscope - 289
Viscous drag - 124
Visualization - 132
Voltammetry – 29

Wafer bonding – 85, 300
Wafer bow – 316

Yaw - 293
Yaw rate sensor - 289
Young's modulus - 312