TECHNICAL DIGEST

SOLID-STATE SENSOR AND ACTUATOR WORKSHOP

Sponsored by the Transducers Research Foundation, Inc.

1996

Hilton Head Island, South Carolina
June 3-6

TRF Catalog Number 96TRF-0001
Library of Congress Number 96-060172

TECHNICAL DIGEST

SOLID-STATE SENSOR AND ACTUATOR WORKSHOP

Sponsored by the Transducers Research Foundation, Inc.

Additional support provided by the National Science Foundation

1996

Hilton Head Island, South Carolina
June 3-6

TRF Catalog Number 96TRF-0001
Library of Congress Number 96-060172

TRF Catalog Number 96TRF-0001
ISBN Number 0-9640024-1-8
Library of Congress Number 96-060172

Copies of this volume may be purchased for $75.00 each plus a shipping and handling charge of $5.00 (domestic) or $10.00 (overseas) per digest by sending the amount in US funds to the Transducers Research Foundation, Inc., P.O. Box 18195, Cleveland Heights, Ohio 44118. Technical digests from previous Workshops are also available.

1996 Solid-State Sensor and Actuator Workshop

On behalf of the entire organizing committee, I am pleased to welcome you to Hilton Head Island for the seventh Solid-State Sensor and Actuator Workshop. Many of us have fond memories of previous "Hilton Head Workshops." If you are attending your first Hilton Head Workshop, I hope that you will share in this enthusiasm for Hilton Head's rich tradition by week's end.

The Hilton Head Workshop is one of the regional meetings held in North America, Europe, and Asia during the even-numbered years, that alternate with the international TRANSDUCERS conferences held in odd-numbered years. TRANSDUCERS '95 was held in Stockholm, Sweden and TRANSDUCERS '97 will take place in Chicago in June 1997.

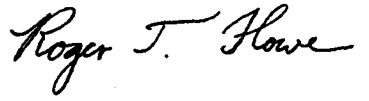
The format of the Workshop is intentionally designed to provide a highly interactive forum for US and Canadian workers to meet and discuss recent advances in emerging technologies for sensing and actuating devices and systems for physical, chemical, and biological applications. It is a single session meeting, with ample time allotted for discussion of each paper, and with blocks of unscheduled time to encourage informal interactions among participants. From the very first meeting in 1984, graduate students have been encouraged to attend and participate in the Workshop. Many of the active professionals in attendance at this Workshop gave their first papers as graduate students here at Hilton Head.

The work of putting the formal technical program in place started last fall with the call for papers. The Program Committee reviewed all the papers that were submitted last winter and considered a number of options for invited papers to make up the formal technical program. Out of 127 submitted abstracts, about half were selected for oral or poster presentations. In order to provide a forum for the presentation of recent results, a "Late News" Poster Session will be held on Wednesday evening before the Rump Session. As in the past, the Rump Session will be devoted to a discussion topic of major importance to the field. In keeping with past Workshops, Tuesday afternoon is left open for participating in leisure activities with your colleagues. I encourage each of you to take full advantage of the various opportunities that the Workshop offers.

We were fortunate to have the support of a number of organizations and individuals in organizing the Workshop. The Workshop is sponsored by the non-profit Transducers Research Foundation, which provided the business infrastructure for the meeting. Special thanks is due to the National Science Foundation and Dr. Deborah Crawford for a travel assistance grant for graduate students presenting papers, and to Richard Muller for his diligence in obtaining and distributing this support on behalf of the Workshop. Sheila Kelly of BSAC provided outstanding support for registration and the General Chairman. The Workshop homepage on the internet was set up and maintained by Jack Judy of BSAC.

In closing, I thank the entire Organizing Committee -- with special thanks to Mark Allen for his arrangements of this fine meeting site, to Joe Giachino for looking after the important financial matters, and to the Program Committee chaired by Tony Ricco for its work in arranging the technical program. Finally, I thank the authors for providing the essential contributions of their research to the Workshop, and to all the participants for being here and taking part to the fullest.

Enjoy your Workshop!



Roger T. Howe
General Chairman

1996 Solid-State Sensor and Actuator Workshop

Organizing Committee

General Chairman	Roger T. Howe University of California at Berkeley
Technical Program Chairman	Antonio J. Ricco Sandia National Laboratories
Treasurer	Joseph M. Giachino Ford Motor Company
Publications	David S. Eddy General Motors Corporation
Local Arrangements	Mark G. Allen Georgia Institute of Technology
Program Meeting Arrangements	Robert J. Huber University of Utah
Student Travel Grants	Richard S. Muller University of California at Berkeley

Technical Program Committee

Luc Bousse Caliper Technologies Corp.	Mehran Mehregany Case Western Reserve University
David W. Burns Honeywell Technology Center	David J. Monk Motorola, Inc.
Gregory T.A. Kovacs Stanford University	Khalil Najafi University of Michigan
Noel C. MacDonald Cornell University	Kurt E. Petersen Cepheid, Inc.
Martin A. Schmidt Massachusettes Institute of Technology	Yang Zhao Analog Devices, Inc.

1996 Solid-State Sensor and Actuator Workshop
June 3 - 6, 1996
Crowne Plaza Resort, Hilton Head Island, South Carolina

PROGRAM and TABLE OF CONTENTS

Sunday, June 2

7:00 - 9:00 pm ***WELCOME RECEPTION*** **with no-host bar**

* * *

Monday, June 3: Session 1

Monday, June 3: Sessions 2 & 3

Monday, June 3

Tuesday, June 4: Session 5

Wednesday, June 5: Sessions 6 & 7

Wednesday, June 5: Session 8

Thursday, June 6: Sessions 9 & 10

Thursday, June 6: Session 11

DESIGN SECRETS AND NEW DIMENSIONS IN FIELD DRIVEN SEPARATIONS FOR MICRO TOTAL ANALYSIS SYSTEMS (μ-TAS)

Andreas Manz
Imperial College of Science, Technology and Medicine,
Zeneca / SmithKline Beecham Centre for Analytical Sciences,
Dept. Chemistry, London SW7 2AY, United Kingdom

INTRODUCTION

Quantitative on-line surveillance of a chemical compound is usually based on sampling, sample pretreatment, including separation of mixtures, detection and interpretation of the obtained results. If all these steps are performed automatically in flowing streams, it is called a Total Analysis System (TAS). Several different types of TAS have been constructed and used for such surveillance tasks. Disadvantages of TAS for liquid samples include slow transport of a sample from the point of sampling to the location of the detection, slow separation processes, poor efficiencies of rapid separations and large consumption of fluids. Miniaturization with respect to the above mentioned criteria leads to a so called Miniaturized Total Analysis System (μ-TAS) and overcomes some of the above problems [1-3].

The recent developments in micromachining and microsystem technology influenced the modern chemistry especially by combinatorial synthesis, chemical sensors, valves and pumps. A spectacular work by Steve Terry on integration of a gas chromatograph on a silicon wafer unfortunately has not attracted sufficient attention then.

Examples of earlier work done with my group at Ciba, e.g., flow-injection analysis, continuous sample pretreatment using free-flow electrophoresis, HPLC, rapid capillary electrophoresis and synchronized cyclic capillary electrophoresis, have demonstrated the advantages and problems of μ-TAS [3-9].

A FEW SIMPLE RULES

For the design of electrophoretic microstructures, it is important to follow a few rules [2]:

Rules for Separation Efficiency.

1. Principally, the absolute voltage drop obtained between the injection and detection points defines the maximum number of theoretical plates N obtainable and the corresponding resolution of two neighbouring peaks ($\sqrt{N}$). The higher the applied voltage, the better.
2. For the separation of ions having a differnt charge, 100 or more theoretical plates are sufficient, but for ions of similar or identical charge, 100,000 or more theoretical plates are needed to give a reasonable resolution.
3. Joule heating inside the capillary and the heat transport across should be kept under control. A power of less than 1 W/m is generally no problem.

Hence, the maximum plate number N_{max} obtainable after a miniaturization of a known system is proportional to the length L of the capillary divided by the diameter d.

$$N_{max} \propto \frac{L}{d}$$

If the miniaturization involves identical factors for L and d, there should be no loss in performance. It can be improved, however, by increasing L or decreasing d.

Rules for Analysis Time.

4. The length of the capillary influences the analysis time. A few centimeters can give a few tens of seconds.
5. The efficiency of the separation column will only be visible, if contributions of injection and detection processes are small compared to volume or length of the separation column. This includes space (length, volume) and time effects (response time).

Hence, the minimum analysis time t_{min} obtainable after a miniaturization of a known system is proportional to the length L of the capillary multiplied by the diameter d.

$$t_{min} \propto L \cdot d$$

0-9640024-1-8 $3.00 © 1996 TRF

Solid-State Sensor and Actuator Workshop
Hilton Head, South Carolina, June 2-6, 1996

This means, the only way of increasing the speed of an analysis and keeping the efficiency constant is to reduce the diameter and the length of a capillary by the same factor.

The above considerations are used for the very simple capillary electrophoresis system. As you can see from Figure 1, there would be other forms of detection (time constant), or other types of running the electrophoresis (free-flow electrophoresis).

In most cases, a spreadsheet calculation of a model case will give sufficient information for the design of a microstructure layout. One has to bare in mind, however, that the calculated efficiency and analysis time values may vary up to a factor 2.

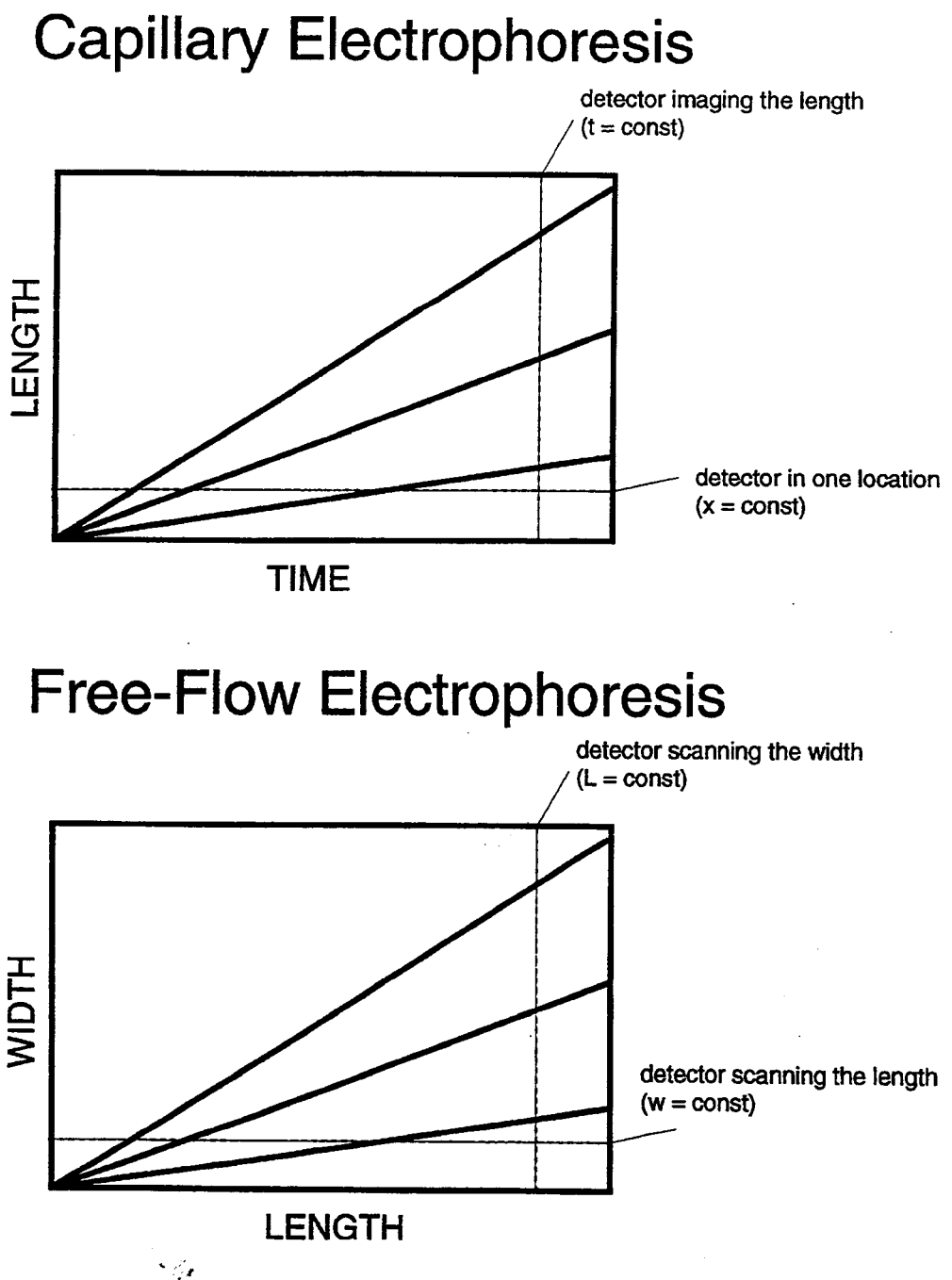


Figure 1. Comparsion of the two most common detection schemes for capillary electrophoresis (CE) and free-flow electrophoresis (FFE).

EXAMPLE: FREE-FLOW ELECTROPHORESIS

For the primary sample preparation steps in a μ-TAS, continuous techniques are clearly advantageous, e.g., filtration, dialysis, SPLITT or free-flow electrophoresis (FFE). A special version of continuous zone electrophoresis in quartz channel structures has been presented by Mesaros et al. recently. The general principle of FFE is given in Figure 2. A pumped flow (carrier) in a first dimension and an electric field in a second dimension allows for the continuous separation of sample components being continuously introduced at the top of the flat separation bed. Conventional FFE is a preparative method used for the isolation of various biological components. Typical separation bed dimensions are 10 cm wide, 50 cm long and 0.4 to 1 mm deep. Ion exchange membranes are used to isolate the electrodes from the separation bed. FFE has been miniaturized and integrated onto a silicon micro structure by reducing all dimensions by a factor of 10 ([6], Figure 3). The ion exchange membranes have been replaced by arrays of narrow 'V'-groove channels each 10 μm deep and 12 μm wide at the base.

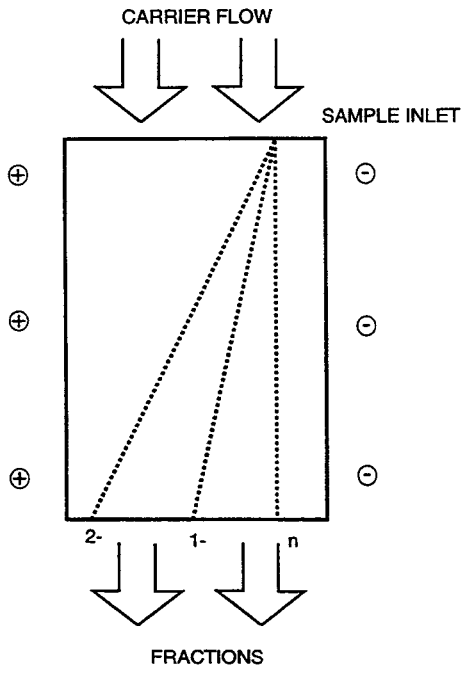


Figure 2. Schematic drawing of a FFE separation.

It is possible to achieve a continuous separation of small ions according to their charge. A voltage drop of 50 V gave a reasonable resolution of ions differing by one charge unit with a response time of 2 to 5 minutes (see Figure 4). The fraction of interest can be isolated and transferred to a next pretreatment step or a detection unit. Furthermore, it is possible to separate small ions from biological cells. Increased voltages should allow for higher separation efficiency and a response time of a few seconds.

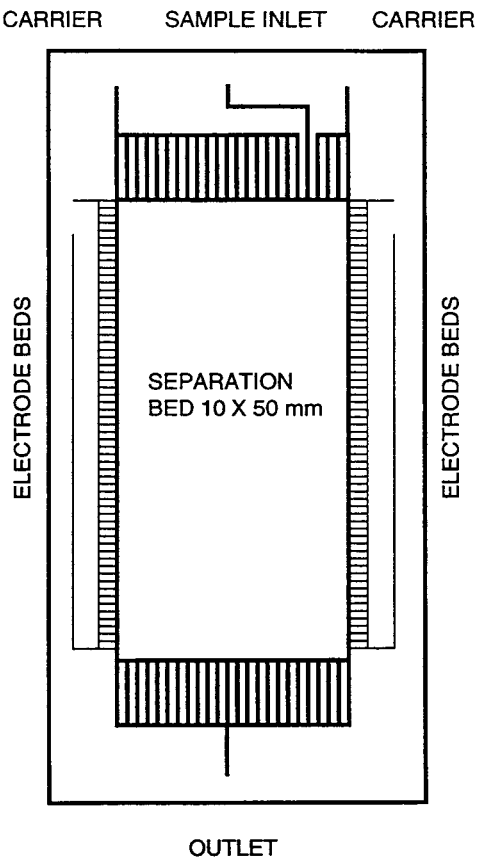


Figure 3. Layout drawing of the silicon part of a FFE separation microstructure.

CONCLUSION

For the design of electrophoretic microstructures, a few simple considerations may be very helpful to reach a performance near the desired optimum. This can be high speed, maximum performance, a compromise between the two, or a hyphenation of several separations. In the future, applications in the bioanalytical area involving molecular recognition, e.g., immunoassays or DNA hybridization, may offer a large number of opportunities which only need a defined intermediate performance. On the other hand, the electrophoretic separation as such and especially

the two-dimensional slab gel electrophoresis can be optimized to its maximum and offers a vast separation efficiency sufficient for thousands of components in a mixture.

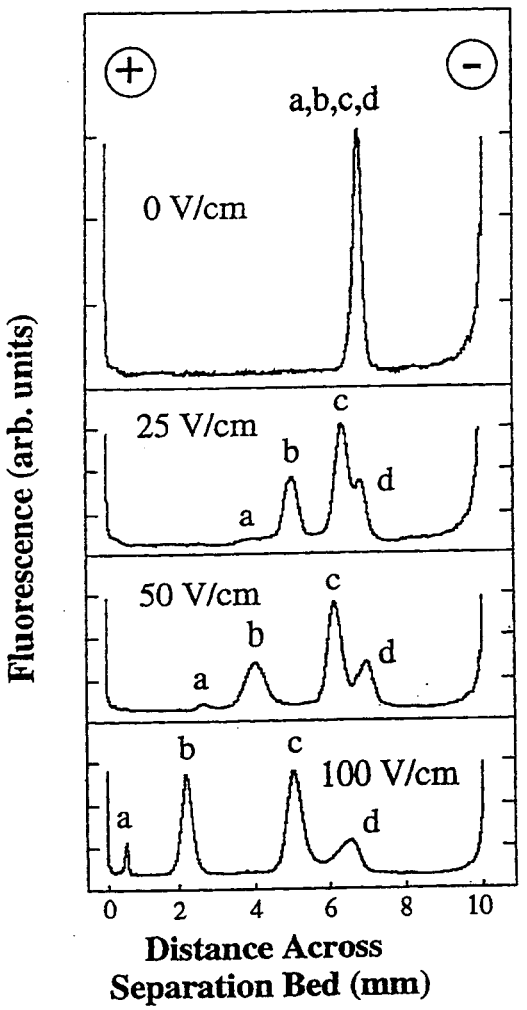


Figure 4. Separation of large biomolecules at various field strengths. Sample (a) FITC label, (b) human serum albumin, (c) bradykinin, and (d) ribonuclease A.Flow rate 15 nL/min, 20 mM borate, 100 mM TRIS buffer (pH 9). Detection after 31 mm, residence time 62 seconds.

References

[1] internet: http://www.ch.ic.ac.uk/manz/

[2] A.Manz, D.J.Harrison, E.Verpoorte, H.M.Widmer, "Planar chips technology for miniaturization of separation systems: a developing perspective in chemical monitoring" in Advances in Chromatography, Vol. 33, Chapter 1, pp 22-36 (1993).

[3] A.Manz, E.Verpoorte, D.E.Raymond, C.S.Effenhauser, N.Burggraf, H.M.Widmer, "μ-TAS - Miniaturized total chemical analysis sytems" in Micro Total Analysis Systems, A.van den Berg, P.Bergveld, eds., MESA Monograph, Vol. 1, Kluwer Academic Publishers, Dordrecht, p. 5-27 (1995)

[4] C.S.Effenhauser, A.Manz, H.M.Widmer, "Manipulations of sample fractions on a capillary electrophoresis chip", Analytical Chemistry 67, 2284-2287 (1995)

[5] G.Ocvirk, E.Verpoorte, A.Manz, M.Grasserbauer, H.M.Widmer, "High-performance liquid chromatography partly integrated onto a silicon chip", Analytical Methods and Instrumentation 2, 74-82 (1995)

[6] D.E.Raymond, A.Manz, H.M.Widmer, "Continuous sample pretreatment using a free-flow electrophoresis device integrated onto a silicon chip", Analytical Chemistry 66, 2858-2865 (1994)

[7] N.Burggraf, A.Manz, E.Verpoorte, C.S.Effenhauser, H.M.Widmer, N.F.de Rooij, A novel approach to ion separations in solution: synchronized cyclic capillary electrophoresis", Sensors and Actuators B20, 103-110 (1994)

[8] R.Freaney, A.McShane, T.V.Keaveny, M.McKenna, K.Rabenstein, F.W.Scheller, D.Pfeiffer, G.Urban, I.Moser, G.Jobst, A.Manz, E.Verpoorte, H.M.Widmer, D.Diamond, E.Dempsey, M.Smyth, "Novel instrumentation for real-time monitoring using miniaturised flow-systems with integrated biosensors", Annals of Clinical Biochemistry, submitted

[9] E.Verpoorte, B.H.van der Schoot, S.Jeanneret, A.Manz, H.M.Widmer, N.F. de Rooij, "Three-dimensional micro flow manifolds for miniaturized chemical analysis systems", J. Micromechanics and Microengineering 4, 246-256 (1994)

IMMUNOASSAY FLOW SYSTEMS ON-CHIP

D. Jed Harrison, Nghia Chiem
Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada, T6G 2G2

ABSTRACT

Integrated microfluidic systems for biochemical analyses and immunoassays were prepared in glass substrates using silicon micromachining techniques. All of the fluid handling steps are performed by applied electric fields, eliminating the need for pumps and microvalves. Electroosmotic pumping was used to drive immunological reagents and samples into a separation channel and the reaction products of the immuno-reaction were evaluated. The extent of reaction can be used to quantitate the amount of sample present when uncomplexed reagents are separated from the product complex. Concentration detection limits of about 30 pM were obtained for fluorescein using fluorescence detection. Detection of the drug theophylline at 5 ng/ml was shown in an immunoassay. Immunoassays for the large proteins bovine serum albumin and human immunoglobulin were also performed on-chip.

INTRODUCTION

Microfluidic devices etched in glass substrates provide a fluidic network in which chemical reactions, sample injection and separation of reaction products can be achieved. The application of high voltages to conductive fluids within these channels leads to electroosmotic and/or electrophoretic pumping, providing both mass transport and separation of components within the samples. In this paper we present on-chip, clinically relevant, immunological assays for proteins such as albumin, and drugs such as theophylline. Electroosmotic pumping is used to deliver reagents and samples, and separation of the reaction products is performed using electrophoresis [1]. This method pumps fluid at velocities up to 1 cm/s in approximately 20 μm diameter capillaries, while also controlling the direction of fluid flow at capillary intersections, without a need for valves or other moving parts. Electrophoretic effects lead to separation: i.e. the differing mobilities of ions result in different migration rates within an electric field, giving rise to separations.

Application of electrophoresis to immunoassays is a newly developing field [2-5]. Antibodies (Ab) are highly selective biological molecules originating from the immune system, capable of binding with high affinity to a particular target molecule known as an antigen (Ag). Immunologically based assays provide a common method to selectively determine many chemicals at low concentration. In common immunosorbent assays, in which the reagents are adsorbed onto surfaces, there are many complex fluid handling steps involved. We expect that the unique fluid delivery capabilities of the micro-electrophoresis chip format will provide a novel, important new method of automating immunoassays in a manner which does not involve the conventional robotics used in clinical labs today. The chip format will also provide much more rapid presentation of results for samples in clinical and environmental applications. The key step towards developing this technology illustrated in this paper is the separation of products and reactants resulting from homogeneous immunoassays, and the demonstration of quantitative analysis of the targeted sample.

EXPERIMENTAL

Devices were fabricated in Pyrex glass or a photomask crown glass [6]. The potential control and measurement instrumentation has also been presented previously [6]. Buffers were adjusted to pH 8 or 9. Antibodies, other proteins, and the fluorescent labels were purchased from Sigma. The theophylline reagents used were part of the Abbott TDx reagent set.

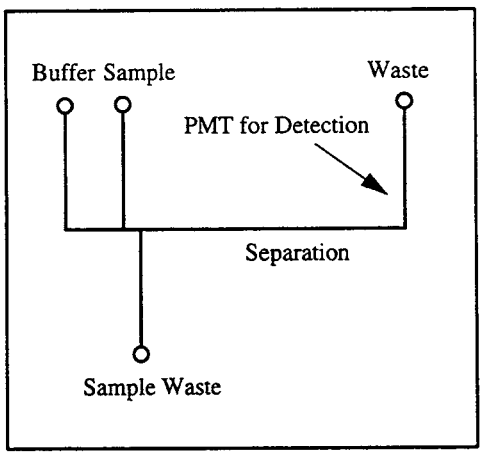


Figure 1 Layout of the flow channels used for immunoassay chip.

RESULTS AND DISCUSSION

Integrated devices for chemical separation employed in this study are based on capillary electrophoresis (CE), a method in which high applied voltages can be used to separate the components of a sample with high efficiency. Figure 1 illustrates the layout of the fluid channels in the chip used for these studies. Sample is introduced through a side channel and driven towards sample waste with 500 to 1500 V, creating a 60-100 pL plug of solution at the intersection, which can be injected into the separation channel. Fluorescence detection is used with a 488 nm laser excitation source and a photomultiplier tube for detection of

0-9640024-1-8 $3.00 © 1996 TRF

the emission from fluorescently labelled molecules. This detection system is highly selective, as only fluorescent compounds will be observed [7].

Sensitivity is a key issue in immunoassays, as most of the compounds of interest are present at very low concentrations. We have improved detection substantially since earlier reports, by optimizing the wafer bonding process to reduce formation of light scattering centers on the surface of the glass devices. Careful alignment of the laser beam to eliminate scatter from the curved walls of the channels was also beneficial. In this way we have reduced our previous detection limits from about 2 nM down to 30 pM. For the optical cell geometry used, this detection limit corresponds to about 200 molecules in the detection volume. Figure 2 shows the signal obtained for 100 pM fluorescein, which gives a signal to noise ratio of 14. Detection limits in the pM range are satisfactory for a number of immunoassays.

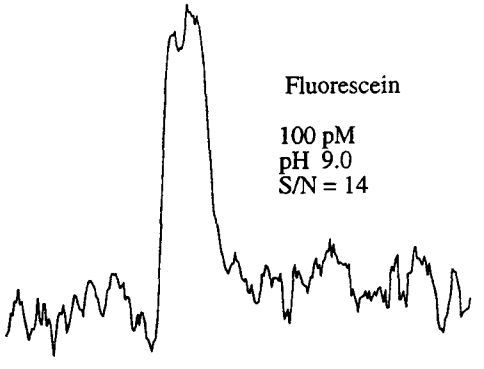


Figure 2 *Fluorescence signal from 100 pM fluorescein in the chip.*

Homogeneous immunoassays are solution phase reactions of an antibody and its a target molecule, the antigen. By labelling either the antigen (Ag) or the antibody (Ab) with a fluorescent tag, it is possible to determine when a complex between antibody and antigen has been formed. (We will designate labelled compounds with a *.) Figure 3 shows an on-chip separation of a labelled antibody (Ab*) before and after reaction with its antigen. In this case the antigen is human immunoglobulin G (IgG, itself an antibody), while the antibody is fluorescently labelled goat antihuman IgG. The complexes formed between the two compounds have a different charge to size ratio than either of the free compounds, so they migrate at different rates in the electric field. The unlabelled 160,000 Dalton target, human IgG, can be determined from its complex formation with the anti-human antibody, while the remaining free antibody is separated from the reacted material. Figure 3 shows there are two product peaks, representing formation of both the Ab*-Ag and Ab*-Ag2 complexes, as the antibody is bivalent.

Figure 4 illustrates an immunoassay involving a lower molecular weight protein than in the above study. This is an assay for the antibody to the 60,000 Dalton protein bovine serum albumin (BSA). To determine the amount of antibody present, it is mixed with 50 µg/ml of flourescently labelled BSA* and

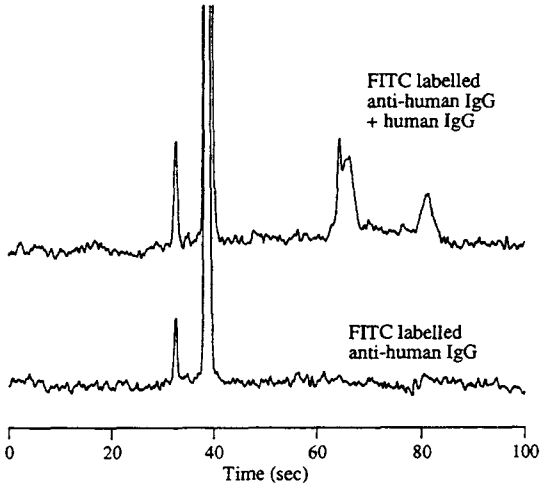


Figure 3 *Bottom trace shows separation of fluorescein labelled anti-human IgG. Top trace shows the separation after addition of human IgG.*

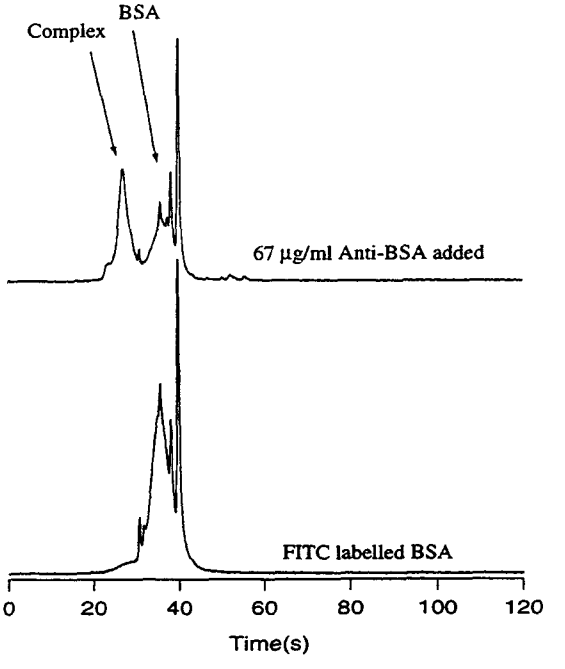


Figure 4 *Bottom trace shows separation of Fluorescein labelled BSA, with a number of impurities present in the commercial product. Top trace shows the separation after addition of mouse ascites fluid containing anti-BSA.*

separated. The complex formed migrates at a different rate, and is separated from the free BSA*. The result is a decrease in signal for BSA* and an increase in the Ab-BSA* complex peak. The increase in peak height or area is proportional to the amount of anti-BSA present.

Figure 5 shows a calibration curve for the assay of BSA antibody. In this study the anti-BSA source was a solution of mouse ascites fluid. This fluid is the extraction fluid as obtained directly from mice during monoclonal antibody harvesting.

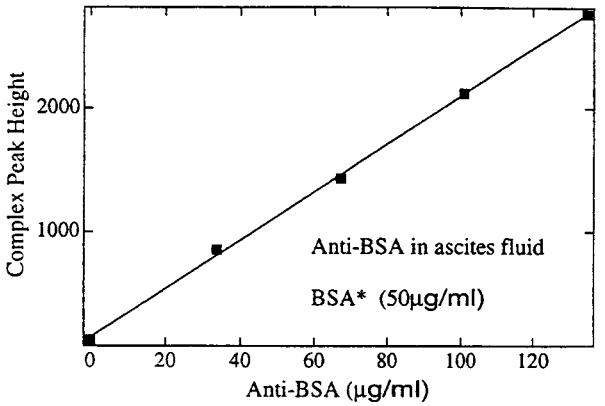


Figure 5 *Calibration curve for anti-BSA in ascites fluid, measured from the complex peak height.*

It is rich in a variety of proteins (27 mg/ml), of which only 10% is the desired antibody. The fact that good separations and quantitative assays can be obtained from this complex fluid illustrates the robustness of the microchip-CE system for these assays. Deliberate addition of other antibodies, which are not specific for BSA, showed no complex formation, indicating the selectivity of the antibody reaction. When combined with the separation power of CE, immunoassay on-chip makes for a very powerful, selective immunosensor system, capable of identifying specific chemicals or biochemicals in a complex sample mixture.

It is also of great interest to analyze for small molecules by immunoassay methods. A typical example is an analysis for a therapeutic drug for asthma, theophylline (Th). In order to get good separations by CE when using small target molecules, it is necessary to label the analyte molecule instead of the antibody. Labelled Th (Th*) is thus mixed with a sample containing unlabelled Th, and the two are allowed to compete for a limited amount of antibody to Th, as illustrated in the scheme below.

$$Ab + Th^* \Leftrightarrow Ab\text{-}Th^*$$
$$Ab + Th + Th^* \Leftrightarrow Ab\text{-}Th + Ab\text{-}Th^*$$

A competitive assay leads to an increase in signal for the free, labelled Th* as Th from a sample increases in concentration. There is a corresponding decrease in signal from the Ab-Th* complex. Figure 6 shows a series of separations performed on-chip in which increasing amounts of Th in a sample were added to a fixed amount of Th* and Ab. The separation occurs in less than 1 minute and the complex is well resolved from free Th*.

Data as in Figure 6 provides a calibration curve, such as is shown in Figure 7. Competitive assays of this type are well known to give non-linear calibration curves, but are extensively utilized in clinical diagnostics. To meet therapeutic criteria, serum samples should contain Th in the range of 10-20 µg/ml. The lower axis shows the original sample concentrations before dilution, and it is clear that high sensitivity is attained in the therapeutic range. It is common to dilute the serum sample before analysis, and in this study we used a 200-fold dilution factor. The upper axis shows the actual concentration in the measured samples, after dilution.

In a conventional assay one would follow the appearance of

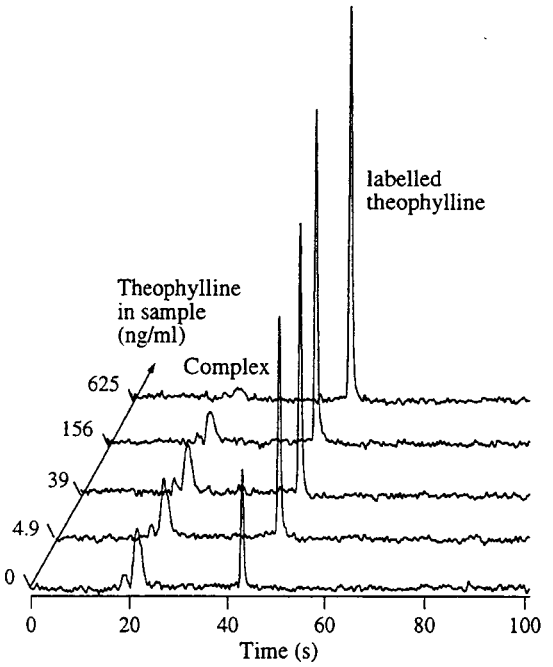


Figure 6 *Competitive immunoassay of theophylline performed on-chip. Peaks for the complex and for free Th* are mixed. The ratio of peak areas changes as Th from a sample is added.*

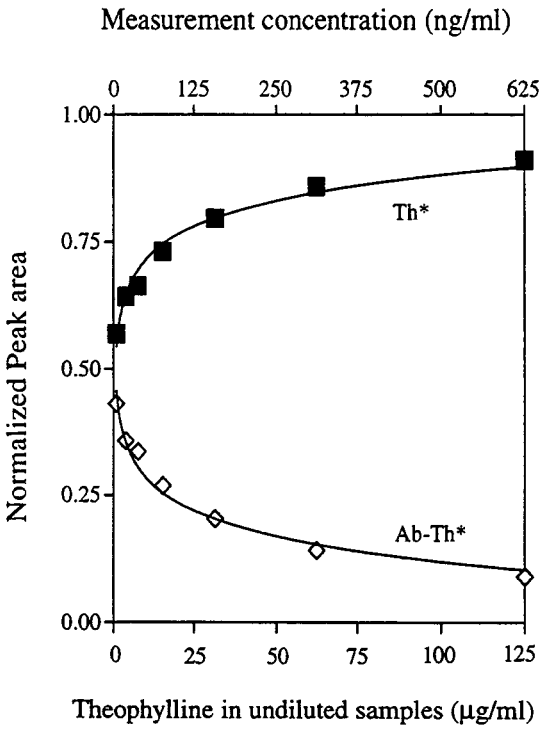


Figure 7 *Calibration curve for unlabelled Th in a sample. Lower axis shows concentrations before dilution, upper axis shows concentrations after dilution.*

free Th*. However, in this microchip format it is possible to observe both disappearance of the complex and appearance of free Th*. This process leads to greater precision of measurement. Both traces are shown in Figure 7. By measuring both components and recognizing that the total fluorescence signal should remain a constant, it is possible to normalize the signals. This compensates for small variations in the amount of material injected and gives better precision. The y-axis in Figure 7 is thus a plot of peak area for one of the two fluorescent species ratioed to the total fluorescent signal.

The calibration curve illustrates that we are readily able to measure theophylline in the therapeutic range, and that the absolute concentrations measured are in the ng/ml range. The precision of the measurements is within 5%, indicating good analytical performance.

CONCLUSIONS

Separation of proteins, in particular strongly adsorbent proteins such as immunoglobulins, offers significant difficulty for capillary electrophoresis methods [2,3]. Demonstration that it can be accomplished on-chip, for a variety of native-form IgG's from different sources is a significant step in developing major applications for CE-based microfluidic systems.

In this work we have shown that separations of a large variety of reactants and products created by immunological reactions can be performed on-chip. The antigens evaluated cover a range from very large proteins (160,000 Dalton IgG) to very small molecules (180 Dalton theophylline). Also, several types of procedures were performed effectively: the direct assay of the antigen human IgG, where the antibody is labelled; the direct assay of an antibody where the antigen is labelled; and the competitive assay of a drug, where a certain amount of the drug is labelled. The assays are highly selective. In the case of the anti-BSA determination, this protein is easily determined in the presence of many others. For the theophylline assay, the drug is determined at low concentration in blood serum, which is a complex mixture of many components. Applications for the assays described range from the determination of antibody production in a fermentation process to analysis of drugs or infectious diseases, indicating the vast array of fields these devices could be applied to.

When combined with our previous studies of pre- and post-separation reactions on-chip [1,4], it is clear that it will be possible to drive the complete process of antibody and sample mixing, incubation and separation within the confines of a chip. This offers the potential for a highly automated, rapid and easily programmed system. With a certain amount of engineering it is likely that portable systems incorporating these chips, using already existing light source, detection and power supply technologies, could be readily achieved.

ACKNOWLEDGMENTS

We thank DRES and NSERC of Canada, and Shimadzu for support. N.C. thanks the Alberta Microelectronic Centre for a graduate fellowship and use of their facilities.

REFERENCES

1. D.J. Harrison, K. Fluri, K. Seiler, Z. Fan, C.S. Effenhauser, A. Manz *Science* **261**, 895 (1993).
2. N.M. Schultz, R.T. Kennedy *Anal. Chem.* **65**, 3161 (1993).
3. K. Shimura, B. L. Karger *Anal. Chem.* **66**, 9 (1994).
4. D.J. Harrison, K. Fluri, N. Chiem, T. Tang, Z. Fan *Transducers 95 Technical Digest Vol 1* (1995) pp 752-755.
5. L.B. Koutny, D. Schmalzing, T.A. Taylot, M. Fuchs *Anal. Chem.* **68**, 18 (1996).
6. Z. Fan, D. J. Harrison *Anal. Chem.* **66**, 177 (1994).
7. M. Albin, R. Weinberger, E. Sapp, S. Moring, *Anal. Chem.* **63**, 417 (1991).

Fluid Circuit Technology: Integrated Interconnect Technology for Miniature Fluidic Devices

Don VerLee, A. Alcock, G. Clark, T. M. Huang, S. Kantor, T. Nemcek, J. Norlie, J. Pan, F. Walsworth, S. T. Wong.

Abbott Laboratories Hospital Products Division
Abbott Park, Illinois, 60064-3537

ABSTRACT

According to many in the MEMS field, medical markets currently represent the second biggest opportunity for microfabricated devices and systems [1]. Certainly, microfabricated blood pressure sensors have already found a foothold in the industry. However, packaging and interconnections between devices remain major barriers to more wide-spread use of MEMS in medical applications.

It has become evident to us that the application of MEMS to medical diagnostics is hindered by the lack of an intermediate scale fluid handling system for feeding microscopic elements. In this paper, we describe a technology that may be able to serve as an interconnect technology between microfabricated devices and as a transition between macroscopic and microscopic scales. We describe a manufacturing process, miniature fluidic elements, and system designs that may offer useful insights into how one might design a microfabricated fluid circuit system.

INTRODUCTION

The future direction for many medical diagnostic analyzers portends increasing fluid processing complexity, miniaturization, and reliability while the costs to manufacture and to operate decrease. The factors that are driving these changes include; cost containment pressures, a steady increase in the menu of tests offered and the resulting number of accessible diagnostic reagents, the need to automate front end sample processing, the need to reduce the fluid volumes of reagent and sample used, and the desire to reduce the amount of solid waste generated as a diagnostic test is run.

Diagnostic instrument and medical device designers have watched developments in MEMS technology for years hoping to apply it to solve this growing list of challenges, yet they still see barriers that limit the spectrum of its applicability. For in-vivo and in-vitro diagnostic use, protein adsorption and clotting of small passageways top the list of concerns about biologically compatible materials and surface treatments. In addition there is a fundamental need to provide an interface between the macroscopic world, where samples are aspirated and dispensed in milliliter quantities, and the microscopic device which has internal volumes of nanoliters to microliters.

It is unrealistic to assume that the macroscopic fluid volumes needed in current instrumentation can all be driven down to microscopic volumes in the future. The lower limit of concentration for some bacterial infections can be as small as one cell per milliliter of blood. It should be clear that there is a lower limit on the volume of sample that must be evaluated to assure statistically relevant results even with multiplication factors that exceed 10^{15}.

To address many of these pressing needs for handling fluids in future clinical diagnostic analyzers and medical devices, Abbott Laboratories has advanced the development of a technology that miniaturizes and integrates fluidic structures and control elements into solid blocks of acrylic - Fluid Circuit Technology.

The objective of the Fluid Circuit Technology program was to investigate and develop fundamental technology to enable the miniaturization and integration of the fluid handling functions needed in a typical medical diagnostic analyzer. The program encompassed research on miniature valves, chambers, channels, fluid metering, and mixing structures and methods, and a manufacturing process to combine these elements into a single integrated network. An additional goal was to demonstrate that this technology is functionally equivalent to existing diagnostic instrumentation while still providing the added benefits of reduced reagent and sample consumption, reduced solid waste, reduced size and cost of the instrument, and increased overall reliability.

The four aspects of Fluid Circuit Technology that will be described below are; A Process for Manufacturing Fluid Circuits, Fundamental Fluid Circuit Elements, Fluid Processing Subsystems, and examples of Fluid Circuit Systems. From the examples, parallels with MEMS technology in wafer bonding, valve design, and system design should become evident.

FLUID CIRCUIT MANUFACTURING PROCESS

Many techniques exist for bonding thermoplastics. Some of the more common techniques include solvent cementing, ultrasonic welding, adhesives, and thermal welding [2]. At the core of Fluid Circuit Technology is a process for bonding layers of thermoplastic into which fluidic structures such as channels and chambers have been machined [3]. This bonding process has the advantage over other techniques in its ability to create solid seamless blocks of plastic, without flash or bubbles, while retaining the internal three-dimensional fluid circuit detail. Originally developed and patented by companies such as Eastern Plastics, Plainville, Conn. and Carville, Dorking, England, we have investigated and optimized this process further for use in medical diagnostic instrumentation.

Fluid circuits have been fabricated out of a variety of thermoplastics, including some injection molded plastics, but the majority of the fluid circuits have been made out of cast polymethyl-methacrylate using the 5 step process that follows.

Preparing Fluid Circuit Raw Materials:

The initial process, called "shrinking", is needed to minimize dimensional drift in the bonding process, that may be caused by unrelieved stresses in the starting material. It involves baking unconfined sheets of the plastic that have been cut down to approximate size, at an elevated temperature (152° C for cast acrylic) for 10 hours, followed by a gradual cool down to room temperature. This usually causes these "blanks" to expand in thickness, and shrink in both of their other axes. The result is a block of plastic that is relieved of internal stresses, and that is relatively dimensionally stable over subsequent heating cycles.

0-9640024-1-8 $3.00 © 1996 TRF

Machining the Fluid Circuit Layers:

To begin the process of machining the channels and chambers that make up a fluid circuit, the blanks that have been "shrunk" are squared on a milling machine, annealed in an oven (90° C for 1 hr. for cast acrylic), then fixtured in a computer numerical controlled milling machine. The fluid circuit design itself is downloaded into the CNC as a list of tools and an extensive sequence of tool path coordinates in much the same manner that apertures and Gerber photoplot vectors are downloaded into a photoplotter to define the layers of a printed circuit board.

To create cylindrical channels and chambers, mirror images of the tool paths are replicated on each opposing face of the acrylic blocks that are to be bonded. The coordinates for the machining of these features are referenced from two alignment holes that are machined into the faces at opposing corners of each block. Dowel pins placed in these holes insure that when the faces are brought together both the top and bottom halves of the cylindrical cavities line up.

Several factors determine how extensive the machining process becomes. The largest contributor to machining time is the surface finish required on the inside of the chambers. In medical diagnostic instrumentation, significant attention is paid to the chemical interactions at the interfaces between solids and liquids to minimize problems such as proteins adsorbing to surfaces, carryover, uneven wetting, bubble entrapment, fluid metering imprecision, and uncontrolled side reactions. We have explored several approaches to minimizing machining time while optimizing surface finish including; flame and vapor polishing, mechanical polishing, and diamond tooling. Of these, the best surface finishes can be achieved in the least amount of time with diamond tooling.

Care is taken during both the layout and machining of the fluid circuit to assure that no channels or holes penetrate to what will eventually become the outside of the bonded block of acrylic. This is done to prevent uneven heating and collapsing of these channels during the bonding process.

A final machining step is to fly-cut the faces to be bonded with a diamond fly cutting tool. The quality of the bond is dependent on the surface finish and flatness of the two faces to be bonded and unless the edges and side walls of the channels and chambers are sharp, a seam defect will be evident after the bonding is complete.

Thermal Diffusion Bonding Fluid Circuit Layers:

After machining, the fluid circuit layers are annealed, washed with detergent and water, and inspected before they are bonded. The flat polished faces of the blocks to be bonded are aligned with dowel pins and stacked in the order that they are to be bonded inside a bonding fixture. A cross section through a bonding fixture is illustrated in Fig. 1 with the layers of a fluid circuit in it ready to be bonded. The bonding fixture is simply a steel base with a cavity in it and a floating lid. The fluid circuit layers to be bonded are machined to achieve a line to line fit when placed inside this cavity at room temperature. A cap piece is bolted onto the base with a series of springs that can be made to apply a known pressure onto the lid by compressing the springs a known distance.

Once the fluid circuit layers are stacked inside the fixture, the bolts are tightened down on the cap piece until a known pressure (~ 45 psi. for acrylic) is applied to the stack. The entire assembly is then placed in a programmable oven. A typical bonding cycle for cast acrylic includes a 2 hr. ramp to 126° C,

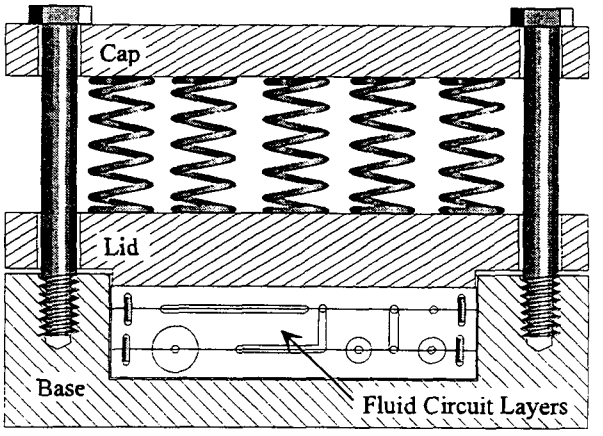


Fig. 1 Fluid Circuit Bonding Fixture

followed by a 4 hr. hold at that temperature, followed by a 2 hr. ramp down to room temperature.

The temperature at which bonding takes place is just above the glass transition temperature of the plastic. To assure that the internal features are maintained, this temperature must not be exceeded. The exact mechanism by which the bonding occurs is not fully understood, but at these elevated temperatures, the layers merge into one solid block, while retaining the fluid circuit detail that was machined in prior to bonding. The network of channels and chambers appear to "float" inside the block.

The best results we have obtained have been with cast polymers such as PolyCast™ Acrylic [4], polycarbonate, α-butylstyrene, and others. In general, highly cross-linked and highly crystalline polymers don't bond well using this method, although regions (e.g. in channels) that have been surface cross-linked through plasma or chemical treatment, and crystalline polymer films have been successfully captured between bonded layers. Bonding injection or compression molded fluid circuit layers is possible but difficult to optimize, because it is difficult to eliminate undesirable characteristics introduced during molding processes; high variance in the glass transition temperature, sink marks in thick cross sections, skinning, high unrelieved internal stresses, and rounded edges on the channel detail.

Post-bond Machining:

When the bond cycle is completed external features such as sensor pockets, external fluid connector ports, valve dimples, and mounting threads are added to the block. Initially any dimensional drift that may have occurred to internal features, during bonding is measured and used to adjust the coordinates of critical features in the original fluid circuit layout tool path code. Since this drift is fairly repeatable (~ ± 25 μm) later "as built" fluid circuits match the design to within ~ ± 10 μm near the center and within ~ ± 100 μm near the edges of a large 20 X 25 cm block.

The block is then aligned in the CNC machine so that the machine's coordinate system references the guide marks. All external features are then machined into the block with the same precision and accuracy as the internal detail of the fluid circuit. After this is complete and the block has been thoroughly washed and inspected the fluid circuit is ready for assembly.

Assembling a Fluid Circuit:

The assembly requirements differ for each fluid circuit and depend entirely on the complexity of the fluid circuit. For a simple fluid circuit the only assembly required, other than to mount fluid connectors to the block, is to create the valves. These

simple valves are assembled in one step [6]. The process sandwiches a thin film (12 μm) of pressure sensitive adhesive backed mylar (Flexcon V95 [5]) between two fluid circuit blocks. These two blocks will be called the "liquid side" and the "air side" (one for handling the biological liquids, the other for plumbing pneumatic control channels to each valve site). The air side fluid circuit is placed on a flat surface with the valve side facing up, and a sheet of the mylar film large enough to cover the block is simply rolled on starting at one edge of the block. Care must be taken not to introduce any wrinkles or dust particles under the film. The film is then carefully cut away from each bolt hole, dowel pin and fluid pass through location. Before the fluid side block is bolted on, a flat and highly polished backing plate is first bolted on and allowed to sit for 4 hrs. to make sure that the adhesive on the back of the film takes a set. The excess film is then trimmed away from the outside of the block and fluid connectors are mounted onto the assembled fluid circuit.

FLUID CIRCUIT ELEMENTS

If one were to draw the analogy that a fluid circuit was the fluidic equivalent of an electronic printed circuit, then a simple fluid circuit would consist of nothing more than channels and interconnection sites for externally mounted active components. However for a fluid circuit to be useful, structures which have no analogy to printed circuits are needed for fluid processing in addition to the channels. For example, it is possible to integrate fluid control elements such as valves and sensors, directly into the fluid circuit block itself and thus greatly reduce the "dead volume", sealing difficulties, and alignment problems encountered with mounting elements outside a fluid circuit.

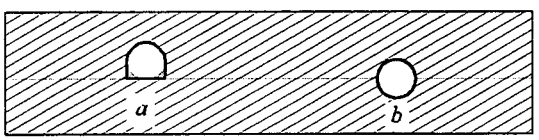


Fig. 2 "D" and Cylindrical Fluid Circuit Channels

The most fundamental fluid circuit element is a channel. The two basic fluid channel designs are illustrated in Fig. 2. Both of these are machined into the plastic using a ball end mill. The "D" channel shape illustrated in Fig. 2a is the most routinely used fluid circuit geometry. "D" channels are used extensively for plumbing pure fluids (reagents, diluents etc.), fluid sources, sinks, and air distribution networks. This channel shape is the least costly to machine since it can be fabricated in one pass on a single side of a block.

The cylindrical channel shape illustrated in Fig. 2b is the preferred geometry for all liquid networks that directly handle biological liquids or the assay liquids. It can be clearly seen that the "D" channel has two corners where liquid menisci remain when a channel is emptied. The fluid entrained in these areas causes higher carryover (or requires a greater volume of wash fluid to purge), which isn't a problem on cylindrical channels

The most common channel sizes used have been 0.8 mm, 1 mm and 1.6 mm in diameter. Smaller channel diameters have been fabricated, (0.5 mm) but still smaller diameters have not been explored extensively since fluid flow resistance, probability of plugging and vapor lock all increase with decreasing diameter.

A variety of valves can be fabricated [6] inside a fluid circuit including those pictured in Figs. 3, and 4. Each of the fluid valves

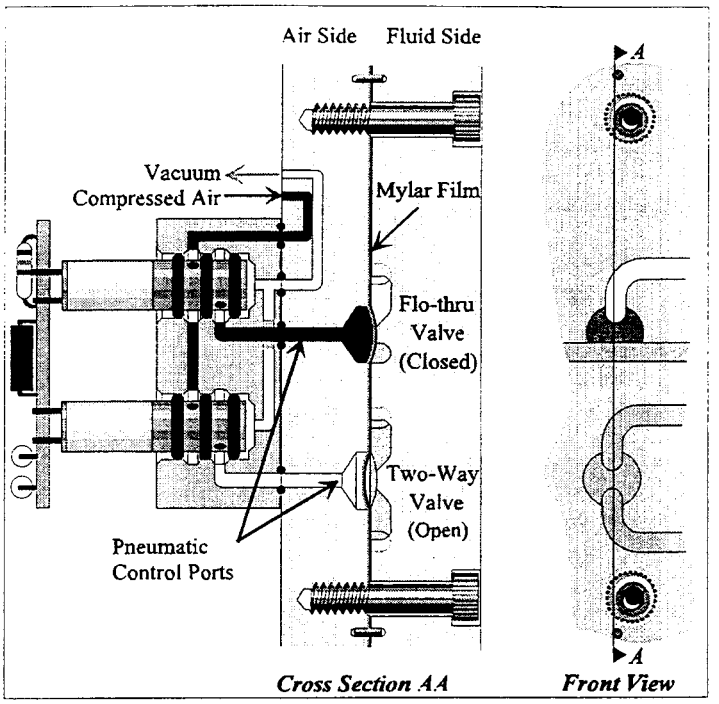


Fig. 3 Two-way and Flow-Through Fluid Circuit Valves

illustrated is pneumatically activated by a remote pilot valve which can be a miniature 3-way solenoid valve such as those from the Lee Company [7] pictured in Figs. 3 and 7. Remote pneumatic actuation results in a significant improvement in reliability of miniature fluid control systems. The difference in reliability between commercially available electromechanical liquid valves, and electromechanical pneumatic valves is from 10 - 100 X number of total cycles before failure

The two general fluid circuit valve types illustrated in Fig 3. are 2-Way and Flow-Through valves. As can be seen in this schematic, 2-Way valves provide simple on/off control of fluid flow. Flow-Through valves, however, provide more interesting and useful control. They enable one fluid to merge into the middle of a free flowing channel at a fixed location and with very little "dead volume". This valve type enable many useful fluid circuit constructs including circuits that wash the channels out, and circuits that enable reagent metering, and mixing. A variant which has two flow through channels creates a cross-over point for fluid

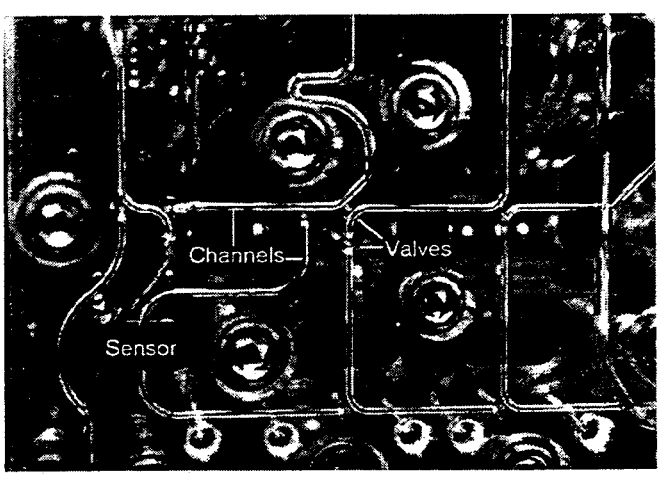


Fig. 4 Fluid Circuit Channels, Valves and Sensors

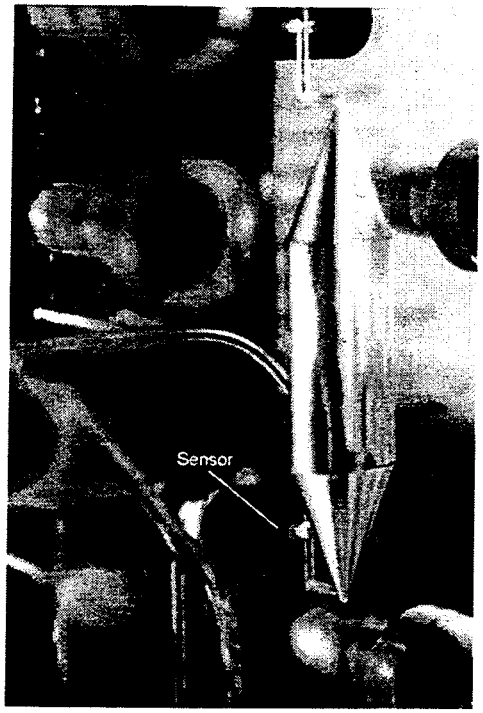


Fig. 5 10 ml Metering Chamber & Sensors

streams to cross from one flow network into another. Examples of these uses for Flow-Through valves can be seen in Fig. 4.

A variety of different chamber sizes and styles have also been fabricated inside a fluid circuit. Chambers used to meter larger volumes (Fig. 5), mix fluids, incubate mixtures (Fig. 6), provide reagent and diluent storage and provide waste sumps have been designed and tested. The incubation chambers shown in Fig. 6 have a cavity machined around them that has been filled with thermally conductive epoxy and assembled contacting a flexible heater attached to an aluminum plate. This created a "hot block" inside the fluid circuit to incubate assay reaction mixtures.

To provide feedback for controlling the fluid progress through a fluid circuit, sensors have been developed which can distinguish between air and liquid in the channel [8]. The sensors pictured in Figs. 4, 5, & 6, are composed of an LED and a photodiode oriented in an opaque housing in such a manner so as to illuminate and reflect off a channel. A logic HI signal is generated if air is present, logic LO if liquid is present.

FLUID CIRCUIT SUBSYSTEMS

The elements that make up a fluid circuit have been combined to create functional subsystems. These have been utilized routinely either as repeated designs copied from one successful fluid circuit to another, or as stand-alone modules that connect to a larger fluid circuit system to carry out a function.

Several of these fluid circuit subsystems are illustrated in the figures. Fig. 7 illustrates a manifold of 16 three-way solenoid valves. This fluid circuit design provides compressed air and vacuum to each of 16 miniature Lee valves and enables a simple face-mount, leak free connection of each of the valves' common outputs to a fluid valve in a fluid circuit. The fluid circuit provides both compressed air and vacuum to the module and makes the connections from the module to each fluid valve control point. The drive electronics enable direct digital access to the module and can be modeled as a 16-bit memory word.

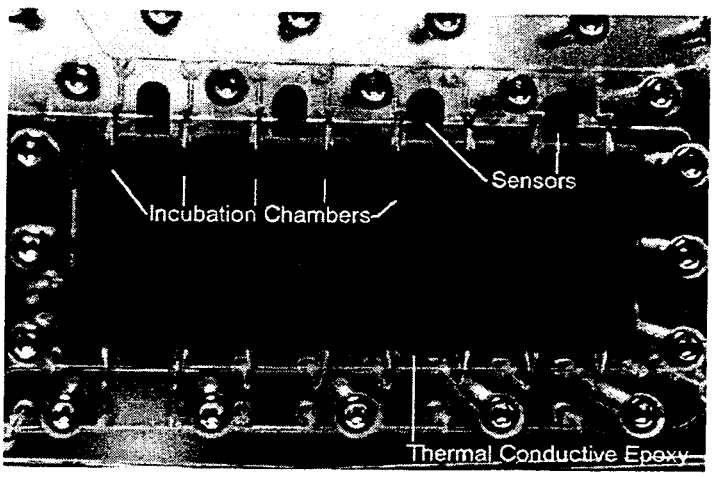


Fig. 6 Incubation Queue of 8 Chambers

Several fluid metering circuits are also pictured in Figs. 4, 5, and 8. In just one example, Fig. 4 illustrates a close-up of a sensor to sensor fluid metering circuit that makes up the center section of the fluid circuit system pictured in Fig. 8. This circuit was designed to allow one fluid (sample or reagent) to be metered across a Flow-Through valve into another fluid by advancing the meniscus of the fluid to be metered from one sensor to another.

FLUID CIRCUIT SYSTEMS

Once initial feasibility had been demonstrated for fluid circuit elements, and subsystems, it became evident that Fluid Circuit Technology could provide a new fluid interconnect and control strategy for entire diagnostic instrument systems. The design approaches used in the development of complex electronic circuits could be applied to the design and development of a complex fluid control system. Proven functional modular subsystems based on simple circuit elements could be simply face or edge connected to other fluid circuits to make a reliable fluid circuit instrument system.

A feasibility project to demonstrate this concept was undertaken in 1992. The objective of the project was to demonstrate that a fluid circuit based instrument could produce comparable performance (throughput, assay sensitivity and specificity) to an existing commercial instrument. Implementing this hierarchical design philosophy meant that the design team had to abandon traditional instrument design approaches that utilized robotics and disposable plastic containers to meter, mix and manipulate fluids, and focus instead on what needed to be done with the fluids to carry out a typical diagnostic assay chemical reaction.

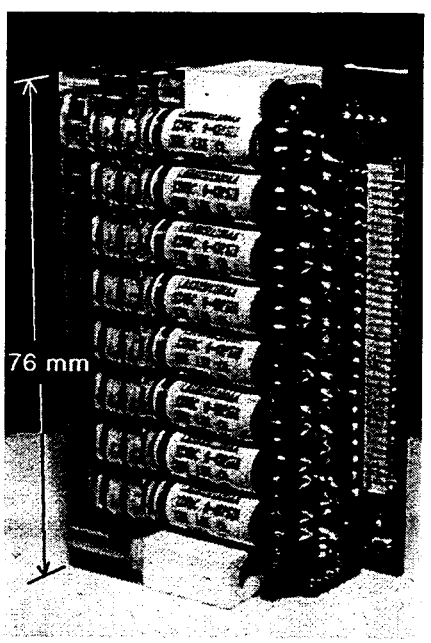


Fig. 7 Miniature 16 Valve Manifold

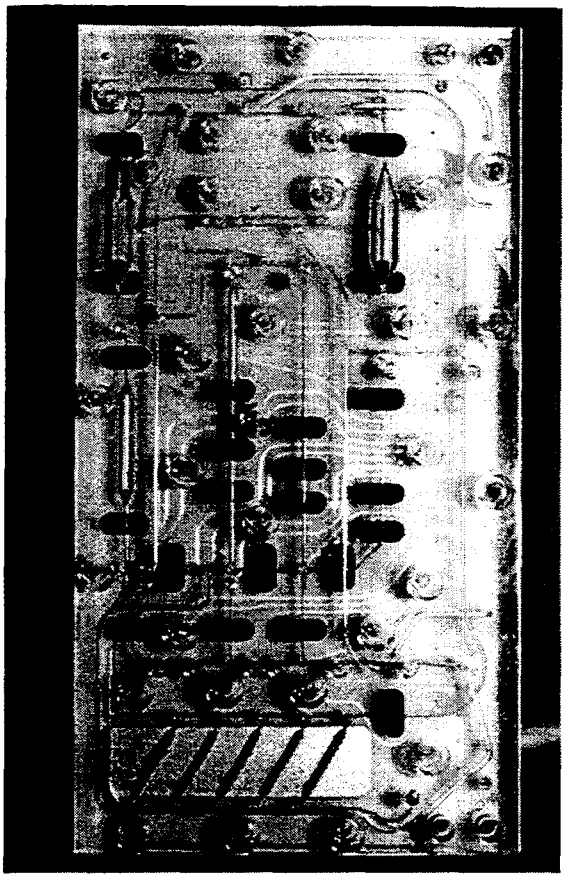


Fig. 8 Miniature Low Volume Clinical Chemistry Fluid Circuit System

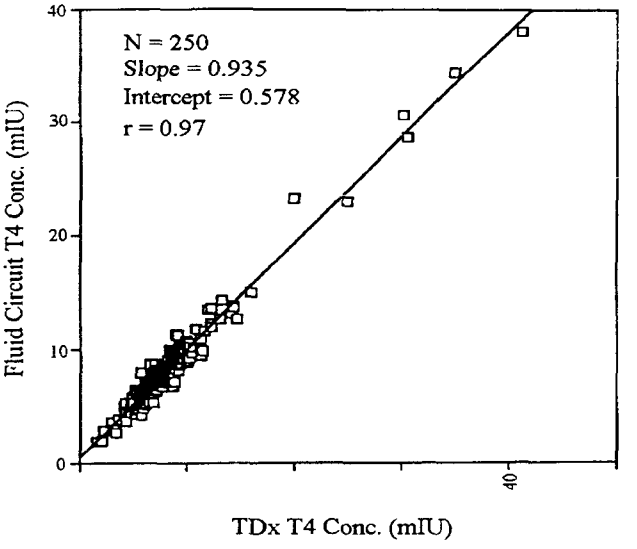


Fig. 9 Fluid Circuit vs. TDx Comparison

The first instrument system chosen for comparison was Abbott's TD$_x$™ instrument. The TD$_x$™ assay system includes a sensitive and completely automated diagnostic instrument, and more than 60 different assays for measuring the concentration of various therapeutic drugs. Each assay's sensitivity is specified in the ng/ml range, and their specificity has been determined against common known interferents. The instrument is a batch analyzer with a carousel capable of handling 20 patient samples at a time with a throughput of 60 - 80 tests per hour.

To develop a fluid circuit equivalent of this analyzer, the fluid metering/mixing, manipulation, incubation and detection steps for each of the TD$_x$™ assays were evaluated. The critical timing for each of these was assessed, and a sequence of 7 fluid processing loops were developed. Each of these could be performed simultaneously in a pipeline fashion in discrete fluid circuit modules. Four different module types were conceived, as well as a connection structure between modules, and a "pin-out" for the connection structure. This resulted in the design of 7 distinct modules that would; 1) acquire sample, 2) dilute the sample, 3) meter and mix a first reagent, 4) pre-incubate the mixture, 5) meter and mix two additional reagents, 6) incubate the final mixture, and 7) present the results to an in-line fluorescence polarimeter.

The fluid circuit that was fabricated is pictured in Fig. 10. Although it could have been designed and fabricated to occupy less space, the circuit is still only 58 cm × 18 cm × 5 cm. There are a total of 254 valves and 24 air/liquid sensors. When operated with a real time multitasking control system the cycle time on the longest fluid processing loop is ≤ 90 sec (45 tests/hr. throughput). This fluid circuit system reduces the volume of any single reagent or sample required to 5 μL (1/5 of TD$_x$™ volumes), requires no plastic disposables, and can continuously perform tests for as long as samples are presented and reagent is available.

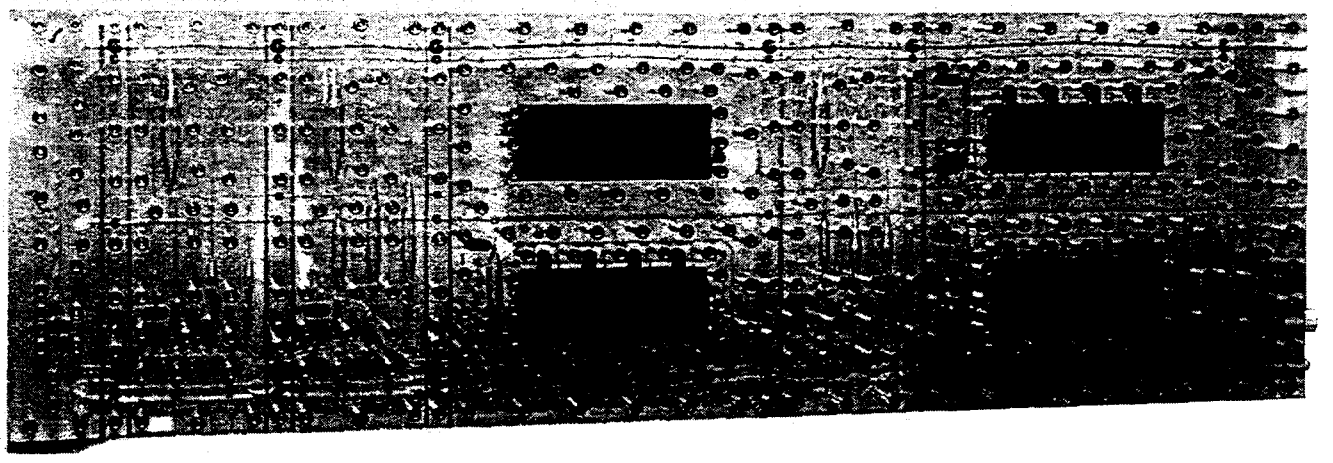


Fig. 10 TD$_x$™ Equivalent Fluid Circuit System

RESULTS

The results of comparing just one of the assays run on this fluid circuit system are illustrated in Fig. 9. This study took >250 random clinical patient samples and ran $TD_x^{™}$ T4 fluorescence polarization immunoassays in triplicate on both the $TD_x^{™}$ instrument, and the fluid circuit equivalent. These and other results proved that the fluid circuit system performed comparable to the $TD_x^{™}$ instrument. Similar results were obtained for THC, Total T3, Theophyline and Digoxin assays. Each of these tests stress a diagnostic analyzer in different ways: THC (Cannabis) stresses a fluid handling systems ability to prevent carryover; T3, T4, and Digoxin stress the system's sensitivity and specificity.

Although the fluid circuit system as fabricated and tested did not quite measure up to the $TD_x^{™}$ in throughput, the cause of this was identified as contention between the two final blocks over valves that were shared during certain protocols. If the final incubation module were redesigned to eliminate this resource contention, the resulting fluid circuit system would have a cycle time of <45 sec and a sustained 80 tests-per-hour throughput.

APPLICATION TO MEMS

There are two areas that Fluid Circuit Technology can help broaden the impact MEMS can have in fluidic applications.

First, it can act as a model for the design of fluidic MicroElectroMechanical Systems. Using this technology we have been able to quickly research new fluid circuit designs, develop fluid circuit design rules, and understand the underlying principles that affect scaling down fluid control networks. The manufacturing processes described above are quite similar to the processes for fabricating printed circuit boards and even to some of the processes for microfabricating MEMS. The materials are different, the scale is larger, and photolithography is absent, but the creation of three dimensional structure by bonding detail created on two dimensional planes (even the bonding process itself), is so similar that it enables the use of the same schematic capture and board layout CAE design tools used for designing printed circuits. Even the designs of some of the fluid circuit elements are similar. Pneumatically actuated valve designs similar to those we have described here have been microfabricated and remotely controlled [9]. These similarities should enable fluidic MEMS designers to first model their concepts on a macro-scale to visualize the fluid flow and develop an understanding of design issues, before they attempt to shrink their design into an opaque silicon device. Some of the subsystems already described here are ideal candidates for significant further miniaturization. The miniature 16 valve manifold, for instance, while it is already very small, could be integrated and miniaturized to micro-scale if rapid (5 ms) , low power (125 mW), inexpensive (<$20), and reliable pneumatic valves could be microfabricated.

Second, Fluid Circuit Technology can serve as an interconnect technology. It should be evident that what has been needed to expand the applicability of MEMS into medical uses has been the equivalent of a printed circuit board for fluidic engineers to enable bridging the micro to macro scales and to enable interconnection of complex fluid networks. Since acrylic/silicon bonding is feasible [10], and LIGA substrates can be solvent welded to cast acrylic [11], the packaging of microfabricated devices into fluid circuits should be relatively straight forward.

We believe it should be possible to utilize a fluid circuit as described here to directly mount, and provide fluidic interconnection for, microfluidic MEMS. An alternative method might be to develop the analog of a "chip-carrier" plastic fluid circuit package to which the microfabricated fluidic device could be bonded. The "chip carrier" could serve three purposes. It could provide; macro to micro fluidic interconnect and fluid access ports on one flat gasketed surface, a wire bonding lead frame for electronic connection and, mounting holes and rigidity so that it could be directly bolted to a fluid circuit.

CONCLUSIONS

The fundamental manufacturing processes, devices, subsystems, and systems that make up Fluid Circuit Technology have been described. Prototype fluid circuit systems have been fabricated to demonstrate the technology's applicability to many of the traditional diagnostic instrument markets, including clinical chemistry, therapeutic drug monitoring, immunochemistry, and hematology. Just as in the $TD_x^{™}$ example discussed, in each application it has been evident that this technology is capable of delivering reduction in size, manufacturing costs, consumption of reagent, and the need for disposables, while demonstrating comparable or improved performance and throughput. The precise manipulation of fluid volumes between 1 µl and 20 ml to accomplish dilution, metering, mixing, incubation, and detection operations have also been demonstrated.

By drawing parallels between fluid circuit bonding and silicon fusion bonding and between macro and micro pneumatically actuated membrane valves, it may be possible to utilize many of the fluid circuit subsystem designs to build analogous microfabricated fluid handling systems. The direct bonding of microfabricated elements into a fluid circuit system should also be possible, by capitalizing on previous work with bonding acrylic to common MEMS materials. We believe that the use of Fluid Circuit Technology to handle bulk fluid movement in a mixed macro- and micro-scale instrument as well as to provide fluidic access to microfabricated elements is the most feasible and attractive solution currently available and will be just as important to the success of fluidic MEMS as the PC board has been to integrated circuits.

REFERENCES

1. Schumann, E., *MEMS Industry Survey*, 1995, SEMI, Mountain View, CA.
2. Berins, M. L., Plastics Engineering Handbook, Society of the Plastics Industry, Van Nostrand Reinhold, New York, NY, 1991, pp. 693-750.
3. Brackett, T., "Method of Bonding Plastics", US Patent #4,875,956
4. Polycast Technology Corporation, Stamford, CT 06904-0141
5. Flexcon Company Inc., Spencer, MA 01562
6. Webster, M., "Valve with Flexible Sheet Member", US Patent #4848722, #4852851, #4858883
7. The Lee Company, Westbrook, CT, 06498-0424
8. VerLee, D., et.al., "Miniature Liquid/Air Interface Sensor", Proc. of O E M '95, Chicago, IL, SPIE Vol. 2622, 1995.
9. Vieider, C., et. al., "A Pneumatically Actuated Micro Valve with a Silicone Rubber Membrane for Integration with Fluid-Handling Systems," *Digest of Technical Papers The 8th International Conference on Solid-State Sensors and Actuators*, Stockholm, Sweden, June 25-29, 1995, vol. 2, pp. 284-285.
10. Brown, Chappell, "Common acrylic is ideal wafer bonder," *Electronic Engineering Times*, Nov. 14, 1994, Issue 823, p.43.
11. Guckel, H., et. al., "Laterally Driven Electromagnetic Actuators," *Tech. Dig. IEEE Solid-State Sensor and Actuator Workshop*, Hilton Head Island, SC, Jun 13-16, 1994, pp. 49-52.

MICROMACHINED CHEMICAL REACTORS FOR SURFACE CATALYZED OXIDATION REACTIONS

Ravi Srinivasan[†‡], I-Ming Hsing[†], J. Ryley[*], M.P. Harold[*], K.F. Jensen[†], and M. A. Schmidt[‡]

[‡]Microsystems Technology Laboratories, Massachusetts Institute of Technology, Cambridge, MA 02139.
[†]Department of Chemical Engineering, MIT.
[*]Dupont Central Research & Development, Experimental Station, Wilmington, DE 19880.

ABSTRACT

A micromachined chemical reactor (μ-reactor) for performing highly exothermic heterogeneous catalytic reactions has been fabricated, and successfully tested for two gas-phase model reaction systems- CO and NH_3 oxidations over Pt catalyst. The results demonstrate the feasibility of chemical production on a micro-scale, and suggest that μ-reactors have the potential of safer operation as compared to conventional reactors.

INTRODUCTION

Micromachined chemical systems have traditionally been fabricated for chemical analysis and sensing, not chemical production. Recent reports have outlined the potential advantages of micromachined chemical systems for chemical production [1,2], and have demonstrated feasibility of conducting reactions in externally heated microchannels [1]. One promising advantage of micromachined chemical reactors over conventional reactors is the increased safety for carrying out highly exothermic reactions that could lead to thermal runaway and explosions in conventional reactors. By carrying out the reaction in small physical dimensions, i.e. increasing the surface area to volume ratio of the reactor, heat can be removed very efficiently from the reactor, potentially preventing a thermal runaway. Moreover, the large surface area to volume ratios serves to inhibit undesired gas phase chemistry by thermal or chemical quenching of free radicals, thus preventing a gas-phase chain branching type of runaway reaction.

The aim of this research is to understand and to explore the fundamental issues in carrying out a particular class of exothermic reactions on a micro-scale, namely catalytic partial oxidation reactions. Partial oxidation reactions are widely used in the chemical industry to produce a large number of useful chemical intermediates such as alcohols, aldehydes, cyanides, and nitriles [3]. These partial oxidation reactions are typically extremely fast and highly exothermic (heats of reaction on the order of 100kJ/mol). These characteristics make the control of conventional 'macroscopic' partial oxidation reactors difficult. Slight changes in reactor parameters such as the inlet temperature can cause a runaway reaction, potentially leading to an explosion. Thus, these reactions are ideal candidates for exploring potential safety advantages of micromachined reactors.

Commercially, μ-reactors could be used for point of use production of toxic or unstable chemicals via partial oxidation reactions. This would eliminate the need for storage and transportation of these hazardous and unstable chemicals [1]. Since some partial oxidation reactions are very fast (reaction time < 1ms), the throughput of one μ-reactor may be sufficiently large that a modest number of μ-reactors operating in parallel could satisfy the requisite production rate for point of use applications. In this paper we will describe our prototype μ-reactor, the test set-up, and initial test results for two model catalytic oxidation reactions.

DEVICE DESCRIPTION

The μ-reactor is a T-shaped channel for gas flow fabricated by bulk micromachining of a silicon wafer (figs.1&2). The channel is sealed from the top by a low stress silicon nitride membrane (fig.1), and is sealed at the bottom by an aluminum base-plate with three gas inlet-outlet holes that is epoxied to the chip (fig. 3). Reactant gases enter the μ-reactor at the two ends of the horizontal flow sensor region, where the gas flow rate is determined using hot film anemometer in either a time of flight or anemometer mode. The reactant gases diffusively mix at the entrance to the reaction zone which is the vertical section of the T-shaped channel. In the reaction zone, the gases react on the catalyst that is coated on the silicon nitride membrane, and the products exit the μ-reactor at the vertical end of the T.

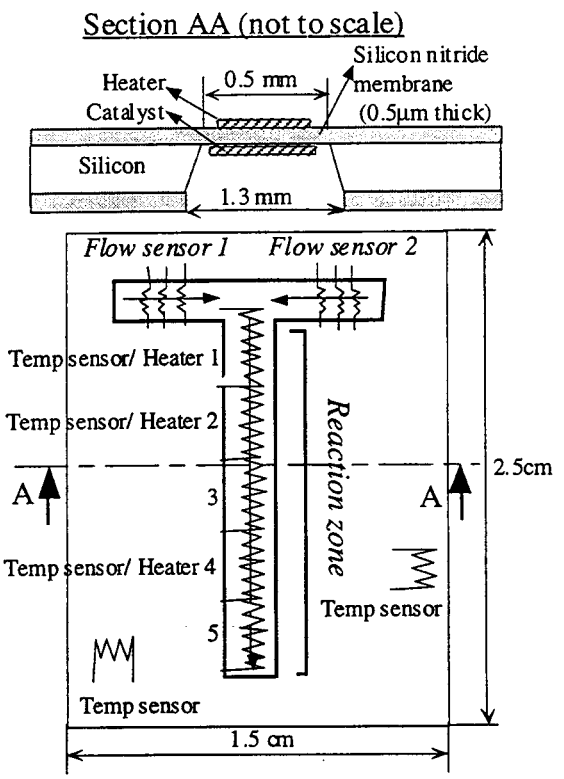


Fig.1 *Schematic top-view of μ-reactor*

0-9640024-1-8 $3.00 © 1996 TRF

Solid-State Sensor and Actuator Workshop
Hilton Head, South Carolina, June 2-6, 1996

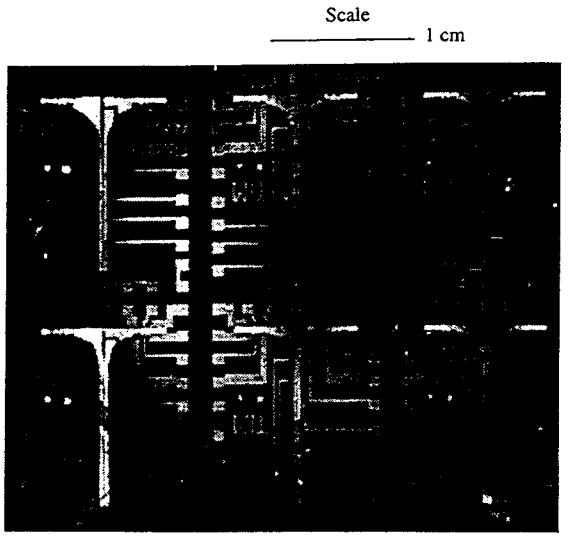


Fig.2 *Photograph of part of wafer showing six devices*

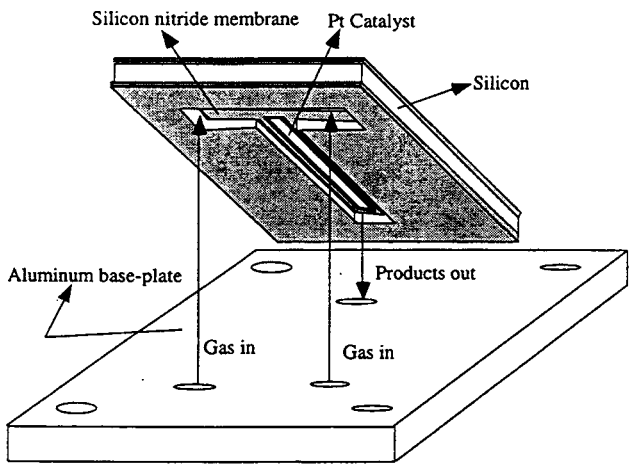


Fig. 3 *Exploded view of μ-reactor package*

The top surface of the silicon nitride membrane has five segmented Pt resistor lines, which serve both to heat the catalyst, and to measure its temperature. Heat transfer simulations indicate negligible temperature difference between heater segments on the top surface, and the catalyst on the bottom surface of the membrane. The temperature is determined by measuring the incremental electrical resistance of the heaters, and then comparing the measured value with previously generated calibration curves of incremental resistance with temperature.

FABRICATION PROCESS

The starting material for the μ-reactor fabrication (fig. 4) was a double side polished (100mm diameter , 550μm thick) silicon wafer LPCVD coated with 0.5μm thick low-stress silicon nitride. The backside of the wafer was patterned and plasma etched to expose the silicon in the T-shaped channel region. The front side of the wafer was then patterned for metal lift-off using infra-red (IR) alignment. E-beam evaporation of 0.1μm platinum (on a 10nm Ti adhesion layer)

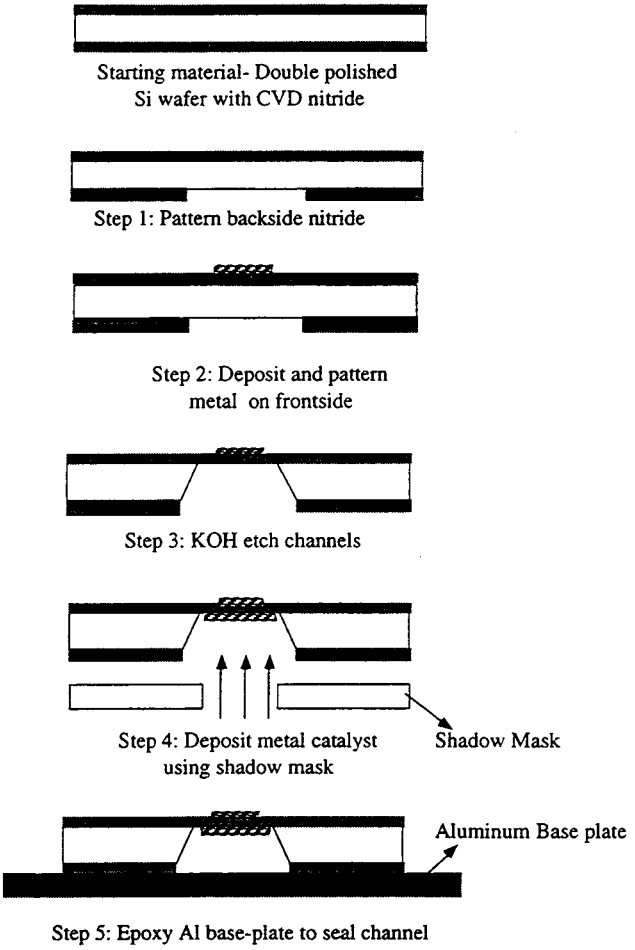


Fig. 4 *Fabrication Process Flow*

for metal heaters, and temperature and flow sensors was carried out. In the next step, the backside of the wafer was exposed to KOH, etching the bulk of silicon, stopping on the nitride membrane. Subsequently, 0.1μm thick platinum catalyst (on 10nm Ti) was e-beam deposited on the silicon nitride membrane in the reactor region via a shadow mask. In the final step, an aluminum base-plate with three gas inlet-outlet holes was epoxied to the chip (figs.3&4).

TEST SETUP

The test-bed for the μ-reactor is shown in fig.5. O_2, NH_3 and CO were transported via external mass flow controllers, which maintain the desired gas flow rates, into a mechanical package (gas input-output chuck). The gas input-output chuck was used to transfer gases from the external 1/16" piping into the μ-reactor. The gas inlet-outlet holes in the μ-reactor base-plate were aligned with the gas inlet-outlet holes in the chuck, and the base-plate was sealed by pressurized contact with Viton® o-rings on the top face of the chuck.

A probe card was used to electrically access the heater and temperature sensor bond pads . The product gas stream exited the μ-reactor via the gas input-output chuck outlet into the exhaust piping. The exhaust gas concentrations were continuously monitored using a quadrupole mass spectrometer. The μ-reactor assembly was placed under a microscope for imaging.

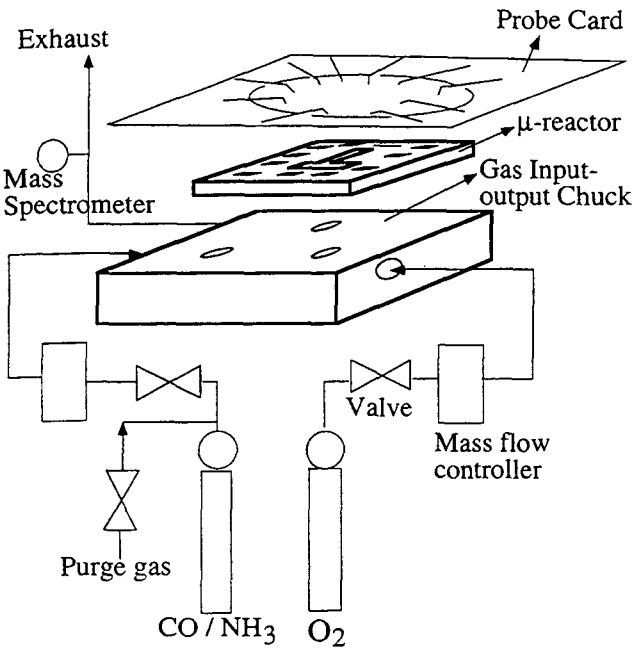


Fig. 5 *Schematic diagram of μ-reactor test-bed*

DEVICE MODELING

We have developed two and three dimensional Galerkin Finite Element (FEM) simulation tools for evaluating and comparing μ-reactor performance with conventional reactors based on similar modeling tools for microelectronics processing[6]. A mesh generator is used to translate reactor geometries into FEM meshes. As an example of the three dimensional fluid flow and heat transfer simulations, fig.6 shows the temperature profile on the silicon nitride membrane, and fig.7 displays the corresponding temperatures in the gas below the heater in the reaction zone. The very good thermal isolation provided by the silicon nitride membrane results in a very narrow hot zone, while the bulk of the μ-reactor chip is at room temperature.

Detailed chemical mechanisms have been coupled with fluid flow, heat and mass transfer FEM simulations to evaluate the chemical behavior of μ-reactors, specifically their ability to prevent runaway reactions and their product yield and selectivity. The computations show that the heat transfer to the bulk Si wafer and wall quenching reactions have the desired effect of eliminating thermal runaway.

TEST RESULTS

The μ-reactor operation has been successfully tested for two catalytic oxidation reactions: 1) Pt catalyzed oxidation of NH_3,

$$NH_3 + \frac{5}{4}O_2 \xrightarrow{Pt} NO + \frac{3}{2}H_2O + 226kJ$$

and 2) Pt catalyzed oxidation of CO,

$$CO + \frac{1}{2}O_2 \xrightarrow{Pt} CO_2 + 284kJ\,.$$

The mass spectrometer data of the normalized NH_3 peak, in the case of the NH_3 oxidation reaction is shown in fig.8. The 2nd heater segment of the μ-reactor (fig.1) was gradually heated under gas inlet conditions summarized in fig.8. The reaction ignited at a catalyst temperature of 230°C, as

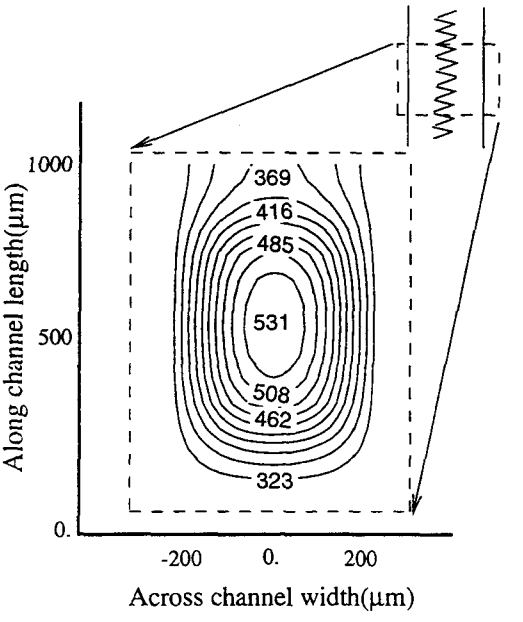


Fig. 6. *Temperature contours on SiN membrane. The figure inset shows the active heater segment on the SiN membrane that was used in the simulation.*

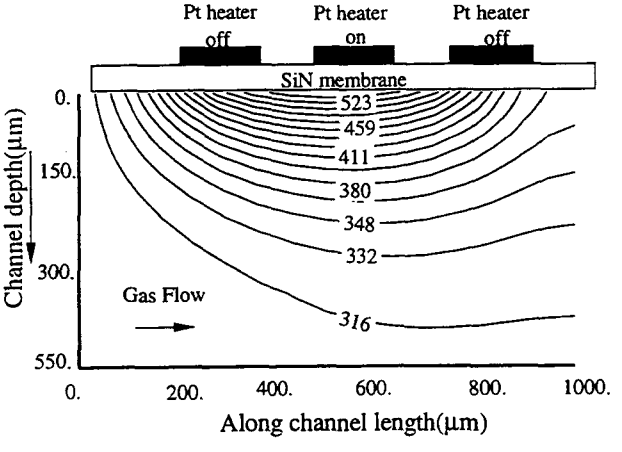


Fig. 7 *Temperature profile along reactor length. The gas flow velocity is 26 cm/s.*

indicated by the drop in the NH_3 peak intensity. Once ignited, the power to the heater was kept constant, but the catalyst temperature rapidly climbed to a steady state value of 680°C due to the heat released by the reaction. The NO peak in the mass spectrometer signal rose while the N_2, N_2O and NO_2 peaks were absent, indicating that there were no side reactions. For this reaction, the catalyst activity remained stable over several hours of operation.

Mass spectrometer data of the exhaust gas stream in the CO oxidation reaction is shown in fig.9. The reaction ignited at a catalyst temperature of 400°C (at t=150s in fig.7), as indicated by the drop in CO concentration and rise in the CO_2 concentration in the product gas stream. The catalyst temperature rapidly climbed to 800°C as a result of the heat

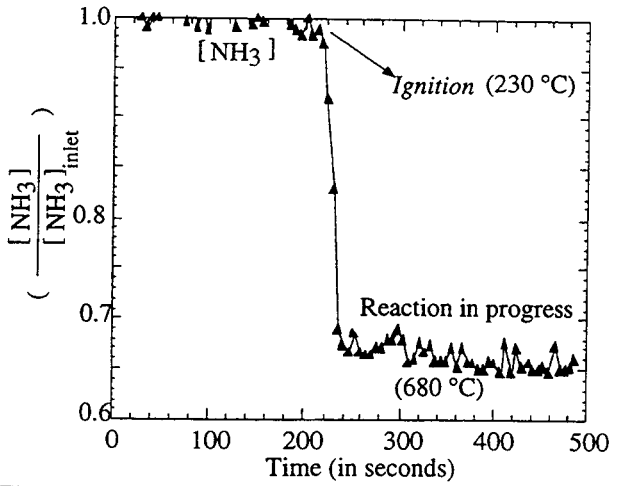


Fig. 8 *Normalized concentrations of μ-reactor exhaust during NH_3 oxidation. Reaction conditions are- Pressure = 1atm, inlet O_2 flowrate = 10sccm, inlet NH_3 flowrate = 1sccm, Power supplied to heater at ignition = 124.5mW.*

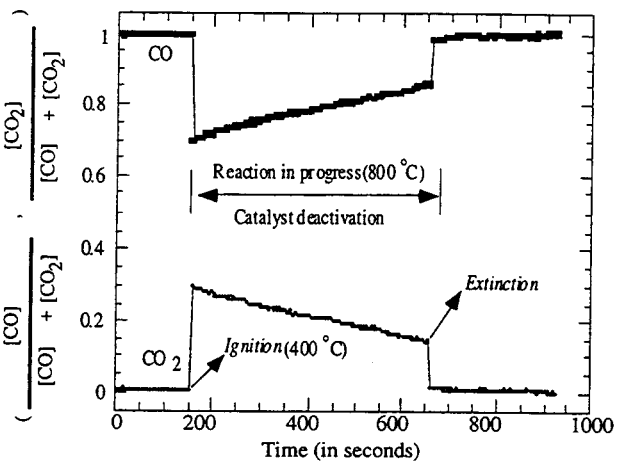


Fig. 9 *Normalized concentrations of μ-reactor exhaust during CO oxidation. Reaction conditions are- Pressure = 1atm, inlet O_2 flowrate = 20sccm, inlet CO flowrate = 2sccm, Power supplied to heater at ignition = 340mW.*

released in the reaction. This ignition behavior is consistent with CO oxidation literature[4]. As the reaction progressed, the catalyst deactivated leading to a steady drop in the reaction rate, and eventual extinction of the reaction. Ex-situ Auger Electron Spectroscopy of the catalyst surface revealed carbon and nickel deposits as possible causes for the catalyst deactivation.

Several observations may be made when comparing the reported behavior of μ-reactor test data with macroscopic laboratory reactors operating with the same inlet gas compositions. The ignition behavior is the same, but the reaction and extinction behavior in the μ-reactor deviated from conventional laboratory reactors. When catalytic NH_3 oxidation is carried out on Pt foils suspended in a 3.2cm diameter pyrex tube, under the same inlet gas conditions as the present study, the reaction is autothermal [5], i.e. it does not require external heat input to sustain the reaction and can only be shut off by stopping the flow of reactants. In contrast, in the μ-reactor, the reaction extinguishes at heater powers below 118mW. Thus, in the event of an emergency, the μ-reactor can be shut-off immediately by turning off the heaters, leading to safer operation.

CONCLUSION

A promising application for micromachined chemical reactors is the processing of potentially explosive and hazardous reactions. A μ-reactor for carrying out an important class of highly exothermic hydrocarbon partial oxidations, has been fabricated, and successfully tested for two model reaction system. Preliminary results demonstrate the feasibility of conducting catalytic reactions on a micro-scale, and more importantly, indicate a potential for safer operation of μ-reactors over conventional reactors.

ACKNOWLEDGMENTS

We thank Dupont and the NSF PYI award program (Martin Schmidt) for funding this work. We also gratefully acknowledge the help of MTL staff, Dupont staff (Mark Wetzel and Jim Ashmead), and students of MAS and KFJ groups at MIT.

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," To be published in the proceedings of the workshop on, *Microsystem Technology for Chemical and Biological Microreactors*, Max-Planck-Institut, Mainz, Germany, 1995.

[2] H. Mobius, W. Ehrfeld, V. Hessel, Th. Richter, "Sensor controlled processes in chemical microreactors," *Proceedings of The 8th International Conference on Solid-State Sensors and Actuators*, 775 (1995).

[3] C.N. Satterfield, "Heterogeneous Catalysis in Industrial Practice," 2nd Edition, McGraw-Hill Inc., New York (1991).

[4] L.F. Razón and R.A. Schmitz, "Intrinsically unstable behavior during the oxidation of carbon monoxide on platinum," *Cat. Rev. Sci. Engng, 28*, 89 (1986).

[5] W.R. Williams, M.T. Stenzel, X. Song, and L.D. Schmidt, "Bifurcation Behavior in Homogeneous-Heterogeneous Combustion: I. Experimental Results Over Platinum," *Combustion and Flame, 84*, 277 (1991).

[6] K.F. Jensen, T.P. Merchant, J.V. Cole, J.P. Hebb, K.L. Knutson, and T.G. Mihopoulos, "Modelling Strategies for Rapid Thermal Processing", in Advances in Rapid Thermal and Integrated Processing, (Ed. F. Roozeboom), Kluwer, The Netherlands (1996).

Versatile Materials for use as Chemically Sensitive Interfaces in SAW-based Sensor Arrays

Richard M. Crooks, David E. Bergbreiter, Merlin L. Bruening, Mona Wells, Yuefen Zhou
Department of Chemistry, Texas A&M University,
College Station, TX 77843-3255.

Antonio J. Ricco, Gordon C. Osbourn
Microsensor R&D and Vision Science Departments, Sandia National Laboratories
Albuquerque, NM 87185-1425

ABSTRACT

The primary research objective of the work described here is to design, synthesize, and characterize new materials for use as chemical sensor interfaces, integrate these materials, using appropriate transducers, into sensor arrays, and then develop appropriate mathematical algorithms for interpreting the array response.

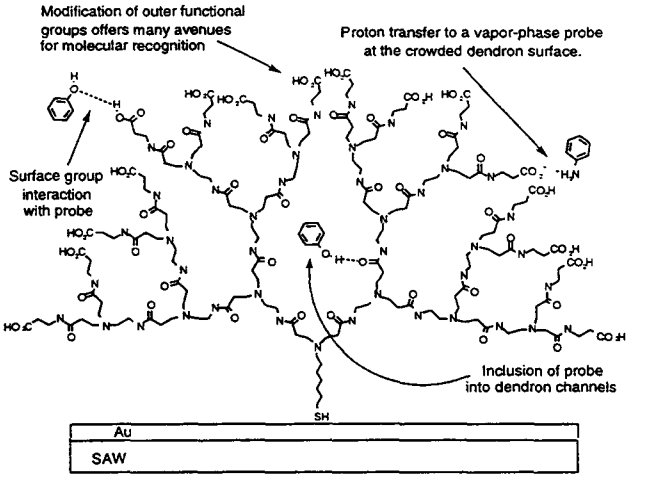
In this paper, we will discuss two new types of materials we have developed that are ideally suited for use as chemically sensitive interfaces for array-based chemical sensing applications, since they: (1) provide general specificity towards classes of functional groups rather than individual compounds; (2) are intermediate in structure between monolayers and polymers; (3) exhibit both endo- and exo-recognition. The first class of materials is surface-confined dendrimers and the second is hyperbranched polymers.

INTRODUCTION

Dendrimer thin films represent a new type of interfacial architecture that resides on the border between self-assembled monolayers (SAMs) and polymer films [1]. The response to vapor-phase dosing experiments we describe here demonstrates how this architecture is particularly well suited to chemical sensing because, like polymers, it displays the sensitivity of thicker films, but, like monolayers, it does not suffer from permeation transients. In addition, we show that the response of these surfaces to vapor-phase probe molecules is dependent in a chemically logical way on both the chemical nature of the dendrimer-probe interaction and the dendrimer size.

Dendrimers are a class of hyperbranched polymer that have regularly repeating branch sequences emanating from a central core that terminate in a large number of surface-end groups [2,3]. These supramolecular species offer a unique and unexplored venue for the creation of chemically-sensitive interfaces that act through dual-mode reception; voids within the dendrimer superstructure serve as endoreceptors and the terminal functional groups of dendrimers serve as exoreceptors (Scheme 1). Importantly, the chemical functionalities of dendrimer exoreceptors and endoreceptors are both subject to rational synthetic control. In spite of the exceptional promise of dendrimers as receptors in thin films, there are no prior published accounts describing the covalent surface-confinement of dendrimers or their integration into chemical sensors, and only three accounts that deal with the formation of organic dendrimer films [4-6].

In addition to dendrimer monolayers, we have also prepared thin films of new highly branched polymers, which we term 'hyperbranched polymers,' that are relevant to boundary layer phenomena such as adhesion, wetting, and chemical sensing [7]. Specifically, we synthesized surface-grafted, hyperbranched poly(acrylic acid) (PAA) films on self-assembled organomercaptan monolayers. These films grow in discrete steps, but unlike most layer-by-layer approaches, film thickness increases nonlinearly as a function of the number of layers because of the branched polymeric architecture. Moreover, these polymer films contain a high density of carboxylic acid groups, which can act as receptor sites themselves or serve as reactive sites for subsequent derivatization.



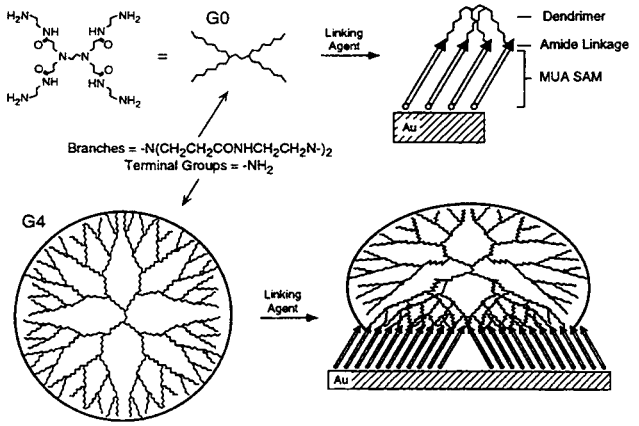
Scheme 1

RESULTS AND DISCUSSION

Scheme 2 illustrates our procedure for attaching poly(amidoamine) (PAMAM) dendrimers to a mercaptoundecanoic acid (MUA) SAM using an asymmetric anhydride intermediate. This intermediate, formed by reacting the MUA SAM with ethyl chloroformate, is highly active and rapidly condenses with the amine-terminated dendrimers to form amide linkages. We characterized the dendrimer surfaces, each of which incorporated one of five generations of PAMAM dendrimers (G0, MW=517; G2, MW=3256; G4, MW=14,215; G6, MW=58,048; G8, MW=233,383), by Fourier transform infrared-external reflectance spectroscopy (FTIR-ERS) and ellipsometry.

To confirm the attachment of dendrimers to MUA SAMs, we obtained FTIR-ERS spectra of G0-, G2-, G4-, G6-, and G8-modified surfaces. Figure 1 shows the amide and carbonyl region of these spectra. There are two important points. First, we do not observe residual IR absorbance from the MUA carbonyls (C=O), which would appear between 1720 and 1740 cm^{-1}, after linking dendrimers to the SAMs. This indicates complete or nearly complete reaction between all five dendrimers and the MUA surface (from control experiments we estimate that residual carboxyl group sites in excess of 10%

0-9640024-1-8 $3.00 © 1996 TRF



Scheme 2

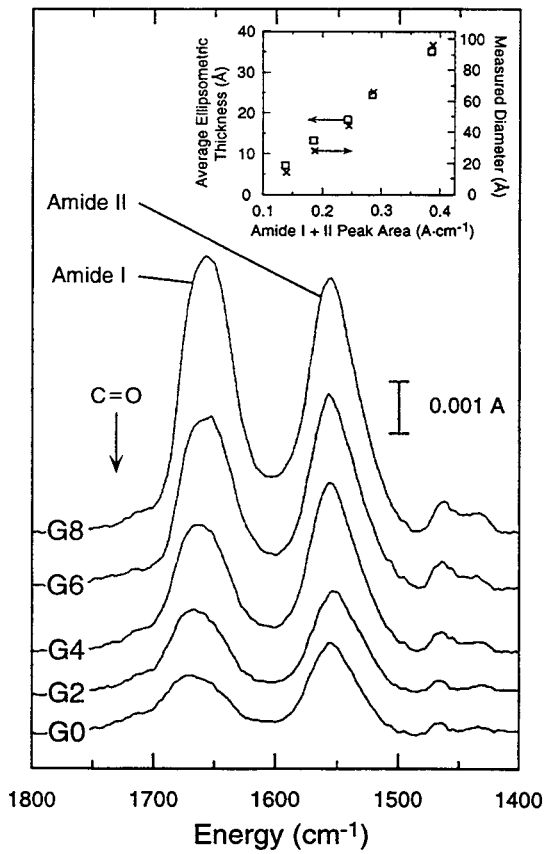


Figure 1. FTIR-ERS and ellipsometric results from analysis of five generations of dendrimer films. The infrared spectra show the quantitative attachment of dendrimers to SAM active sites, via an absence of residual carbonyls, and concomitant increase in dendrimer amide bands and generation. The inset illustrates how both the average ellipsometric film thickness and dendrimer diameter scale linearly with the total amide peak area. Taken together these data show that, for a constant two-dimensional dendrimer coverage, three-dimensional growth is dependent upon dendrimer generation. Data in the inset represents a mean of five or more separate experiments.

would be discernible). Second, amides are the principal infrared-active functionality of PAMAM dendrimers, and the total area of amide I and II bands increases with increasing dendrimer generation. Since the two-dimensional projection of the dendrimer overlayer onto the surface is the same for all five generations, this result demonstrates the anticipated three-dimensional film structure.

The inset of Figure 1 shows that there is a correlation between the diameters of the bulk-phase dendrimers and the thickness of the dendrimer monolayers since both the dendrimer diameter, measured by size-exclusion chromatography, and the measured ellipsometric thicknesses of the dendrimer monolayers scale linearly with the amide I and II peak area. The FTIR-ERS and ellipsometric results attest to the approximately monolayer-level dendrimer surface coverage and three-dimensional film growth that scales with dendrimer generation.

For the reasons discussed earlier, dendrimers of different size and chemical composition are ideally suited for integration into array-based chemical sensors [8], and therefore we monitored the response of these surfaces to vapor-phase probe molecules. In the first experiment of this type, we used methyl acrylate to probe the primary amine functionalities at the dendrimer terminal arms. Michael addition of methyl acrylate to primary amines is known to target only the primary amines of the PAMAM structure. The resulting product is a methyl ester whose signature is evident in the carbonyl region of the infrared spectrum (1720-1740 cm^{-1}). Figure 2 shows the relationship between surface reactivity, gauged by the area of the carbonyl band, and the square of the dendrimer radius (r_d^2).

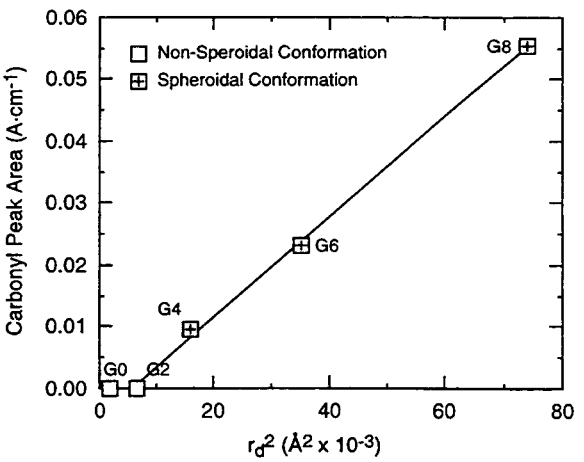


Figure 2. The relation between r_d^2 and the reactivity (indicated by carbonyl peak area) of dendrimer film exoreceptors with vapor-phase methyl acrylate. Nonspherical dendrimers (G0 and G2) react completely with MUA SAMs at the time of attachment and produce dendrimer films that are unreactive to methyl acrylate. Spheroidal dendrimers (G4-G8) retain a globular structure subsequent to surface attachment and hence, for the constant two-dimensional dendrimer coverage observed in these experiments, have reactivity with methyl acrylate that is directly related to the amount of three-dimensional surface area of the dendrimer film.

These data illustrate two important points. First, below the threshold at which PAMAM dendrimers are known to adopt a globular geometry (G0 and G2), the dendrimer films show no reactivity. Since G0 and G2 dendrimers are geometrically malleable, all or most of the outer functional groups react with the MUA surface during attachment, and therefore there are no amines left to react with methyl acrylate. A second notable feature of Figure 2 is that at or above the threshold at which PAMAM dendrimers adopt a globular geometry (G4-G8), the number of dendrimer-film reactive sites increases linearly with r_d^2 ($r_d^2 \propto$ exoreceptive surface area of a spheroid). This is a result of the spheroidal geometry of G4-G8 surface-confined dendrimers, which prevents all the terminal functional groups

from reacting completely with the MUA surface (Scheme 2). Thus increased reactivity with vapor-phase methyl acrylate is a direct consequence of the enhanced number and accessibility of exoreceptors in dendrimer films with three-dimensional film structure.

We conducted a second experiment to investigate the response of dendrimer-modified surfaces to unreactive vapor-phase probes. In this experiment we sequentially dosed dendrimer-modified surface acoustic wave (SAW) devices with several molecular probes of varying chemical nature. To examine functional group specificity, we chose benzene, 3-methyl hexane and *n*-heptane as hydrophobic probes, methyl propanol and butanol as hydrogen-bonding probes, and propionic acid, which is a hydrogen bonding probe capable of proton transfer.

SAW devices are best known for their ability to function as highly sensitive mass balances capable of quantitatively detecting as little as 100 pg/cm^2, and as excellent platforms for chemical microsensor systems [**8**]. In this experiment, we first modify the Au-coated active area of a SAW device with a dendrimer film, and then load the device into a flow cell for vapor phase dosing [**9**]. Figure 3a is an example of unprocessed data from a typical SAW experiment. It illustrates how the dendrimer-modified device response in these experiments possesses three of the essential attributes of an ideal chemical sensor: 1) the response to dosants in these experiments is rapid with no detectable permeation transient, 2) the signal-to-noise ratio of the data is excellent, and 3) the response is typically completely reversible [**10**].

Figure 3b summarizes the results from vapor-phase dosing experiments of dendrimer-modified surfaces carried out as described for Figure 3a. We dosed a total of three independently prepared devices for each dendrimer generation, and each device was sequentially dosed three or more times with each dosant. All dendrimer-modified devices are most responsive to propionic acid, to which, as with methyl acrylate, the response of G4-, G6-, and G8-modified devices is much greater than that of G0- and G2-modified devices. The response of the G4-modified device to propionic acid is statistically differentiable from the responses of both the G0- and G2-modified devices. However, unlike methyl acrylate, the response to propionic acid does not increase with increasing dendrimer size. This is an interesting point inasmuch as the methyl acrylate probe targets the dendrimer outer functional groups (exoreceptors) *only*, whereas propionic acid may probe dendrimer endoreceptors as well. Partial or complete intercalation of probes into the dendrimer structure (endoreception) generates a mass loading that would not yield the linear response shown in Figure 2. Although the void volume, and hence endoreceptive capacity of dendrimers, increases greatly with increasing generation, the larger dendrimers have a very high density of outer surface groups that may sterically obstruct probe penetration into interior dendrimer voids. In this context, we note that the G4-modified surface is simultaneously the most responsive surface to all probes used in the SAW device study as well as the least constricted of the spheroidal dendrimers that we investigated. That is, while the surface area of G4 is the smallest of the spheroidal dendrimers, its interior is most accessible.

Another key point about the SAW device responses in Figure 3b is that the response to probes decreases in the order acid > alcohols > hydrophobic dosants. This response order is more pronounced for the G4-G8 modified surfaces and is dictated by the PAMAM structure which possesses hydrogen-bonding exoreceptors and endoreceptors. G0 and G2 dendrimer films are not as effective at discriminating between the three different classes of probes since these surfaces have few or no free amine terminal groups, and no coherent endoreceptive ability.

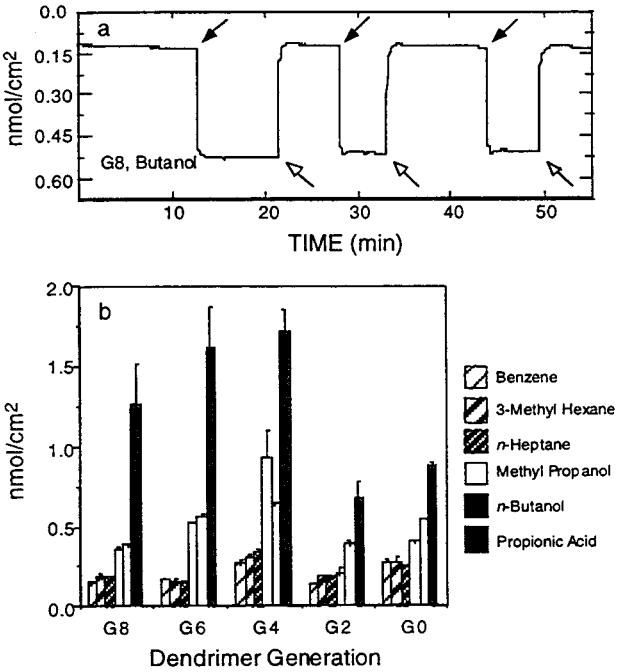


Figure 3. a) Unprocessed response from a G8-modified SAW device dosed with vapor-phase butanol (25 % of saturation in N_2). This trace shows how a typical response of dendrimer modified devices is instantaneous, stable, and reversible. The solid arrows indicate butanol dosing and the closed arrows indicate purging with pure N_2. b) Results from dosing dendrimer-modified SAW devices with six VOCs. The response order acid > alcohols > hydrophobic dosants is a result of the high number of polar receptor sites associated with the PAMAM structure. Higher generation dendrimers (G4-G8) discriminate the chemical nature of the probe more effectively due to the three-dimensional structure of these films, however the magnitude of response is not directly related to dendrimer size since the dosants in this experiment probe the dendrimers' dual-mode receptive capacity.

As with the dendrimer monolayers, synthesis of hyperbranched polymer films (Figure 4) begins with a mercapto-undecanoic acid (MUA) SAM. The acid groups react with an α,ω-diamino-terminated poly(*tert*-butyl acrylate) (H_2NR-PTBA-RNH_2) to yield a grafted polymer layer. Hydrolysis then forms a grafted PAA layer. Repetition of these steps produces additional grafting at multiple sites on each prior graft leading to a layered, hyperbranched polymer film.

This procedure has several desirable attributes. First, it compensates for inefficiencies in reactions at surfaces. Because there are many grafting sites on each polymer chain, this method produces a thick polymer layer even if reactions proceed in relatively poor yield. Second, the grafted polymer films contain a high density of reactive functional groups that are suitable for further elaboration. Finally, because branching infers that each subsequent layer contains more polymer chains, this method leads to increasingly thicker and more tightly packed polymer layers.

Because of the high density of functional groups, PAA films can serve as specific metal-ion binders . After exposure to an equimolar ethanolic solution of $Fe(ClO_4)_3$ and $Ni(ClO_4)_2$, the intensity of the acid carbonyl FTIR absorption (1731 cm^{-1}) decreased by >80% and new peaks corresponding to the symmetric and asymmetric stretches of carboxylate appeared (1582 cm^{-1} and 1440 cm^{-1}) (Figure 5). These changes demons-

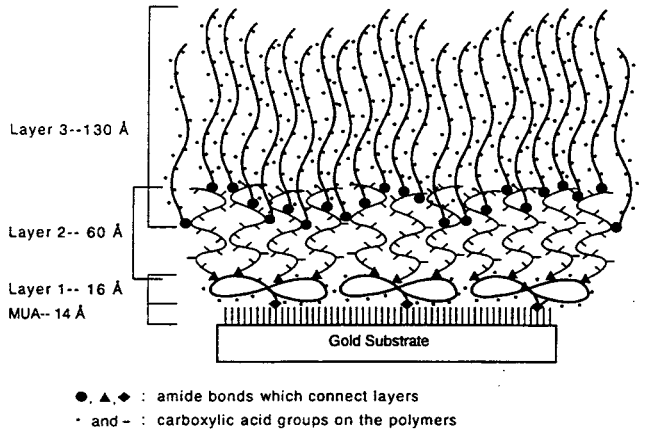


Figure 4. Illustration of a three-layer hyperbranched polymer attached to a mercaptoundecanoic acid (MUA) self-assembled monolayer. The layers are composed of poly(acrylic acid) polymer chains covalently linked to one another through terminal amide bonds. The asterisks denote carboxylic acid receptor groups.

strate metal complexation by the carboxylate receptors. In this competitive binding experiment, the films complex high levels of Fe^{3+} (XPS: O / Fe = 6, consistent with each Fe^{3+} coordinated to three carboxylates), but no detectable Ni^{2+}. This result is in accord with the formation constants of Fe^{3+}- and Ni^{2+}-carboxylate complexes.

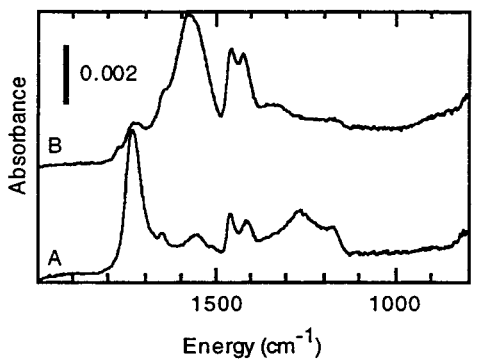


***Figure 5.** FTIR-ERS spectra of (A) 1 layer of PAA grafted on MUA and (B) the same sample after immersion in an equimolar ethanolic solution of $Fe(ClO_4)_3$ and $Ni(ClO_4)_2$ (0.006 M) followed by an ethanol rinse.*

SUMMARY AND CONCLUSIONS

In summary, we have demonstrated for the first time that monolayers of dendrimers can be covalently linked to surfaces, and we have shown that films incorporating higher generation dendrimers, that is, those possessing characteristic exo- and endo-structure, show superior response (methyl acrylate, propionic acid) and selectivity (propionic acid > alcohols > hydrophobic dosants) to vapor-phase probes than films incorporating non-spheroidal molecules (G0 and G2 PAMAM dendrimers) lacking endo- and exo-structure.

Hyperbranched PAA films can also be easily synthesized on a MUA SAM via sequential grafting with polymeric building blocks. The film thickness increases rapidly as a function of the number of grafting steps. In addition, these films contain a high density of carboxylic acid groups that can selectively bind metal ions or be chemically modified. Hyperbranched polymer films thus provide new platforms for chemical sensing applications and for tailoring polymer surface properties for a wide variety of technological applications.

We are grateful for financial support of this research from the National Science Foundation (DEB: DMR-9308414 and RMC: CHE-9313441), the U.S. DOE under contract DE-AC04-94AL-85000, and the state of Texas (Texas Higher Education Coordinating Board) through the Advanced Technologies Program. MLB gratefully acknowledge an NIH postdoctoral fellowship and ML acknowledges support through an EPA graduate fellowship.

REFERENCES AND NOTES

1. M. Wells, R. M. Crooks, "Preparation and Properties of Chemically Sensitive Dendrimer Surfaces", *J. Am. Chem. Soc.*, in press.

2. D. A. Tomalia, A. M. Naylor, W. A. Goddard III, "Starburst dendrimers: control of size, shape, surface chemistry, topology and flexibility in the conversion of atoms to macroscopic materials", *Angew. Chem., 102*, 119 (1990).

3. C. J. Hawker, J. M. J. Fréchet, "Preparation of Polymers with Controlled Molecular Architecture. A New Convergent Approach to Dendritic Macromolecules", *J. Am. Chem. Soc., 112*, 7638 (1990).

4. S. Watanabe, S. L. Regen, "Dendrimers as Building Blocks for Multilayer Construction", *J. Am. Chem. Soc., 116*, 8855 (1994).

5. B. Alonso, M. Morán, C. M. Casado, F. Lobete, J. Losada, I. Cuadrado, "Electrodes Modified with Electroactive Films of Organometallic Dendrimers", *Chem. Mater., 7*, 1440 (1995).

6. P. M. Saville, P. A. Reynolds, J. W. White, C. J. Hawker, J. M. J. Fréchet, K. L. Wooley, J. Penfold, J. R. P. Webster, "Neutron Reflectivity and Structure of Polyether Dendrimers as Langmuir Films", *J. Phys. Chem., 99*, 8283 (1995).

7. Y. Zhou, M. L. Bruening, D. E. Bergbreiter, R. M. Crooks, M. Wells, "Preparation and Properties of Hyperbranched Polymer Films Grafted on Self-Assembled Monolayers", *J. Am. Chem. Soc.*, in press.

8. A. J. Ricco, "SAW Chemical Sensors", *Electrochem. Soc. Interface, 3*, 38 (1994).

9. H. C. Yang, D. L. Dermody, C. Xu, A. J. Ricco, R. M. Crooks, "Molecular Interactions Between Organized, Surface-Confined Monolayers and Vapor-Phase Probe Molecules. 8. Reactions Between Acid-Terminated Self-Assembled Monolayers and Vapor-Phase Bases", *Langmuir, 12*, 726 (1996).

10. In some cases, particularly for higher generation dendrimers, we observe that up to 30% of the total mass loading for propionic acid is bound irreversibly. This is likely due to proton transfer from propionic acid to the dendrimer primary amine terminal groups.

SELECTIVE, PULSED CVD OF PLATINUM ON MICROFILAMENT GAS SENSORS

Ronald P. Manginell, James H. Smith, Antonio J. Ricco, Daniel J. Moreno, and Robert C. Hughes
Sandia National Laboratories
Albuquerque, NM 87185-1080

Robert J. Huber
University of Utah
Salt Lake City, UT 84112

Stephen D. Senturia
Massachusetts Institute of Technology
Cambridge, MA 02139

ABSTRACT

A post-processing, selective micro-chemical vapor deposition ("micro-CVD") technology for the deposition of catalytic films on surface-micromachined, nitride-passivated polysilicon filaments has been investigated. Atmospheric pressure deposition of platinum on microfilaments was accomplished by thermal decomposition of platinum acetylacetonate, an organometallic Pt precursor; deposition occurs selectively only on those filaments which are electrically heated. Catalyst morphology, characterized by scanning electron microscopy, can be controlled by altering deposition time, filament temperature, and through the use of pulsed heating of the filament during deposition. Morphology plays an important role in determining the sensitivity of these devices when used as combustible gas sensors.

INTRODUCTION

Combustible gases such as methane and hydrogen pose a safety threat when they accumulate at explosive concentrations. A well-developed technology to detect such species in air is the use of an electrically heated Pt filament or Pt-loaded ceramic bead [1-3]. These catalytic gas sensors measure the temperature change induced in the catalyst during combustion by means of a resistance variation [4]. Despite successful implementation for many years, these conventional combustible gas sensors have several limitations. Due to their relatively large size, they have high power consumption and slow response time. Often, hand manufacture or hand sorting of these devices is required [5,6]. For these reasons, they can be expensive, unreliable and not reproducible. Our goal is to overcome these limitations by taking advantage of Si microfabrication technology. Furthermore, this method of fabrication allows for both improved selectivity through the development of sensor arrays, and the integration of electronics for control and data acquisition [7, 8].

Micromachined combustible gas sensors have been reported elsewhere [5-6,9-10]. Using surface micromachining and "micro-CVD" techniques, we have designed, modeled, fabricated, and tested Pt-coated polysilicon microfilaments for the calorimetric detection of combustible gas mixtures. Similar microfilaments have been previously explored as potential incandescent light sources and hot-wire anemometers [11,12] but have as yet to be exploited for chemical sensing. Our filament fabrication procedure is similar to that developed by Mastrangelo [13]. The filaments also have much in common with "micro-hotplates," [5,7,14-16] but offer the potential advantage of less thermal mass and consequent reductions in power consumption and response time.

This work focuses on a post-processing, selective "micro-CVD" technology for the deposition of catalytic films on silicon nitride-passivated polysilicon filaments (Figure 1). We

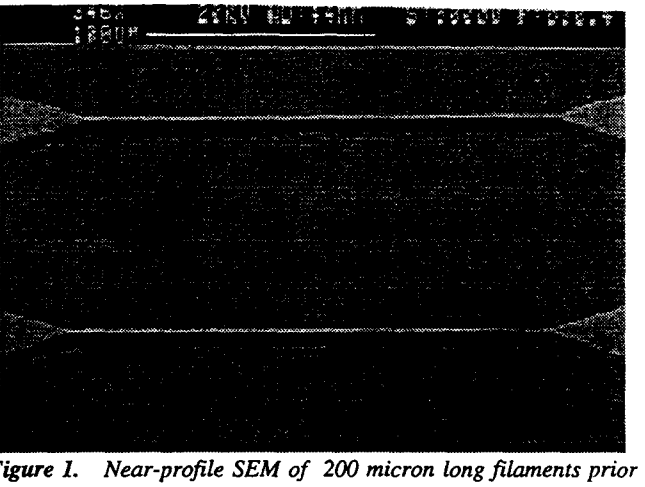


Figure 1. *Near-profile SEM of 200 micron long filaments prior to Pt deposition The filaments are 2 μm thick, 10 μm wide and elevated 2 μm above the substrate. They are encapsulated in a passivating layer of silicon nitride.*

believe this is an important technology for many chemical microsensors, because (a) film deposition prior to release limits film choice due to possible attack by the etchant, and (b) the fragility of released microstuctures renders them unsuitable for many conventional deposition techniques [17]. We use atmospheric-pressure deposition of Pt on surface-micromachined microfilaments by thermal decomposition of an organometallic Pt precursor, platinum acetylacetonate. This post-packaging deposition occurs only on the selectively heated filament(s).

Because catalyst morphology plays a role in determining the sensitivity, selectivity, and longevity of combustible gas detectors, control of platinum film microstructure is critical. Scanning electron microscopy (SEM) studies of Pt films are presented here and show that control of film morphology can be achieved by altering deposition time, filament temperature, and importantly, through the use of pulsed heating of the filament during Pt deposition. Sensitivity of these devices to various concentrations of hydrogen in oxygen and nitrogen is also reported and shows that film morphology is a critical parameter.

The interplay between gas transport kinetics and surface reaction rates can affect the sensor response as a function of catalyst film morphology and, indeed, can influence the catalyst morphology during its deposition. Higher surface-area films contain a larger number of reaction sites, tending to drive the rate of surface reaction, and hence the heat deposited in the filament, towards the mass-transport limit, thus improving sensitivity to combustible mixtures. During the CVD process, the relatively larger accessible gas volume that supplies organometallic precursors to "sharp" filament features such as corners and edges yields a higher mass-transport-limited flux than in the planar

0-9640024-1-8 $3.00 © 1996 TRF

Solid-State Sensor and Actuator Workshop
Hilton Head, South Carolina, June 2-6, 1996

regions of the filament, often resulting in edges covered by thicker and/or higher surface area deposits than found on top of the filament.

These variable-mass-transport effects can be manipulated by operation in a regime where depletion of the gas-phase concentrations of precursor (or reactants) is controlled: pulsed operation of the filament allows "replenishment time" between pulses, during which the concentrations of gaseous species recover to their bulk values. Manipulation of pulse width and duration therefore offers the opportunity to control the reaction rates for catalyst deposition, as well as combustible species detection, at various points between the two limits of mass transport and surface reaction rate.

EXPERIMENT

The micro-CVD apparatus is diagrammed in Figure 2. The packaged filament is placed in the CVD cell with a small amount of the Pt precursor. The cell is inserted in an oven at 140 °C. This temperature provides adequate vapor pressure of the Pt precursor in the cell, but no deposition occurs, except on those filaments that are electrically heated by the external control circuit. Deposition of platinum is optimized where filament temperature exceeds 450°C [18].

The control circuit is designed to maintain the filament at constant resistance and can provide continuous and pulsed reference resistances. In order to evaluate the effects of deposition time and filament bias mode (pulsed versus continuous), a matrix of deposition conditions was examined. For a given filament temperature distribution, deposition time and duty cycle of the heating pulse were varied. Deposition times of 0.7, 1.5, 3,7,15 and 25 min. were investigated using a 20 Hz pulse train with duty cycles of 5, 10, 20, 50 and 100% (continuous).

To detect combustible gas mixtures, the current and voltage required to maintain a Pt-coated filament at constant resistance were monitored during exposure to combustible hydrogen mixtures. Measurements were made in a steady-flow gas test bed in which total flow was regulated to 1 slm. Gas flows were prepared by mixing nitrogen and 20% oxygen with various concentrations of hydrogen using mass-flow controllers.

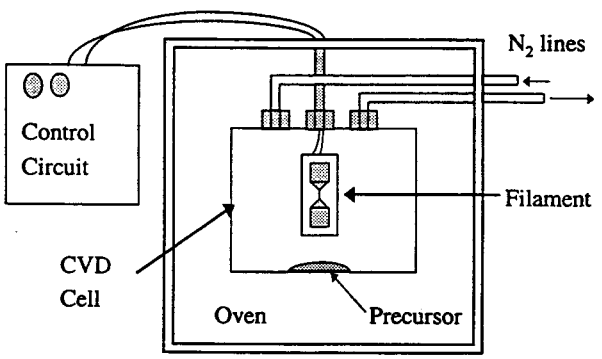


Figure 2. The micro-CVD apparatus: A teflon "CVD cell" houses the packaged filament and a small amount of the Pt precursor. The cell is equipped with electrical access to provide power to those filaments selected for deposition. Nitrogen lines insure an inert ambient in the cell. To vaporize the precursor, the CVD cell is inserted into an oven at 140 °C.

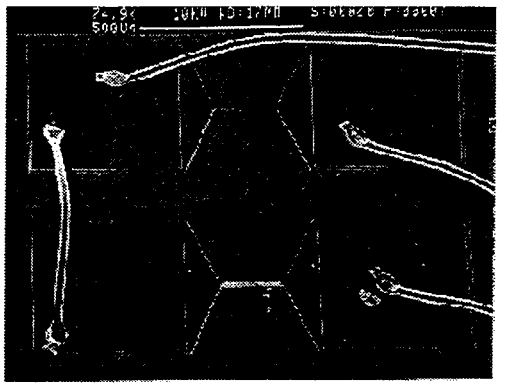


Figure 3. SEM top view of a pair of filaments. Al bond wires connect the square bond pads to a conventional dual in-line package (DIP). Only the lower filament which was electrically heated in the CVD cell contains a 0.2 μm layer of Pt

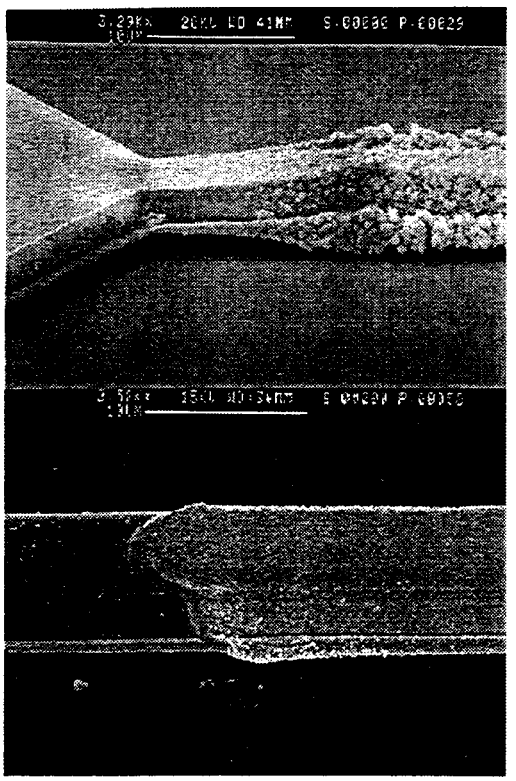


Figure 4. Near-profile SEMs of the ends of two filaments subjected to different deposition conditions. The high surface-area film on the top filament was deposited in continuous mode for 10-minutes. The lower film underwent a 9% duty-cycle (5 msec on, 50 msec off) deposition for a total elapsed time of 7 min., and is relatively smooth in comparison to the upper film.

RESULTS AND DISCUSSION

Figure 3 shows a top view of side-by-side filaments, only one of which (the lower) was electrically heated and consequently coated with Pt. Figure 4 illustrates two extremes of catalyst morphology that can be obtained by varying the deposition time and the duty cycle of the heating pulse.

To better understand the role these factors play in controlling the morphology, one must make a fair comparison between films

deposited under different conditions. One approach compares films formed using equivalent total energy inputs into the filaments during deposition. For a given filament temperature distribution, this energy is proportional to the amount of time the filament is heated during deposition (product of duty cycle and total deposition time). SEMs of two filaments, shown in Figure 5, demonstrate that although both had the same input energy, the higher duty-cycle heating increased the amount of Pt at that filament's edge. This result is consistent with the higher mass-transport limit of catalyst precursor that is expected at the edges of these filaments. In the extreme case demonstrated in Figure 4, heavy deposition on the upper filament's edge threatens to thermally connect the device to the substrate, and thereby defeat its mode of operation. This scenario underlines the importance of film morphology control.

A slight increase in surface roughness with increased duty-cycle is also evident in Figure 5, a result we have generally observed. Deposition time, on the other hand, mainly controls the film thickness, with a secondary effect on surface roughness. Figure 6 shows two devices with the same duty-cycle deposition (continuous), but different deposition times. An increase in film thickness with deposition time, along with an increase in roughness, is demonstrated.

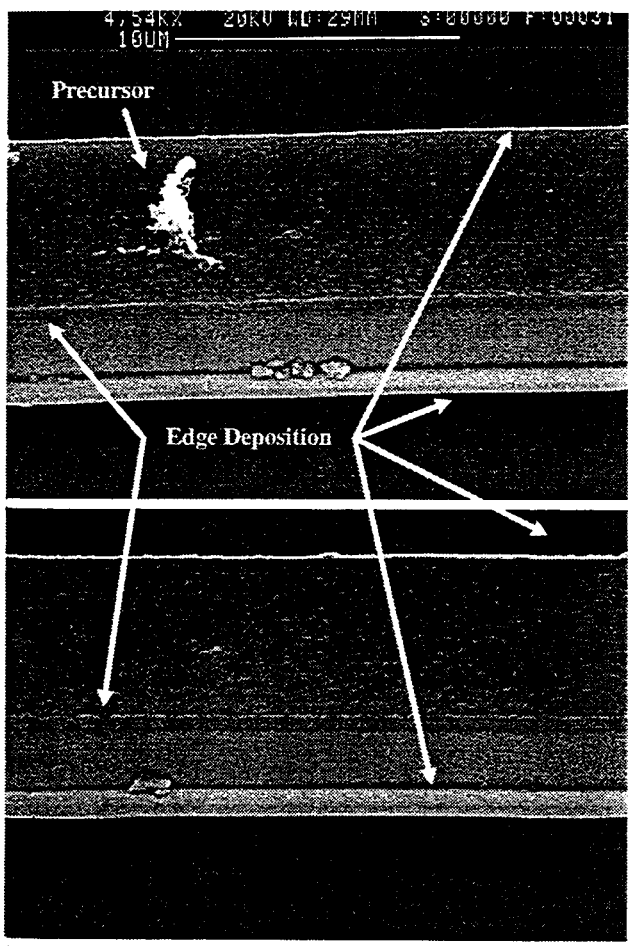


Figure 5. Two filaments with the same energy input. The upper filament underwent a 15 min. 10% duty-cycle, 20 Hz deposition. The lower filament had a 7.5 min. 20% duty-cycle, 20 Hz deposition. The overall Pt surface roughness and Pt deposition on the filament edge have both increased with duty cycle. A particle of precursor remained on the upper filament subsequent to deposition. Both SEMs were taken at the same magnification

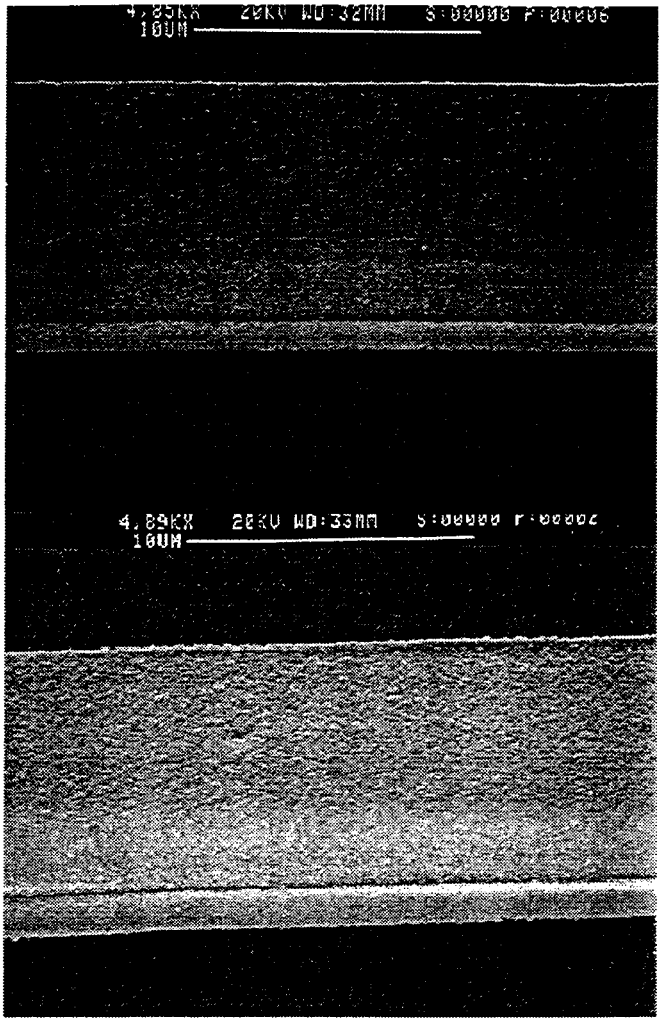


Figure 6. High magnification, plan view of two filaments. Both films were deposited in continuous mode. Deposition time and surface roughness increase from top to bottom in the figure. (From top to bottom, deposition times are 1.5 and 3.0 min.)

Typical sensor response to mixtures of hydrogen in oxygen is given in Figure 7. These data demonstrate that for small gas concentrations, response is rapid (about 30 seconds for a full-scale response). Thermal response data show that the filament actually reaches operating temperature in 0.5 msec. The detection limit is 100 ppm hydrogen in synthetic air. Furthermore, when operated in continuous mode, the filament consumes less than 70 mW of power; operation in a pulsed mode can further reduce total power consumption.

Figure 8 is a comparison plot, illustrating the relative effects of different deposition conditions (and therefore different catalyst morphologies) on sensor response. Prior to gas sensing, film morphology was evaluated qualitatively from SEM photographs. Film 1 had a smoother catalyst morphology and, as shown in Figure 8, roughly half the response of film 2. From these preliminary data it is clear that higher surface area films, which correspond to greater duty cycles and longer deposition times, are more sensitive. This result is consistent with mass-transport limited reaction rates.

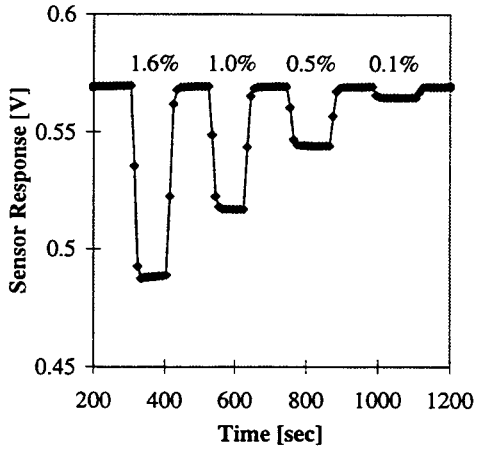


Figure 7. Sensor response of a Pt coated filament when subjected to various concentrations of hydrogen in 20% oxygen and a balance of nitrogen. The baseline signal is taken in pure nitrogen. Data was taken at 10 sec per point.

SUMMARY AND CONCLUSIONS

We have demonstrated an atmospheric pressure, post-processing, selective "micro-CVD" technology for the deposition of catalytic films on surface-micromachined, nitride-passivated polysilicon filaments. Catalyst morphology can be controlled by altering deposition time, filament temperature, and through the use of pulsed heating of the filament during deposition. Both deposition time and duty cycle can be varied to control the film thickness and surface roughness. Duty cycle also affects the amount of catalyst deposited at the filament edge. When used in the detection of combustion gases, these devices are low power, sensitive and fast. The most sensitive films have higher surface area, requiring longer duty cycles and deposition times. Future work will be directed at investigating the effects of filament temperature, precursor temperature, and pressure on morphology and sensitivity. By controlling these parameters, in addition to deposition time and duty cycle, film morphology and sensor response can be optimized.

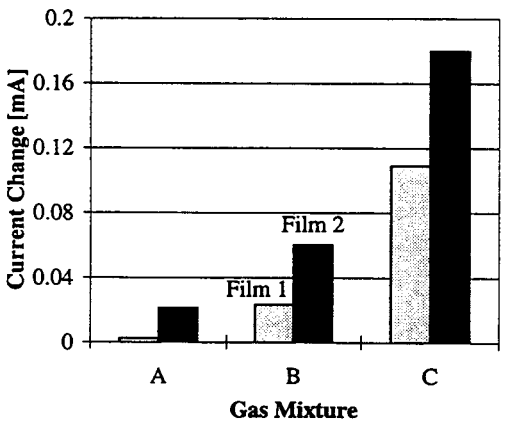


Figure 8. . Response of two films, both deposited with a 20 Hz heating pulse, to different concentrations of hydrogen (A=0.5%, B=1% and C=2%) in 20% oxygen and a balance of nitrogen. Film 1 was deposited for 7 min. using a 10% duty cycle pulse. Film 2 used a 20 % duty-cycle, 7 min. deposition.

ACKNOWLEDGMENTS

The excellent technical assistance of A.W. Staton, M.W. Jenkins, M. Hill, and the engineers, operators and technicians of the Sandia Microelectronic Development Laboratory is gratefully acknowledged. Thanks also to P. Shea for the SEMs shown in this paper. This work, performed at Sandia National Laboratories, was supported by the U.S. Department of Energy under contract DE-AC04-94AL85000.

REFERENCES

1. E. Jones, "The Pellistor Catalytic Gas Sensor", *Solid State Gas Sensors*, P.T. Moseleyand and BC. Tofield eds., A. Higler (Bristol and Philadelphia) 1987.

2. R. Dalla Betta and W.O. Bymum, "A New Combustion Control Sensor", *Sensors*, April 1989.

3. A. Accorsi, "Gas Sensors", *Proc. Capteurs '86*, Paris, France, 6/17-19/86, pp. 183-192.

4. S.J. Gentry and T.A. Jones, "The Role of Catalysis in Solid-State Gas Sensors", *Sensors and Actuators, 10*, 141 (1986).

5. M. Zanini, J. H. Visser, L. Rimai, R. Soltis, A. Kovalchuk, D. W. Hoffman, E. M. Logothetis, U. Bonne, L. Brewer, O. W. Bynum, and M. A. Richard, "Fabrication and Properties of a Si-based High Sensitivity Microcalorimetric Gas Sensor", *Tech. Digest 1994 Sol.-State Sensor & Actuator Workshop*, Hilton Head Isl., SC, 6/13-16/94, pp. 176 - 79.

6. C. Vauchier, D. Charlot, G. Delapierre and A. Accorsi, "Thin-Film Gas Catalytic Microsensor", *Sensors and Actuators B*, 33 (1991).

7. R.E. Cavicchi, J.S. Suehle, P. Chaparala, K.G. Kreider, M. Gaitan, and S. Semancik, "Micro-hotplate Gas Sensor", *Tech. Digest 1994 Sol.-State Sensor & Actuator Workshop*, Hilton Head Isl., SC, 6/13-16/94, pp. 53-56.

8. J. Smith, S. Montague, J. Sniewgoski, J. Murray, and P. McWhorter, "Embedded Micromechanical Devices for the Monolithic Integration of MEMS with CMOS", *Proc. IEDM '95*, pp. 609-612 (1995).

9. M. Gall, "The Si-Planar-Pellistor: A Low-power Pellistor Sensor in Si Thin-film Technlogy", *Sensors and Actuators, B4*, 553 (1991).

10. P. Krebs and A Grisel, "A Low Power Integrated Catalytic Gas Sensor", *Sensors and Actuators, B12*, 155 (1993).

11. Y. C. Tai, R. S. Muller, and R. T. Howe, *Transducers '85. 1985 International Conference on Solid-State Sensors and Actuators. Technical Digest of Papers*, p. 445 (1985).

12. C. H. Mastrangelo and R. S. Muller, "A Constant Temperature Gas Flowmeter With A Silicon Micromachined Package", *Tech. Digest 1988 Sol.-State Sensor and Actuator Workshop*, Hilton Head Isl., SC, June, 1988, pp. 43 - 47.

13. C.H. Mastrangelo, "Thermal applications of microbridges," Ph.D. Thesis, UC Berkeley, 1991.

14. C. L. Johnson, J. W. Schwank, and K. D. Wise, "Integrated Ultra-thin-film Gas Sensors", *Sensors and Actuators B*20, 55 (1994).

15. J. S. Suehle, R. E. Cavicchi, M. Gaitan, and S. Semancik, "Tin oxide gas sensor fabricated using CMOS micro-hotplates and in-situ processing", *IEEE Electron Device Letters.*, *14*, 118 (1993).

16. N. R. Swart and A. Nathan, "Design optimisation of integrated microhotplates", *Sensors & Actuators A*, *43*, 3 (1994).

17. S. Majoo, J.W. Schwank, J.L. Gland, and K.D. Wise, "A Selected-area CVD Method for Deposition of Sensing Films on Monolithically Integrated Gas Detectors", *IEEE Electron Device Letters*, *16*, 217 (1995).

18. B.S. Kwak, P.N. First, A. Erbil, B.J. Wilkens, J.D. Budai, M.F. Chisholm, and L.A. Boatner, "Study of epitaxial platinum thin films grown by metalorganic chemical vapor deposition", *J. Appl. Phys.*, *72*, 3735 (1992).

CHEMICAL CLASS SPECIFICITY USING SELF-ASSEMBLED MONOLAYERS ON SAW DEVICES

Ross C. Thomas and Antonio J. Ricco
Microsensor R&D Department, Sandia National Laboratories
Albuquerque, NM 87185-1425

Huey C. Yang, Dan Dermody, and Richard M. Crooks
Department of Chemistry, Texas A&M University
College Station, TX 77843-3255

ABSTRACT

We have studied the chemical selectivity and sensitivity of surface acoustic wave (SAW) sensors covered by $(COO^-)_2/Cu^{2+}$-terminated interfaces by examining the response of self-assembled monolayer (SAM) films formed from the solution phase for 36, 84, and 180 h adsorption times. These SAMs were prepared on thin-film Au surfaces having variable, controlled grain size. Our results have important implications for designing and reliably fabricating chemical sensors that respond to specific organic analytes.

INTRODUCTION

Chemically sensitive SAW devices provide a platform for sensors that respond to specific organic analytes.[1,2] Custom-designed and synthesized molecular recognition sites are an appealing means for obtaining chemical specificity, but the time-intensive nature of this approach, coupled with the plague of non-specific adsorption and the large number of molecules for which sensors are sought, means that custom "guest-host" complexes may ultimately find application in a small fraction of successful chemical sensors. An alternate approach is to relax selectivity requirements significantly, using arrays of devices to provide a distinctive "fingerprint" pattern of responses for each analyte.[3]

We have designed chemically sensitive interfaces, of moderate class selectivity, that make use of known, reversible, bulk-phase interactions between functional groups in the film and a given analyte.[4] These interfaces, which we immobilize on 97-MHz ST-quartz SAW delay lines (Figure 1), are formed from composite SAM films employing organomercaptans. The rationale for using this kind of SAM is to provide a simple methodology to quickly and reliably fabricate chemically active vapor/solid interfaces that respond reversibly and reproducibly to a wide spectrum of organic analytes. For example, it is well documented that SAM films prepared using millimolar concentrations of *n*-alkanethiols in ethanol produce monolayers having reproducible properties (i.e., surface coverage, structure, and orientation).[5] Moreover, research has shown that the chemical reactivity of the vapor/solid interface is easily controlled by altering the SAM terminal groups.[5]

Organomercaptan SAM films are most often formed onto Au surfaces. However, the deposition method and the substrate materials used to produce the Au layer for attaching the SAM film commonly varies for different research groups. For example, Au films might be prepared by thermal, electron-beam, or sputter deposition. In many instances, a Ti, Cr, or organic adhesion layer is used to bond the Au to Si, glass, quartz, or mica surfaces. Additional steps completed prior to forming the SAM film may include annealing or electropolishing the Au surface. Although

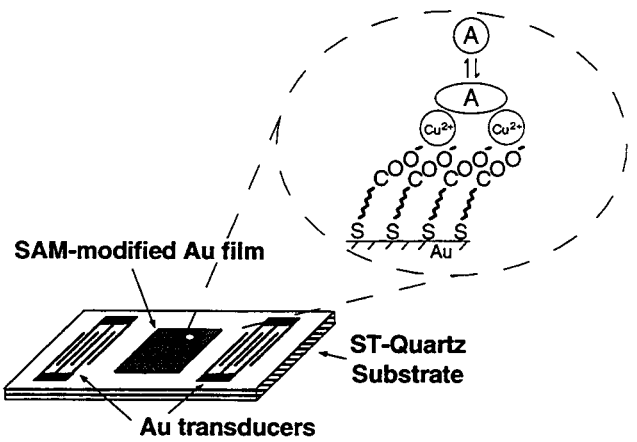


Figure 1. Schematic diagram of a self-assembled composite monolayer-modified chemically sensitive SAW device. The carboxylate coordinated Cu^{2+}-terminated interface depicted here displays significant selectivity, in addition to reversibility and durability, for detecting organophosphorous compounds.

the wide range of substrate materials and deposition conditions produce Au films having different grain sizes and surface roughnesses, SAM films having similar properties are nevertheless formed. The ordering of the S head group onto the primarily (111) crystallite orientation found to predominate for all of these Au surfaces, along with the densely-packed self assembly of the alkane portion of the organomercaptan molecules, most likely accounts for the consistent characteristics of the Au-supported SAMs.[5] Subtle effects (e.g., number of defects) within a SAM film, however, can result from differences in its microstructure, which is affected by the monolayer solution-phase adsorption time and/or the grain boundaries of the underlying Au film.[6]

In this paper, we compare the SAW response from a carboxylate coordinated Cu^{2+}-terminated SAM to that from a methyl-terminated SAM, as these films interact with a vapor-phase organophosphonate analyte as well as the vapors of a number of common organic solvents. (Selective detection of organophosphorous compounds, components of such substances as pesticides and chemical warfare agents, is important since these materials can pose a significant threat to health and environment.) We also examine in detail the effects of the solution-phase SAM adsorption time and the grain size of the supporting Au surface.

EXPERIMENTAL METHODS

The 97-MHz ST-quartz SAW devices were designed and fabricated at Sandia National Laboratories using interdigital transducers composed of fifty finger-pairs with a periodicity of 32

0-9640024-1-8 $3.00 © 1996 TRF

μm.[7] The transducers were defined photolithographically from 200 nm-thick Au. A 15 nm Ti adhesion layer was used to bond the Au to the quartz substrate. The composite SAMs were prepared on 100 nm-thick Au films, which are electron-beam evaporated between the transducers. No Cr or Ti adhesion layer was used, since both of these metals are known to diffuse through Au films; this could interfere with comparisons between composite SAM films formed for different lengths of time and on Au surfaces having variable grain sizes. Au grain size was controlled by annealing the films at different temperatures.

Composite monolayer films were prepared on newly deposited Au films by soaking the substrates in a dilute solution of 11-mercaptoundecanoic acid (MUA) in ethanol for 36, 84, and 180 h followed by a 10 min solution-phase exposure to ethanolic $Cu(ClO_4)_2 \cdot 6H_2O$. Atomic force microscopy was used to determine the grain size of the SAM-modified Au film after completing the SAW measurements. A computer-controlled flow system was used to measure adsorption/desorption isotherms.[8]

The relationship between the frequency shift of a SAW delay-line oscillator circuit and the mass loading is described in detail elsewhere.[9] Briefly, the change in frequency, Δf, is related to the change in adsorbed mass per area, $\Delta(m/A)$, when the SAW velocity is perturbed only by mass loading variations:

$$\frac{\Delta f}{f_o} = -\kappa c_m f_o \times \Delta(m/A) \qquad (1)$$

where c_m is the mass sensitivity (1.33 cm^2/g-MHz for ST-quartz), f_o is the unperturbed oscillator frequency (97 MHz), and κ is the fraction of the center-to-center distance between the transducers covered by the Au film; its present value is 0.7.

RESULTS AND DISCUSSION

Figures 2A and 2B show adsorption isotherms as a function of vapor-phase concentration for a series of organic analytes, as well as water, interacting with a SAW device covered by either a $(COO^-)_2/Cu^{2+}$- or a CH_3-terminated monolayer film, respectively. For clarity, only a small percentage of the recorded data points are plotted. Since "raw" frequency changes are biased in favor of analytes with higher molecular weight, we determine the extent of selectivity by normalizing the data for analyte molecular weight, expressed as the number of moles adsorbed/area.

The adsorption isotherms in Figure 2A reveal preferential response to diisopropylmethylphosphonate (DIMP) and acetone relative to other common organic solvents. These two analytes are expected to have moderately strong interactions with the $(COO^-)_2/Cu^{2+}$-terminated surface because of their available oxygen lone pairs; the coordinated Cu^{2+} is a Lewis acid and these two analytes are Lewis bases. Only the DIMP analyte, however, shows an appreciable amount of specific binding with this composite monolayer film when the results from the methyl-terminated SAM are taken into account. The degree of specificity of adsorbate/SAM interactions is most readily ascertained by taking the differences between the isotherms of Figures 2A and 2B, shown in Figure 2C. The data clearly indicate that the $(COO^-)_2/Cu^{2+}$-terminated SAM is almost 6 times more sensitive to DIMP than to acetone. The high level of specificity should apply to organophosphonates as a class.

The most important conclusion to be drawn from comparison of Figures 2A and 2B is that a pair of chemically sensitive SAW devices, one functionalized with the $(COO^-)_2/Cu^{2+}$- and the other with the CH_3-terminated monolayer film, readily distinguishes chemically specific from non-specific adsorption of analytes. For example, the similar responses for both acetone and water adsorbed on the two different SAM surfaces are implicative of interactions between the adsorbate molecules, not specific interactions with the SAM terminal groups. Water has a tendency to form H-bonded networks with itself, which probably explains the relatively large response. Since acetone has no H-bonding protons, a similar effect would not be expected. However, the acetone was not rigorously dried, so the entrained vapor undoubtedly contains a small amount of water, which we suspect may participate in the formation of a H-bonded Network.

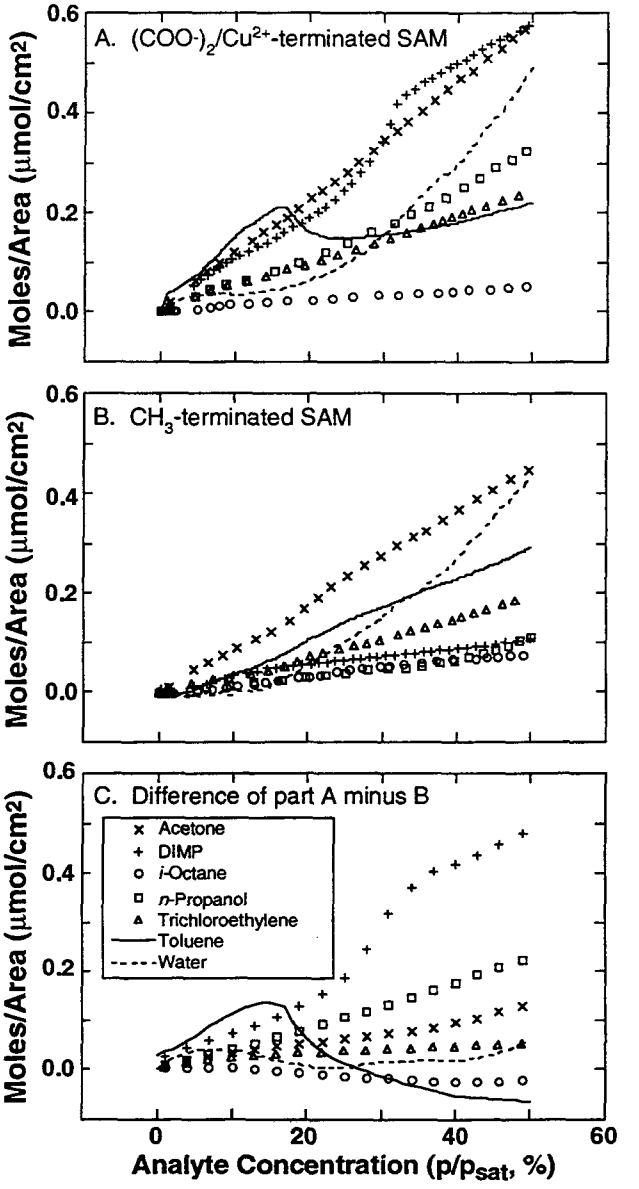


Figure 2. Adsorption isotherms for two SAW devices, covered with (A) $(COO^-)_2/Cu^{2+}$- and (B) CH_3-terminated monolayer films, exposed to six different organic analytes and water. (Responses are normalized for molecular weight). The degree of specificity of adsorbate-SAM interactions is most readily ascertained by comparing the difference isotherms obtained from part (A) minus part (B), shown in (C).

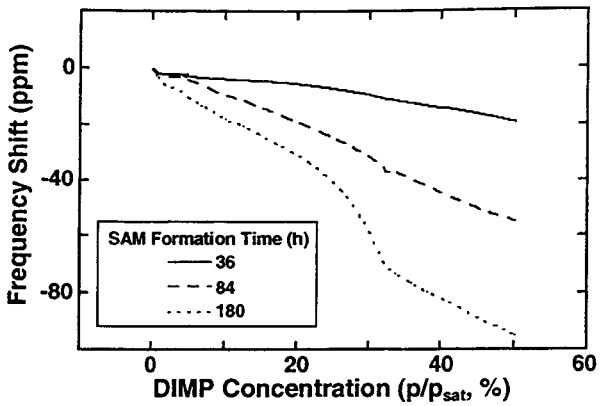


Figure 3. Adsorption isotherms of SAW devices covered by $(COO^-)_2/Cu^{2+}$-terminated interfaces exposed to DIMP. The MUA SAMs were prepared for 36, 84, and 180 h. The frequency shift as a function of concentration increases with the SAM formation time. Enhanced sensitivity results from structural ordering effects at the vapor/solid interface of the composite monolayer film.

The responses measured for *n*-propanol and trichloroethylene further support our conclusion. The SAW device with the $(COO^-)_2/Cu^{2+}$-terminated SAM shows about three times the response for propanol as the methyl-terminated SAM, an indication of specific interaction between the propanol (and probably alcohols as a class) and the more polar monolayer film. For trichloroethylene, the response is only about 30% larger for the $(COO^-)_2/Cu^{2+}$-terminated film, indicating a minor role for specific interactions. Only two of the species examined actually show a larger surface coverage at $p/p_{sat} = 0.5$ for the methyl-terminated surface than for the $(COO^-)_2/Cu^{2+}$ surface: *i*-octane and toluene, the least polar of the analytes. In the case of toluene, the precise extent of the difference is clouded by the partially non-mass-related response observed in Figure 2A. For *i*-octane, though the response is small compared to the other analytes, it is nonetheless about 60% greater for the CH_3-terminated surface, in accord with the very similar chemical nature of the surface and the analyte in this case.

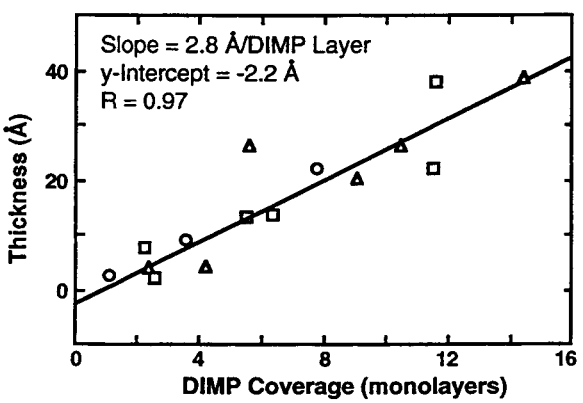


Figure 4. Ellipsometry data (y-axis) taken at the same time as SAW measurements (x-axis) for composite SAMs with MUA formation times of 36 (o), 84 (□), and 180 (▲) h. The DIMP vapor concentration was varied from 0-75% of saturation. The estimated coverage was determined from the frequency response using Eq. 1 and the bulk density of DIMP. The solid line is a linear fit thru all of the data. These results directly confirm multilayer adsorption of the DIMP analyte.

Figure 3 shows a series of DIMP adsorption isotherms obtained from SAW devices covered by $(COO^-)_2/Cu^{2+}$-terminated monolayer films that were assembled for 36, 84, and 180 h adsorption times. The binding affinity of this selective coating for DIMP increases dramatically as the MUA monolayer film is allowed to form for longer periods of time. Our results, which are consistent with DIMP multilayer formation (using Eq. 1), suggest that long adsorption times allow (re)organization of the SAM to provide a lower energy structure and an orientation of the $(COO^-)_2/Cu^{2+}$-terminated layer that is more favorable for inducing ordering in adsorbed DIMP. The frequency shift for the 180-h film at $p/p_{sat} = 50\%$ is equivalent to 17 layers of adsorbed DIMP.

We have used in-situ ellipsometry combined with SAW measurements to provide detailed information about the multilayer adsorption response of the chemical sensor. The data indicate that the thickness of all the composite SAMs is 1.9 ± 0.5 nm, which is in agreement with previous measurements and confirms that only a single monolayer of MUA forms (regardless of the solution-phase adsorption time). Figure 4 shows ellipsometry data (y-axis) taken at the same time as SAW measurements (x-axis) for composite SAMs with MUA formation times of 36 (o), 84 (□), and 180 (▲) h.. The DIMP vapor concentration was varied from 0-75% of saturation. We estimated the coverage from the frequency shift using Eq. 1 and the bulk density of DIMP.

As expected, the data show a linear relationship between the mass loading and film thickness (independent of the composite SAM film studied). The solid line represents a linear fit thru all

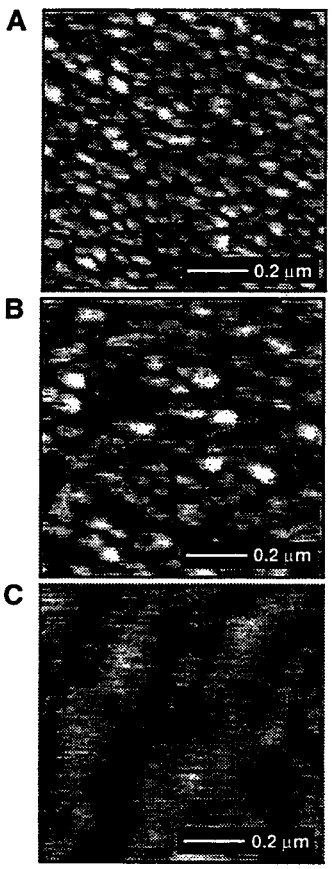


Figure 5. Constant repulsive-force images of 100 nm-thick Au films evaporated onto ST-quartz SAW devices. The average grain size, which we controlled by annealing the devices at different temperatures, was (A) 50 nm, (B) 80 nm, and (C) 240 nm.

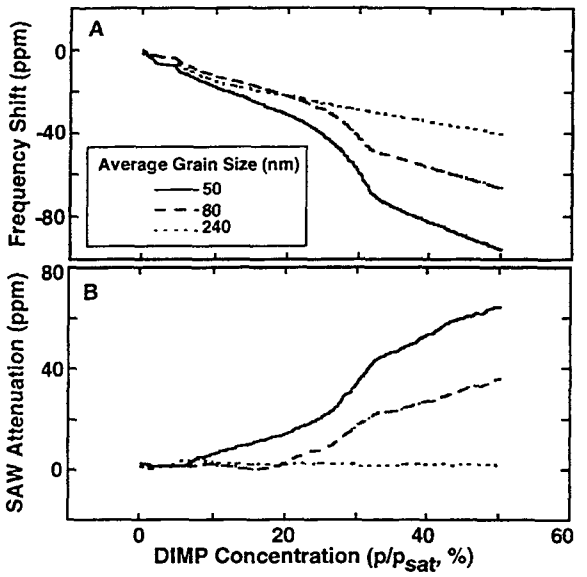


Figure 6. DIMP adsorption isotherms for composite monolayer films prepared on the Au surfaces having the variable grain sizes shown in Figure 5. (A) Frequency shift and (B) SAW attenuation versus the vapor-phase concentration of DIMP illustrate that sensitivity of SAW devices with composite SAMs prepared on Au films increases as grain size decreases.

of the data with a slope of 2.8 Å/DIMP layer and a y-intercept of -2.2 Å. Because each DIMP layer should be about 6.0 Å thick, the results suggest that the SAW measurements overestimate the mass loading by a factor of 2. We have not, however, taken into account non-specific adsorption of the DIMP analyte over the active region of the SAW device not covered by Au. Control experiments using bare quartz SAW devices exposed to DIMP show a measurable response that is partly irreversible, irreproducible, and dependent on surface cleanliness. Thus, the data do confirm multilayer adsorption, but additional measurements are necessary to reconcile quantitatively the SAW and ellipsometry responses.

The composite-SAM-modified Au surfaces shown in Figures 5A, 5B, and 5C produced the DIMP isotherm data shown in Figure 6 by the solid, dashed, and dotted lines, respectively. Clearly, as the grain size of the Au film *decreases* (by a factor of about 5), there is a significant *increase* in the extent of adsorption of DIMP at a given p/p$_{sat}$. We used electrochemical measurements to verify that surface roughness is *not* the dominant factor in these results: while the roughness changes by 11%, the SAW response varies by a factor of 2.5 at p/p$_{sat}$ = 0.5 from the largest to smallest grain size. Since the crystallite orientation of the Au is predominantly (111) and is not altered by annealing, the total coverage of the MUA SAMs for the three grain sizes should be very similar. The attenuation of the acoustic wave (Figure 6B) indicates that the grain size also has a strong influence on the structural properties of the SAM and its consequent interactions with adsorbed analyte.

SUMMARY AND CONCLUSIONS

Our results illustrate that a pair of SAM-functionalized SAW sensors, one bearing a methyl- and the other a $(COO^-)_2/Cu^{2+}$-terminated monolayer, allow clear distinction to be drawn between chemically specific and nonspecific adsorption of analytes. The results further show that the composite SAM films (1) preferentially adsorb particular classes of organic analytes in a manner that follows from simple concepts such as Lewis acid/base, H-bonding, and polar vs. nonpolar interactions, (2) adsorb a greater quantity of DIMP at a given partial pressure for longer solution-phase formation times of the organomercaptan monolayer, and (3) adsorb a greater quantity of DIMP at a given partial pressure as the grain size of the supporting Au film decreases. While result (1) is straightforward and to be expected, we believe results (2) and (3) are consequences of the extent and nature of the ordering of the SAM and its outer surface: the ordering affects the thermodynamics of adsorption of multilayers. The results from these measurements are an important step toward reliably fabricating chemical sensor arrays that respond to organic analytes with controlled class and/or molecular specificity.

We gratefully acknowledge the excellent technical assistance of Al Staton and Mary-Anne Mitchell of Sandia National Laboratories. This work was supported by the U.S. DOE under contract DE-ACO4-94AL85000 and by the National Science Foundation (CHE-9312441).

REFERENCES

1. R. C. Hughes, A. J. Ricco, M. A. Butler, and S. J. Martin, "Chemical Microsensors", *Science*, *254*, 74 (1991).

2. J. W. Grate, S. J. Martin, and R. M. White, "Acoustic Wave Microsensors", *Anal. Chem.*, *65*, 940A (1993).

3. A. J. Ricco, G. C. Osbourn, J. W. Bartholomew, R. M. Crooks, C. Xu, and R. E. Allred, "New Materials and Multidimensional Cluster Analysis for SAW Chemical Sensor Arrays", *Technical Digest of the 1994 Solid-State Sensor and Actuator Workshop*, Hilton Head Isl., SC, 6/13-16/94 Transducer Research Foundation, Cleveland (1994), pp. 180-183.

4. R. C. Thomas, H. C. Yang, C. R. DiRubio, A. J. Ricco, and R. M. Crooks, "Chemically Sensitive Surface Acoustic Wave Devices Employing a Self-Assembled Composite Monolayer Film: Molecular Specificity and Effects Due to SAM Adsorption Time and Au Surface Morphology," *Langmuir*, accepted for publication (12/95).

5. L. H. Dubois and R. G. Nuzzo, "Synthesis, Structure, and Properties of Model Organic Surfaces", *Annu. Rev. Phys. Chem.*, *43*, 437 (1992) and references therein.

6. G. E. Poirier and M. J. Tarlov, "Molecular Ordering and Gold Migration Observed in Butanethiol Self-Assembled Monolayers", *J. Phys. Chem.*, *99*, 10966, (1995).

7. A. J. Ricco and S. J. Martin, "Thin metal film characterization and chemical sensors: monitoring electronic conductivity, mass loading and mechanical properties with surface acoustic wave devices", *Thin Solid Films*, *206*, 94 (1991).

8. N. W. Ashcroft and N. D. Mermin, *Solid State Physics*; Saunders College (1976), Chapter 4, p. 63.

9. S. J. Martin, A. J. Ricco, D. S. Ginley, and T. E. Zipperian, "Isothermal Measurements and Thermal Desorption of Organic Vapors Using SAW Devices", *IEEE Trans. on UFFC, UFFC-34*, 142 (1987).

WAFER-TO-WAFER TRANSFER OF MICROSTRUCTURES FOR VACUUM PACKAGING

Michael B. Cohn, Yiching Liang*, Roger T. Howe, and Albert P. Pisano*
Departments of Electrical Engineering & Computer Science and Mechanical Engineering*
Berkeley Sensor & Actuator Center
Berkeley, CA 94720-1770

ABSTRACT

Vacuum packages based on wafer-to-wafer transfer of polysilicon shells are described. Transfer is accomplished by Au/Si eutectic bonding, in conjunction with break-away tethers. The low bonding temperature (370°C) enables end-process application. The short anneal time (~30 seconds) and wafer-scale parallelism promise high manufacturing throughput. Inexpensive packages of this type are desired for resonant sensors and implantable devices. However, a more general application may lie in assembly of the devices themselves, i.e. by transfer of pre-fabricated sensor structures, circuits, and other components.

INTRODUCTION

Resonant microstructures require a vacuum environment to achieve a high mechanical Q. Examples of this sensor class include resonant accelerometers [1] and vibratory rate gyroscopes [2], [3], [4]. Volume manufacturing of these devices will most likely require a micromachined vacuum encapsulation process, which would allow the sensor chip to be enclosed in a conventional, low-cost electronic package.

Several groups have demonstrated micropackaging processes, based on surface micromachining, wafer bonding, eutectics, and other methods. The current work borrows elements from all of these, but introduces a new technique, namely the batch transfer of released microstructures from one wafer to another. In this method, micromachined shells are fabricated on a handle wafer, using a sacrificial layer process such as HexSil [5]. As shown in Figure 1, the shells are anchored to the substrate by tethers, and coated with gold. The handle wafer is then contacted to a second wafer and heated to the Au-Si eutectic point. When cool, the two wafers are separated, breaking the tethers and leaving the shells welded to the second substrate.

This transfer process achieved 100% yield in chip-level transfers, with 30 packages per chip. Vacuum packaging was achieved by annealing in vacuum. At the end of a 5-month test, substantial vacuum persisted. Precision alignment was obtained by mechanical interference fit between lithographically defined features on the two substrates. An unexpected advantage of this process is its robustness to surface roughness and particulate contamination.

BACKGROUND

Micromachined packaging dates to Wise's 1981 single-crystal silicon hemispherical microshells for inertial fusion targets [6]. Petersen [7] has shown pressure sensors with sealed reference cavities fabricated by silicon fusion bonding. Schmidt [8] produced sealed pressure sensors by Si/SiO_2 wafer bonding and etchback. Ikeda [9] and Guckel [10] have demonstrated surface micromachined encapsulation for resonant beams, evacuated by reactive sealing.

Various glass packages have been shown, typically anodically bonded to millimeter-scale sensor chips or thin-film structures [11], [12].

Mastrangelo [13] and Lin [14] have demonstrated vacuum

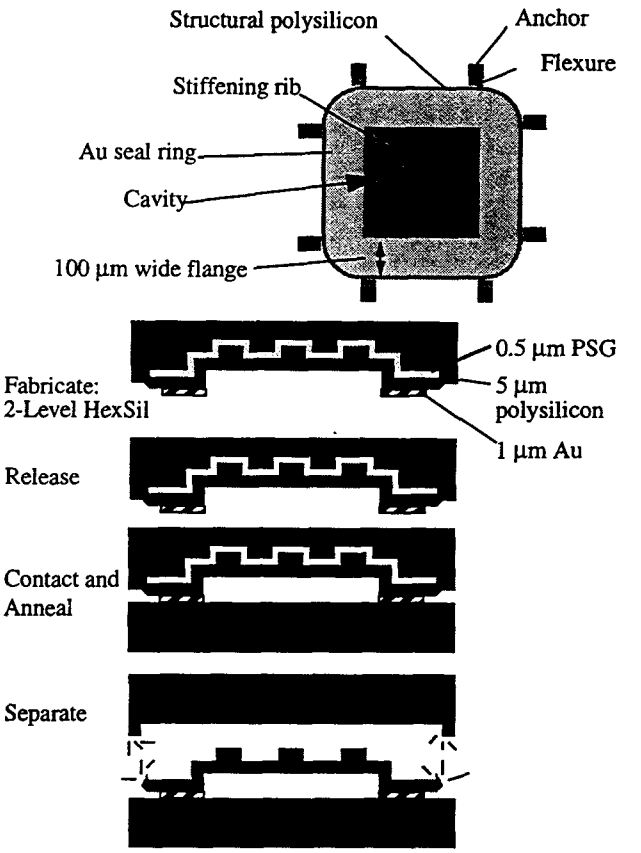


FIGURE 1. Above: Layout of a typical polysilicon cap for vacuum packaging. Below: Cross-section view of bonding process.

packages for incandescent microlamps and microresonators, respectively. In this work, a silicon nitride film was deposited over a sacrificial PSG layer; the resulting shells were etched out with HF and sealed by CVD. The final pressure is limited by residual CVD gases.

Several researchers have employed eutectic bonding, both for assembly and packaging [15], [16]. An advantage of this process is its low temperature: e.g. 183°C for the Pb/Sn system, and 363°C for Au/Si.

PROCEDURE AND RESULTS

The basic packaging procedure consists of two parts: fabrication and transfer. The packages – essentially polysilicon hemishells – are fabricated by the HexSil process [5]. Starting with a <100> silicon substrate, an anisotropic plasma etch defines a mold for the package's stiffening ribs. These have a width of 8 µm and a depth of 5 µm, and are laid out in a cartesian grid. A second plasma etch defines cavities 300 µm square and 4 µm deep for the packages themselves. This depth should not be exceeded without special resist coating equipment, or step coverage problems may occur in subsequent lithography.

The etched surfaces are smoothed by a wet silicon etch

0-9640024-1-8 $3.00 © 1996 TRF

Solid-State Sensor and Actuator Workshop
Hilton Head, South Carolina, June 2-6, 1996

(50 HNO_3:20 H_2O:1 NH_4F, 5 min), followed by growth of 0.5 μm of wet thermal oxide, and removal of the oxide in concentrated HF.

A sacrificial oxide is then deposited. PSG is preferred, since release of the caps will require 250 μm of undercut. LTO may be used as a diffusion barrier to protect adjacent layers from doping. The film is densified at 1000°C for 1 hour in nitrogen.

The oxide is patterned with anchor holes. A 20 μm square geometry proves adequate, though other sizes may be possible. In photoresist application, care should be taken to ensure adequate coverage around the cavities. Multiple coats may be required, particularly if plasma etching. Adhesion promoters such as HMDS may be detrimental. If a spin coater is used, "stale" photoresist must be purged from the feed line. A minimal hardbake should be used, as resist tends to pull back from steps.

Polysilicon deposition is preceded by a 15 second buffered HF dip and 2 minute DI rinse, to remove native oxides from the anchor contacts. Particular care must be taken, because of the long HF release step which will follow. Doping and grain size in the polysilicon film may be chosen to suit; undoped, coarse-grained (605°C, 100 nm roughness) material was mainly used, since this gave the highest growth rate. Undoped material should be isolated from PSG to prevent doping-induced stress gradients. Films were annealed in nitrogen at 1000°C for 1 hour to remove residual stress. A thick film – e.g. 4-5 μm – is desired, to avoid damage during release and bonding. The film should fill the trenches defined for the stiffening ribs. A test structure consisting of a long cantilevered beam should be included for thickness measurement, since optical methods fail beyond ~ 1 μm.

Polysilicon is defined by plasma etch. Again, photoresist step coverage is vital, to avoid "burnthrough" at the edges of the cavities. Tethers were designed to be robust, to prevent structures from detaching during the release. 4 μm wide x 10 μm long, and 30 μm wide x 30 μm long designs both worked satisfactorily. Eight anchors were provided for each 500 μm square package. Flanges on the packages were 100 μm wide. Narrower designs may be possible.

Gold deposition is preceded by a BHF dip, as above, to ensure adhesion in the release. In addition, 100 nm of chrome was deposited deposited as an adhesion layer. This also seems to prevent excessive diffusion of the gold during bonding, and the layer is apparently not undercut in the HF release. An evaporated gold film of 0.7 - 1.0 μm thickness yields good results in bonding. Problems were encountered with films in the 100 nm range.

Following metallization, wafers were diced and released. Release required 25 minutes in concentrated (49%) VLSI-grade HF. A surfactant (Triton X-100) was added to prevent collapse due to surface tension. A 5 minute DI rinse was followed by critical point drying in CO_2 [17].

Target substrates consisted of 2 μm thick polysilicon on a single-crystal substrate. Roughness was 10 nm peak-to-peak. Targets were prepared with the BHF dip procedure, followed by blow-dry with nitrogen. Handle and target chips were then contacted in a Class 100 environment. Contact was performed by laying one edge of the handle chip down on the target chip, then allowing the handle chip to fall flat. The handle could then be positioned laterally to align the two chips. The chips were placed in the bonding apparatus and clamped. The target chips were diced about 10% larger than the handle chips to simplify handling of the sandwich.

Approximately 60 kPa of contact pressure was applied to the chips in the apparatus, shown in Figure 2. Note that this force is distributed only over the bonded areas, making the actual contact pressure about 1.2 MPa. Dummy structures were placed at each corner of the layout to ensure parallelism of the chip faces.

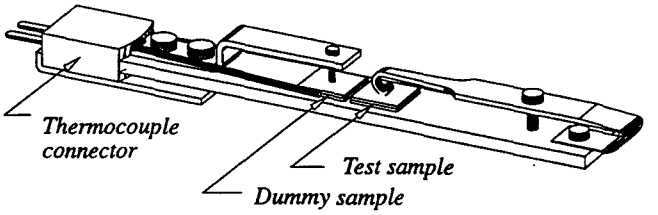


FIGURE 2. Bonding apparatus. The apparatus is placed in a quartz vacuum chamber and heated from below with an external lamp The clampling spring (right-hand side) applies contact pressure to the test sample.

Bonding was performed in a vacuum of ~ 10^{-5} T. The anodized aluminum surface of the bonding apparatus provides a susceptor, allowing consistent heating by the external lamp. A temperature of 370°C was found to be effective for Au/Si eutectic bonding. Temperature was monitored by a thermocouple mounted inside a sandwich of two silicon chips, clamped adjacent to the sample (see Figure 2). The thermal ramp-up to 370°C required approximately 5 minutes, after which cooling was begun immediately. Approximately 30 - 60 seconds was spent above the eutectic temperature. After cooling, the two chips were carefully separated to avoid shearing the microstructures. Separation breaks the polysilicon tethers (see Figure 1), leaving the shells bonded to the target. Results are shown in Figure 3.

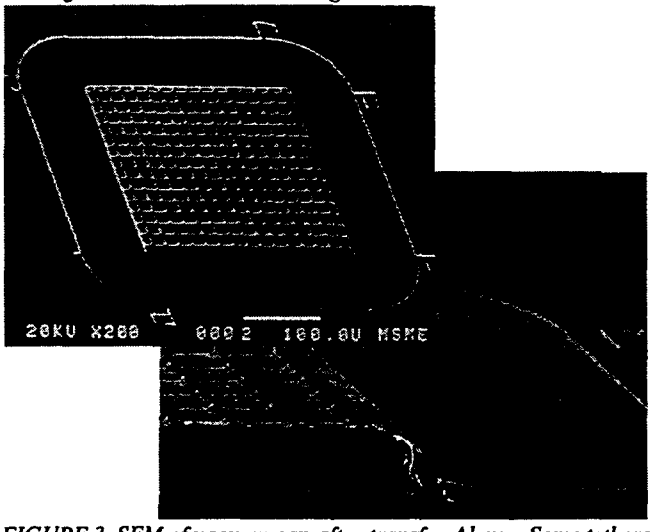


FIGURE 3. SEM of vacuum cap after transfer. Above: Some tethers are still partially attached to edge of cap. Flange (surrounding ribbed square) is 100 μm wide. Below: Cleaved cross-section of ribbed cap.

Most of the packages were designed with ribs, so that there would be minimal deformation under atmospheric pressure. However, several packages were designed without ribs so that Nomarski imaging (Figure 4) could reveal a displacement of the shell. Profilometry indicated a displacement of up to 680 nm on 300 μm square, 4.47 μm thick structures. The displacement disappeared when the package was punctured. Substantial vacuum persists at the end of a five-month test, as shown in Figure 5.

Profilometer measurements after transfer indicate significant compression of the seal structure. From an initial thickness of 730 nm, the gold layer is compressed to an apparent thickness of 346 nm. Some of this compliance may have occurred in the adjacent

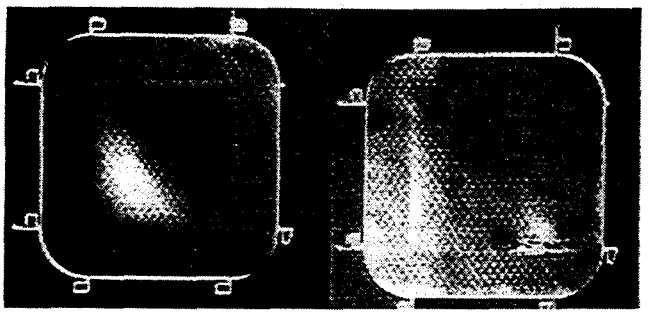


FIGURE 4. Cap without ribs was designed to deflect under pressure differential. Left: after assembly and anneal. Right: same, after being punctured. Nomarski imaging.

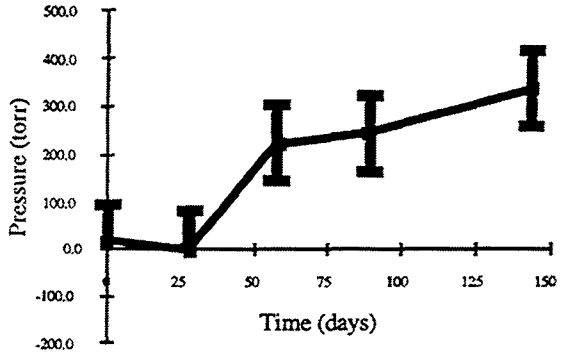


FIGURE 5. Package internal pressure as a function of time.

layers of the structure (refer to Figure 1).

DISCUSSION

The transfer and bonding process described here displays several important features and unexpected advantages. Since a high degree of roughness is tolerated on the mating surfaces, packages can be pumped out after being assembled. This simplifies the apparatus considerably. In addition, the compression of the seal – more than 500 nm in some cases – suggests application to feedthroughs and other planarization problems.

There is evidence that complex, high-aspect structures can be transferred. The sample in Figure 6 was prepared by a "wet transfer" process, in which release occurs after the two chips are contacted. Residual SiO_2 from the DI rinse most likely causes bonding [18]. Subsequent processing included sputter deposition of 100A of gold.

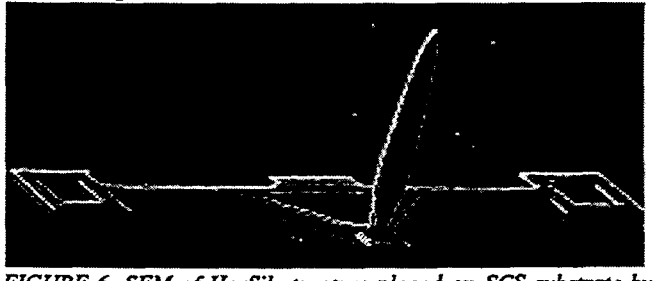


FIGURE 6. SEM of HexSil structure placed on SCS substrate by "wet transfer" method. The vertical feature is 80 μm in height.

Although these transfer processes appear robust to surface roughness and a modicum of particulate contamination, released structures are generally fragile and are potential particle generators themselves, if accidentally detached from the substrate. Fortunately, robust anchors can be liberally distributed, because of the high strength of the eutectic bond.

Shear strength of the bond was measured by bonding two unpatterned silicon chips. The resulting sandwich was supported along two opposite edges and a load applied at the middle. A crack formed and propagated through the sandwich, with no evidence of de-bonding. A value of at least 0.39 GPa may be calculated for the shear strength in this case. This is given by the shear stress at the neutral axis of the two-chip sandwich, when the maximal stress reaches the yield strength of silicon. The value seems reasonable in light of Lee's result of 5.5 GPa for the tensile strength, achieved under similar annealing conditions [16].

To ensure survival of the caps during release and contact, it may also be useful to avoid stress concentrators in the design. For example, hexagonal ribbing is preferred to square, and rounded corners to right angles.

A process has been developed for patterning the shells with reinforcement ribs. In fabrication, the rib structures are etched first, followed by the package cavities. In this way, the mask stepper need only focus at the plane of the wafer surface. Trench widths should be chosen as with HexSil [5]. Oxide thickness was minimized, on the premise that a thin sacrificial layer would prevent damage during contact, by avoiding excessive strain at the anchors. Thicker layers may be usable.

Patterning of the sacrificial oxide is complicated by the topography. Poor resist coverage tends to occur at the edges of the cavities, leading to loss of the oxide and failure to release. Softening of these steps with an isotropic etch may be beneficial. Generous use of photoresist is recommended, e.g. 4 μm or more, depending on step height. Alternately, the cavity depth may be reduced, since sagging of the ribbed packages is negligible.

Some latitude is possible in choice of structural material. Both undoped and doped, coarse- and fine-grained polysilicon recipes have been used. Silicon nitride has been shown by Lin [14] in a similar application, and gold is also a possibility.

Gold adhesion in the release step was variable. The preceding BHF dip seems vital, however. In one trial, a short plasma etch ($CH_3F/CF_4/He$) was used instead, with the intent to maintain a dehydrated surface for the gold evaporation. Poor adhesion resulted in this case.

Accurate temperature control is recommended. Significantly exceeding the eutectic temperature causes distortion of the polysilicon structures, apparently due to consumption by the liquid eutectic. An example is shown in Figure 7. Control of temperature and the initial gold distribution may lead to novel capabilities, such as plasticly deformed shells, welded links, or fusible supports.

FIGURE 7. SEM of a corner of a vacuum package, after bonding on a 700 nm Au film at ~ 400°C. Note "smearing" and grainy appearance of material.

In the sealed structures, the differential pressure across the cap may be determined from the expression for plate deformation under a distributed load [19]:

$$\Delta z = \frac{0.00126 q a^4}{D}$$

where

$$D = \frac{Eh^3}{12(1-\nu^2)},$$

the Poisson ratio of polysilicon, ν, is taken as 0.23, the Young's modulus, E, is 170 GPa, Δz is the measured deformation of the top surface of the cap, the width of the plate, a, is 300 μm, and the thickness of the plate, h, is 4.47 μm. Rearranging, we find q = 816±75 T. The uncertainty is contributed mainly by E.

Since the ambient pressure at the time of measurement was ~ 752 T, the internal pressure is expected to be in the range of 0-12 torr. The variation of this value with time is plotted in Figure 5. The long time constant suggests outgassing rather than leakage; a more aggressive bakeout, or gettering, might yield improvement.

WORK IN PROGRESS

Alignment of handle and target was achieved by patterning a 1.7 μm oxide step on the target, contacting the two chips, and abutting features on the handle against the step. A sharp piece of silicon was dragged across the back of the handle chip to generate a lateral force. It was expected that this force would be limited by the Si-Si friction coefficient, and thus would be commensurate with the forces present on the front side of the chip. Misalignment after bonding was less than the detection limit of 5 μm.

Encapsulated resonator structures were designed using a periodic pattern of conductive feedthroughs to favor planarization [12]. A 300 nm doped polysilicon film was used in a 5 μm line-space pattern, shown in Figure 8. A 2 μm LTO/PSG layer was overlaid for planarization, followed by 1 μm of low stress silicon nitride to protect it from the release. The same oxide film was also used for the resonator sacrificial layer. 2 μm polysilicon doubles as the resonator structural layer and a buffer layer over the feedthroughs for additional planarization. This surface displayed a 170 nm height modulation due to the feedthroughs, suggesting that a seal should be readily achieved, considering the 500 nm compliance described above. The die also includes multi-bore needles with internal electrical feedthroughs, hinge structures, and transferred resonators.

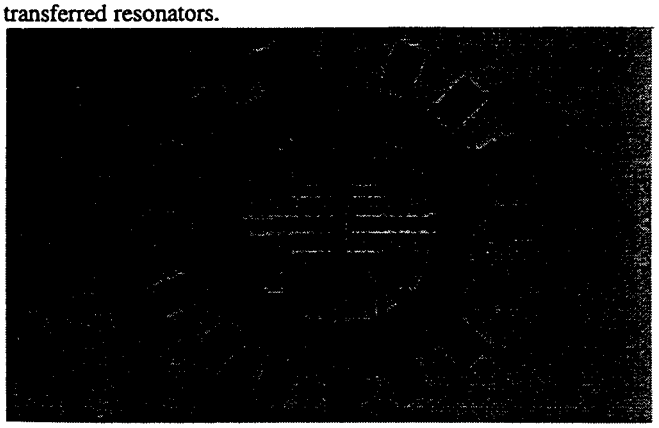


FIGURE 8. Feedthrough structure for encapsulated resonator. 3000A doped polysilicon connections are supplied for an electrostatic comb drive, sensing capacitors, and fusible supports.

Subsequent work will pursue both improved vacuum packaging and more general applications. For instance, the transfer process may be proposed for sensor structures, i.e. to combine resonators with CMOS. GaAs devices are also a candidate. This approach avoids many process-compatibility problems, and offers to simplify MEMS development: pre-fabricated, VLSI-scale modules (op-amps, microcontrollers) could be incorporated into many systems by assembly, rather than monolithic integration.

CONCLUSION

A technique has been demonstrated for batch transfer of complex polysilicon structures between two wafers. A gold/silicon eutectic forms a high-strength bond and hermetic seals. The bonding process is robust to surface roughness and contamination, and can be performed at low temperature. Packages can be evacuated after assembly, simplifying equipment design.

ACKNOWLEDGMENTS

The authors thank Mark Noworolski for sharing his results on gold adhesion. This work was funded by ARPA under contracts FA-ADI-DTD8/5/93 and DABT63-95-C-0028, and by the Berkeley Sensor & Actuator Center.

REFERENCES

1. Roessig, T., Pisano, A. P., and Howe, R., T., "Surface Micromachined Force Sensor," *Proceedings of the ASME Dynamic Systems and Control Division,* ASME, 1995, p. 871.

2. Bernstein, J.; Cho, S.; King, A.T.; Kourepenis, A.; and others, "A micromachined comb-drive tuning fork rate gyroscope," *Proceedings, IEEE Micro Electro Mechanical Systems.*, Fort Lauderdale, FL, 7-10 Feb. 1993, p. 143-8.

3. Clark, W. A., "Surface Micromachined Z-Axis Vibratory Rate Gyroscope," *Technical Digest of the 1996 Solid-State Sensor and Actuator Workshop,* Hilton Head, SC, 6/2-6/96.

4. Juneau, T. and Pisano, A. P., "Micromachined Dual Input Axis Angular Rate Sensor," *Technical Digest of the 1996 Solid-State Sensor and Actuator Workshop,* Hilton Head, SC, 6/2-6/96.

5. C. G. Keller and M. Ferrari, "Milli-Scale Polysilicon Structures," *Technical Digest of the 1994 Solid-State Sensor and Actuator Workshop,* Hilton Head, SC, 6/13-16/94.

6. Wise, K. D., Robinson, M. G., and Hillegas, W. J., "Solid-State Processes to Produce Hemispherical Components for Inertial Fusion Targets," *J. Vac. Sci. Technol.*, 18, April 1981, p. 1179-1182.

7. Petersen, K., Barth, P., Poydock, J., and others, "Silicon Fusion Bonding for Pressure Sensors," *1988 Solid-State Sensor and Actuator Workshop,* Hilton Head, SC, 6/6-9/88, p144-7.

8. Huff, M. A., Nikolich, A. D., and Schmidt, M. A., "A Threshold Pressure Switch Utilizing Plastic Deformation of Silicon," *1991 International Conference on Solid-State Sensors and Actuators,* IEEE, San Francisco, CA, 6/24-27/91, p.177.

9. Ikeda, K., Kuwayama, H., et al, "Three-dimensional Micromachining of Silicon Pressure Sensor Integrating Resonant Strain Gauge on Diaphragm," *Sensors and Actuators,* A21-23, 1990, p. 1007-10.

10. Sniegowski, J.J.; Guckel, H.; Christenson, T.R. "Performance characteristics of second generation polysilicon resonating beam force transducers." *IEEE Solid-State Sensor and Actuator Workshop* Hilton Head Island, SC, 6/4-7/90 p. 9-12.

11. Hoogerwerf, A. C. and Wise, K. D., "A Three-Dimensional Neural Recording Array," *1991 International Conference on Solid-State Sensors and Actuators,* IEEE, San Francisco, CA, 6/24-27/91, p. 120.

12. Ziaie, B., Gianchandani, Y, and Najafi, K., "A High-Current IrOx Thin-Film Neuromuscular Microstimulator," *1991 International Conference on Solid-State Sensors and Actuators,* IEEE, San Francisco, CA, 6/24-27/91, p. 124

13. Mastrangelo, C. H., Muller, R. S., Kumar, S., "Microfabricated Incandescent Lamps," *Applied Optics,* 1 March 1991, vol.30, (no.7):868-73.

14. Lin, L., McNair, K. M., Howe, R. T., and Pisano, A. P., "Vacuum Encapsulated Lateral Microresonators," *1993 International Conference on Solid-State Sensors and Actuators,* IEEE, p. 270-3.

15. Ko, W. H., Suminto, J. T., and Yeh, G. J., "Bonding Techniques for Microsensors," *Micromachining and Micropackaging of Transducers,* Elsevier, 1985, p. 41.

16. Lee, A. P., Ciarlo, D. R., and others, "A Practical Microgripper by Fine Alignment, Eutectic Bonding, and SMA Actuation," *Digest of the 1995 International Conference on Solid-State Sensors and Actuators,* Stockholm, Sweden, 6/25-29/95, p. 368-71

17. Mulhern, G. T., Howe, R. T., and Soane, D. S., "Supercritical Carbon Dioxide Drying of Microstructures,"*1993 International Conference on Solid-State Sensors and Actuators,* IEEE, p. 296-9.

18. Judy, M. W., Ph.D. Thesis, EECS Department, University of California, Berkeley, 1994, p. 46 ff.

19. Roark, R. J., and Young, W. C., Formulas for Stress and Strain, McGraw-Hill, 1982.]

20. Minami, K., Moriuchi, T., and Esashi, M., "Cavity Pressure Control for Critical Damping of Packaged Micro Mechanical Devices," *Transducers '95,* 1995, Stockholm, Sweden.

21. Monk, D. J., Krulevitch, P., Howe, R. T., and Johnson, "Stress-Corrosion Cracking and Blistering of Thin Polycrystalline Silicon Films in Hydrofluoric Acid," *Symposium on Thin Films: Stresses and Mechanical Properties IV,* MRS Spring Meeting, San Francisco, CA, 5/93.

MEDIA COMPATIBLE PACKAGING AND ENVIRONMENTAL TESTING OF BARRIER COATING ENCAPSULATED SILICON PRESSURE SENSORS

David J. Monk, Theresa Maudie, Dennis Stanerson, John Wertz, Gordon Bitko, Jeanene Matkin, and Slobodan Petrovic
Motorola, Semiconductor Products Sector, Sensor Products Division
Phoenix, AZ 85008

ABSTRACT

Three specific examples of media exposure testing of pressure sensors with barrier coatings are presented. The examples are experiments used to simulate specific automotive and white goods applications. Experimentation was performed on an apparatus that was developed specifically for the exposure of these devices with *in situ* output voltage measurement in organic or aqueous environments. The first example illustrates the swelling of fluorosilicone gels in fuels and establishes a solubility parameter for one fluorosilicone gel between 6 and 8 $(cal/cm^3)^{1/2}$. While exposure to organic solutions has not been observed to cause catastophic failure of fluorosilicone gel filled devices, corrosion is accelerated in subsequent aqueous solution exposure. The second experiment was used to simulate automotive exhaust gases and water by exposing devices to a fuel mixture followed by an acidic solution. The third experiment was performed to study corrosion under parylene coatings during exposure to an alkaline test solution. This process has been modeled as a cathodic delamination phenomenon. Acceleration factor expressions have been estimated considering parylene coating thickness, solution pH and applied device supply voltage as acceleration means. These expressions have been used to evaluate parylene coated pressure sensors against a benchmark lifetime requirement: 2000 hours of exposure in a $NaOH/NaHCO_3$ buffer solution at 55 °C with a pH of 11 and a supply voltage of 5 Vdc applied continuously. For a 1 % failure rate, parylene coated pressure sensors survived approximately 500 hours, whereas an alternative, multiple barrier coating survived over 2000 hours.

INTRODUCTION

Recently, questions about the viability of MEMS devices in the marketplace have begun to arise [**1, 2**]. Lack of adequate infrastructure is considered a major limitation for the future success of MEMS products [**3, 4**]. A significant amount of infrastructure is necessary to develop sensors and actuators into high volume products, including: manufacturing; design, fabrication, packaging, testing, reliability and quality assurance facilities and personnel; marketing and applications; sales; management; etc. Analogies with the traditional IC industry can be used as guidelines for developing these functions; however, the uniqueness of MEMS devices often requires that these areas be addressed specifically for the device being produced. For example, this paper will propose a formal reliability method for characterizing media compatible pressure sensors. Reliability engineering techniques [**5**] are adapted to the media compatibility of pressure sensors by applying a physics-of-failure reliability approach [**6**]. This paper will: 1) provide motivation for addressing this problem, 2) identify the failure mechanisms observed, 3) show examples of using reliability engineering fundamentals to predict lifetime behavior for specific package types in specific media exposure experiments, 4) develop acceleration models for an observed failure mechanism, and 5) compare these results to a customer requirement.

Silicon pressure sensors have been offered by several companies since the late 1970s. A significant portion of the pressure sensor business has been specified for dry, non-corrosive gas environments or for short-term (24 - 48 hours), room temperature saline exposure environments for the medical industry. However, considerable opportunities are available for low-cost devices that can withstand long-term (i.e., several years) media compatible applications. To quantify acceptable long-term media compatibility, a general definition for media compatibility was formulated that can be applied to specific applications: the ability of a pressure sensor to perform its specified electromechanical function over an intended lifetime in the chemical, electrical, mechanical, and thermal environments encountered in a customer's application.

All sensors must provide this capability in at least dry, non-corrosive gas media. Specific markets that require additional media compatibility include automotive and/or white goods. Automotive compatible pressure sensors must survive fuel (i.e., organic) exposure in a temperature range from -50 to 150 °C for underhood applications. In addition, many are required to withstand exposure to salt water and/or strong acids (e.g., NO_x or SO_x and water), which form electrolytic aqueous solutions. Failure mechanisms are often different for exposure to these two classes of solutions.

White goods applications most often require exposure to aqueous solutions. For example, a washing machine application requires that the sensor is compatible with an alkaline solution because of the detergents added to tap water [**7, 8**]. One current specification, for a customer who is interested in a water compatible pressure sensor for another white goods application, has been used as a benchmark for our development purposes. The exposure requirements are 2000 hours in a $NaOH/NaHCO_3$ buffered solution [**9**] with a pH of 11 at 55 °C.

There has been a significant amount of literature, especially patent literature, defining media compatible pressure sensor construction (e.g., [**10-17**]). Structures include backside exposure, designing the silicon to minimize exposure of the electronics, creating packages to minimize the exposure at crucial areas (including silicon-to-substrate bonding techniques), and the use of barrier coatings. Moreover, several companies sell media compatible silicon pressure sensors housed in stainless steel with a silicone oil transfer media and a stainless steel diaphragm. However, very little data exists for quantifying the media compatibility of a given device [**12-14, 18, 19**]. Some of this may be the result of the legality of publishing such data and some may be the result of the difficulty associated with performing such an experiment. Yet, analogous

0-9640024-1-8 $3.00 © 1996 TRF

humidity experiments are the standard for quantifying reliability in IC devices. With this in mind, the goals for this work are: 1) to develop experimental test procedures for sensor exposure in these environments; 2) to determine the failure mechanisms during exposure testing and to model these results; and 3) to develop a product that has customer applicability.

EXPERIMENTAL

Our initial constraint in this work was to use existing silicon bulk micromachined, piezoresistive pressure sensors in the standard Motorola unibody package so that neither existing fabrication nor assembly manufacturing lines were impacted severly. Furthermore, low-cost, high-volume differential pressure sensor devices were required. Low-cost implies no more than approximately a 15 % adder for media compability, and high-volume implies the capability of producing hundreds of thousands of devices each year.

Therefore, to provide media protection, barrier coatings were investigated initially. Specifically, fluorosilicone gels and/or parylene coatings have been used to produce these low cost devices. Several additional coatings were investigated, but they either did not satisfy the cost or the performance targets. Fluorosilicone gels and parylene coatings are significantly different materials, so a brief description of each is necessary.

Silicone gels are one or two part siloxanes that are drop dispensed and require a curing step [**20-22**]. Several vendors provide these materials, each with a slightly different formulation, and each with a slightly different curing mechanism. An example are the two-part fluorinated siloxanes that were used for these experiments. Often, these two-part materials consist of one part with a base polymer and a catalyst and a second part with the same base polymer, a multifunctional moiety, and an inhibitor. When the two parts are mixed and heated, the base polymer reacts (e.g., vinyl addition polymerization) with the multifunction species to form a gel-like network. These materials are deposited in relatively thick coatings (on the order of millimeters); but because they are extremely low modulus materials, they have little effect on the device electrical performance.

Parylene C (poly(monochloro-*para*-xylylene)) coatings are deposited in a chemical vapor deposition chamber [**23-25**]. Dimer starting material is vaporized and then pyrolyzed to create a monomer gas that is drawn into a vacuum chamber containing the sensors. Because of the vacuum in the deposition chamber (approximately 80 mtorr), the material deposits and polymerizes in a very conformal manner. The nature of the near-room-temperature deposition process allows for deposition on fully assembled devices with reasonably repeatable thicknesses. Thickness control is critical because parylene materials are much higher modulus materials than silicone gels, and they do produce a noticeable effect on electrical output. However, the repeatability of these shifts allows modeling of the effect of parylene on device parameters [**26**].

In the present work, experiments are presented to show the effects of organic solutions on gel coated devices, the effect of acidic solutions on gel coated devices, and the effect of alkaline solutions on parylene coated devices.

In the first experiment, various vendor supplied fluorosilicone gel materials used in coating pressure sensor packages were exposed to fuel solutions in a test apparatus that was designed as an automated organic solution exposure chamber (Fig. 1).

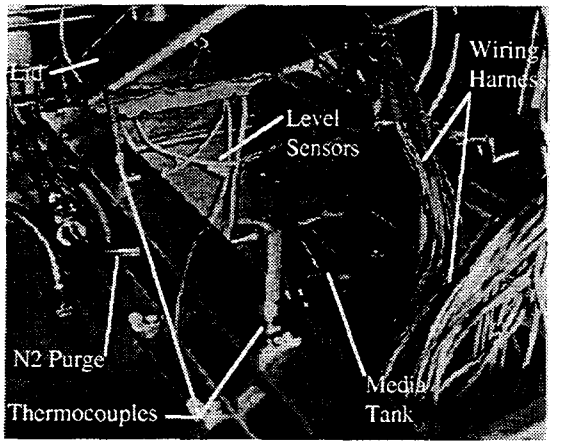


Figure 1. Organic solvent (fuel) media exposure system for monitoring pressure sensor offset in situ.

The test apparatus allows *in situ* offset (i.e., zero applied pressure output voltage) measurements with the ability to provide a supply voltage to the devices continuously. An automated mixing and circulating system is used to concoct various test solutions, often per SAE guidelines [**27**]. For this experiment reference fuel C with 5% corrosive water and fuel C with 20% ethanol and 5% corrosive water were evaluated. The corrosive water consists of NaCl, Na_2SO_4, and $NaHCO_3$ in DI water. This apparatus is equipped with an external reservoir for temperature control from approximately 10 to 90 °C. Most importantly, considerable safety interlocks were designed into the system so that ultimately, if needed, the apparatus could shut itself off unattended. Visual and weight loss measurements were performed before and after these exposure experiments. In addition, manufacturing swelling data has been compared with solubility parameter information on the given solvents to provide an estimate for the solubility parameter of a representative gel [**28**].

In the second experiment, pressure sensors with various fluorogels as barrier coatings have been evaluated during exposure to fuel, acid, and combined fuel and acid solutions. The fuels and acids are known to react with the barrier materials [**29**]. Failure mechanisms expected during the evalution included swelling and dissolution of the polymer, and subsequent corrosion of any exposed metallization. Various fluorosilicone materials were evaluted for each test and lifetime statistics for the various tests computed.

In the final experiment, parylene coated pressure sensors have been used in alkaline exposure experiments. Corrosion under barrier coatings has been observed on several occasions to be the primary failure mode for devices when exposed to electrolytic aqueous solutions. Corrosion under parylene coatings in alkaline solutions was assumed to be analagous to the delamination of paint from metal surfaces and the subsequent corrosion of the metal (e.g., automobile bodies [**30, 31**]). The latter process has been modeled by diffusion of water through the polymer to the metal surface [**31**]. By solving Fick's First Law of Diffusion, with Ohm's Law as a boundary condition, a "delay time", t, prior to the onset of delamination can be established [**31**]:

$$t = \frac{\pi b^2}{4D}\left(1 - \frac{V}{4 R_o DF C_o}\right), \qquad (1)$$

where b is the thickness of the coating, D is the diffusion coefficient of water through the coating, V is the applied voltage, R_o is the resistivity of the coating, F is Faraday's constant, and C_o is the concentration of water at the surface of the coating. Our assumption was that delamination occurs prior to the onset of corrosion, similar to the model proposed by Leidheiser *et al.* (eqn. 1 and **[31]**); therefore, the corrosion failure times observed for these devices should be a function of coating thickness, applied bias, and solution concentration (eqn. 1). Unfortunately, we know very little about the delamination mechanism for parylene coatings in these environments, so we are assuming that the corrosion failure times are proportional to the delay times in the model proposed by Leidheiser *et al.* **[31]**. Therefore, the exposure experiment was designed to examine these variables by using a design of experiments technique **[32]**. Three parylene deposition runs were performed to create sets of devices with three different average parylene thicknesses. A modified Nanometrics Nanospec program was used to measure thickness. All fabrication, assembly, and surface preparation procedures remained constant within the existing manufacturing tolerances to minimize the effects of these variations on this experiment. Exposure of these devices was performed in individual exposure compartments (Fig. 2). All of the closed "cells" were placed in an oven at 55 °C. Continuous applied bias and offset measurements were performed during the 3022 hours of the experiment.

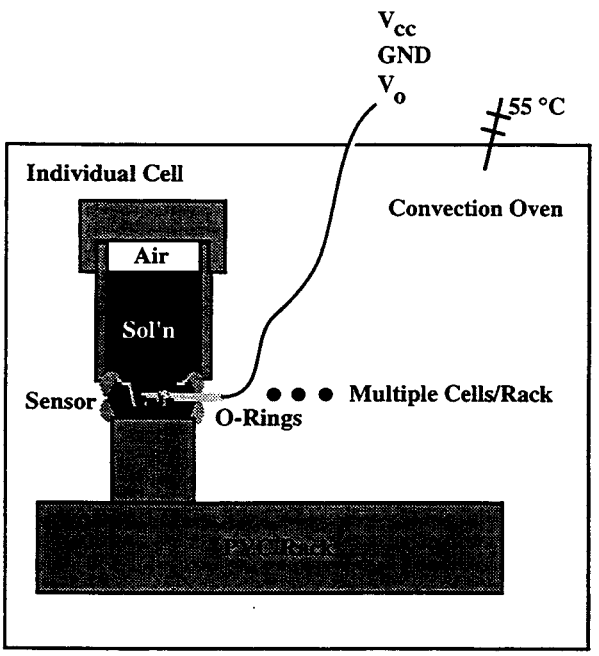


Figure 2. Individual compartments for exposure of barrier coated pressure sensors to aqueous solutions. A rack of 64 of these "cells" is placed into an oven at 55 °C.

RESULTS AND DISCUSSION

A more general discussion of the observed failure mechanisms with these type of devices is presented elsewhere **[33]**. Uniform, galvanic, and local corrosion; silicon etching; polymer swelling and/or dissolution; interfacial permeability (e.g., lead leakage **[34]** and/or adhesive delamination); and mechanical failures (i.e., cracking, creep. fatigue, etc.) all have been observed. In our experiments, polymer swelling and corrosion are observed and modeled in detail.

The swelling of fluorosilicone gels has been analyzed by attempting to determine the solubility parameter of a particular gel. Figure 3 shows a plot of manufacturer published swelling data plotted against the solubility parameter of the particular solvent **[28]**. The solubility parameter is a measure of the interaction energy between species:

$$\delta_i = \left(\frac{\Delta E_i^v}{V_i}\right)^{1/2}, \qquad (2)$$

where ΔE_i^v is the energy of vaporization and V_i the molar volume of the species. Because "like dissolves like", a peak in the curve represents the approximate solubility parameter of the polymer. Therefore, the solubility parameter for this particular fluorosilicone gel is between 6 and 8 $(cal/cm^3)^{0.5}$.

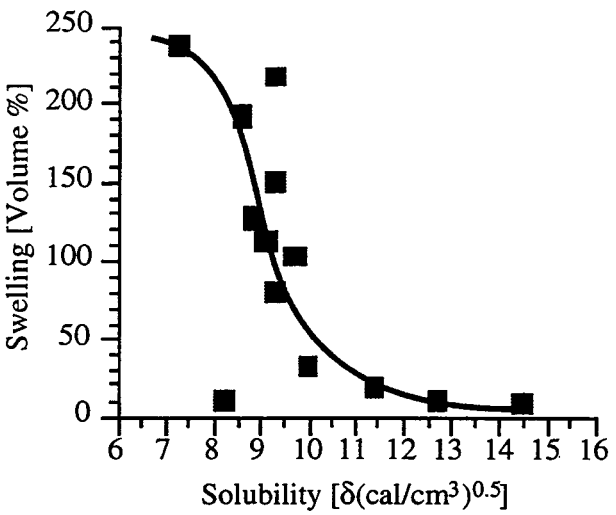


Figure 3. Swelling as a function of solubility parameter for a typical silicone gel using material supplier's swelling data following 8 days of exposure at 20 °C in various solvents and solubility parameters from ref. **[35]**.

In Figure 4, the weight loss, assumed to be caused by loss of low molecular weight species into the solution, is shown for four fluorosilicone gels. The results show that the weight loss for a variety of fluorosilicone gels is not consistent. At least two potential causes for this result can be proposed. First, the extent of reaction of the gels varies, so certain gels have more low molecular weight species that will dissolve into the solution. Second, the solubility parameters of the polymers are different, resulting in a different degree of dissolution. Obviously, both of these can contribute to the final result. Regardless, each gel formulation must be treated as a separate material and

experiments to understand these results more fully are being prepared. The most significant concern with this variability is the potential for subsequent corrosion that results from the exposure, especially if dissolution (i.e., weight loss) is great enough for metal regions to become exposed (e.g., Fig. 5).

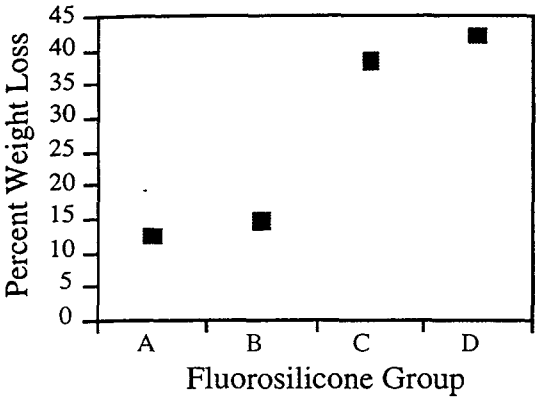


Figure 4. Weight loss measurements on four different fluorosilicone gels after 250 hours of exposure to reference fuel C with 5% corrosive water at 65 °C.

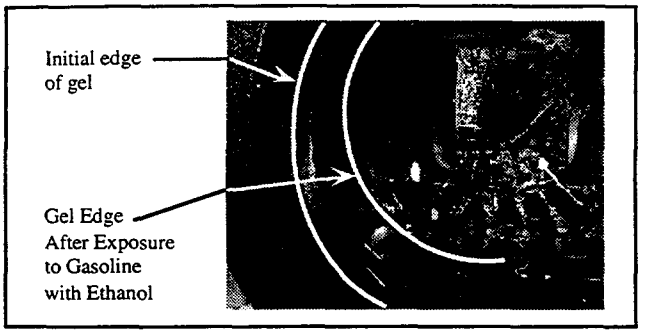


Figure 5. Photograph of a pressure sensor device after extended exposure to fuel containing corrosive water [27], *followed by exposure to a strong acid. Evidence of the gel swelling during the test, and corresponding shrinkage after removal from the test media can be seen by the gel retracting away from the sidewall of the package.*

The second experiment was performed to address this concern. Automotive applications require that sensors survive both organic solutions (e.g., fuels, oils, etc.) and aqueous solutions (e.g., salt water, NO_x + water = HNO_3, SO_x + water = H_2SO_4, etc.). In this experiment, fluorosilicone gel-filled devices were exposed to fuel ([27] for 250 hrs.) and then to nitric acid with a pH of 1.8 at 85 °C, which is a standard specification to simulate under-the-hood operation. Failure lifetimes were recorded with *in situ* offset monitoring. For each group, this lifetime data was assumed to follow a Weibull cumulative failure distribution:

$$R(t) = exp\left[-\left(\frac{t}{\theta}\right)^{\beta}\right]. \qquad (3)$$

Statistical likelihood analysis of the lifetime data was used to determine the θ factors (Table 1). These results can then be used to compare the performance of devices coated with different fluorosilicone gels. However, this technique can be plagued with the same limitations as traditional corrosion testing (i.e., the difficulty of simulating the field environment with a laboratory experiment) unless the environment is extremely well understood. As a supplier of pressure sensors, we are attempting to become as adept as possible at understanding customer applications and failure mechanisms. Assuming the test environment produces the same failure mechanism with similar kinetics, these experiments are an invaluable tool to compare media compatibility schemes and to estimate lifetime in a customer application provided the specifics of that application are known.

Table 1. Weibull Distribution Parameters for Fluorosilicone Gel Coated Pressure Sensors Exposed to Fuel and then to Nitric Acid (pH 1.8 at 85 °C).

Fluorosilicone Group	Characteristic Life, θ [hours]	R(t=48 hours)
A	291	85%
B	51	39%
C	27	16%
D	70	51%

This is the case in the final experimental example. The qualification procedure for a media compatible pressure sensor for a particular white goods customer has been estimated to be 2000 hours of exposure in a $NaOH/NaHCO_3$ buffered solution [9] with a pH of 11 at 55 °C. Therefore, experiments with parylene coated sensors have been performed to 3022 hours with a continuous supply voltage and *in situ* offset monitoring in this environment. Again, failure distributions are modeled assuming a Weibull cumulative failure function. If all of the Weibull slope parameters are assumed to be 3, indicating a "wear-out" failure mechanism (e.g., corrosion under a coating), the characteristic lifetimes can be modeled using analysis of variance to develop a linear expression that takes into account the effects of coating thickness, applied bias, and solution concentration (Fig. 6a):

$$\begin{aligned}\theta(hrs) = {} & (15000 \pm 5800) - (400 \pm 340)\, t_p[\mu m] + \\ & (100 \pm 480)\, V_{cc}[V] - (1400 \pm 500)\, C_o[pH] - \\ & (40 \pm 16)\, V_{cc}^2[V^2] + (40 \pm 31)\, t_p \cdot C_o[\mu m \cdot pH] + \\ & (50 \pm 33)\, V_{cc} \cdot C_o[V \cdot pH] \qquad R^2 = 0.846\end{aligned} \qquad (4)$$

In addition, these distributions can be used to obtain estimates of the 1% failure rate (Fig. 6b) or the failure rate at 2000 hours (Fig. 6c).

From these experimental results, the current parylene processing techniques do not produce a device that will satisfy the requirements of this customer. At a pH of 11 and a 5 V supply voltage, the characteristic lifetime (Fig. 6b) is approximately 450 hours for a 1% failure rate, which was assumed to be our constraint. Devices with fluorosilicone gel encapsulation alone performed similarly to the parylene encapsulated devices.

During these experiments, however, additional devices with a multiple organic barrier coating were observed to survive 3022 hours of exposure without failure. A subsequent experiment with 37 composite coatings with a "worst-case" manufacturing process have been exposed for 550 hours to $NaOH/NaHCO_3$ at pH of 12.5, 55 °C, and 15 Vdc without

incident. By using a conservative acceleration factor of 4X from equation 4 (or Fig. 6), this represents approximately 2200 hours of exposure in this environment. This estimate assumes that these devices fail with the same mechanism as the parylene encapsulated devices used to develop equation 4.

Finally, additional techniques, to be presented elsewhere, are being developed to obtain a better understanding of why this combination of materials provides an enhanced barrier coating, and to determine if it still provides acceptable electrical performance. It is postulated that reduced amount of corrosive reactants and/or temperature at the corrosion sites prior to reaching the steady state may affect the results. In addition, perhaps the upper coating minimizes delamination of the of the lower coating, which is required before corrosion can begin. Further exposure testing is being performed to repeat these results and to better understand the barrier coating failure mechanism.

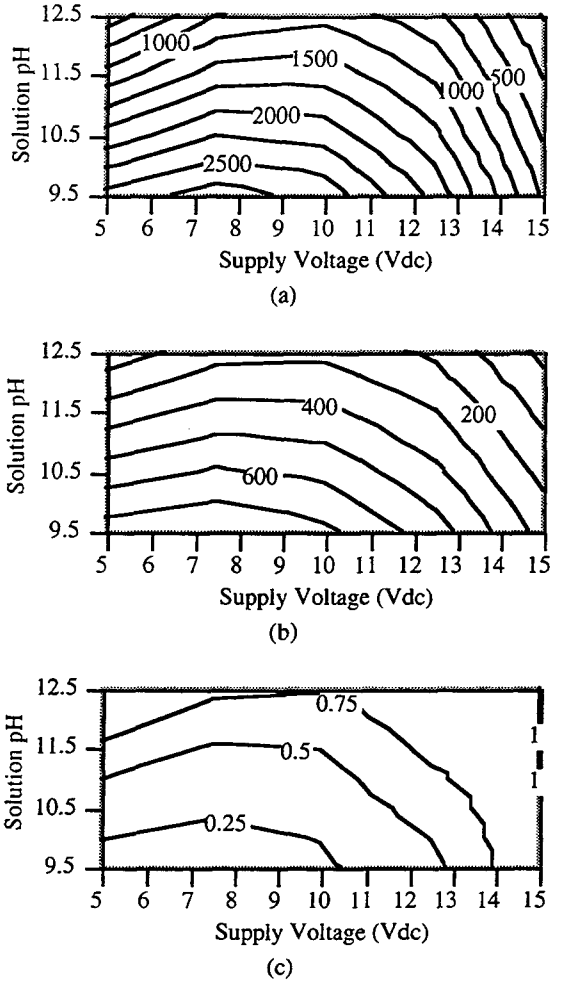


Figure 6. Failure analysis for ≈ 5 μm parylene C coated pressure sensors exposed to $NaOH/NaHCO_3$ at 55 °C with continuous supply voltage. (a) Weibull characteristic life, θ, as a function of supply voltage and solution pH. (b) Characteristic lifetime for a failure rate of 1% as a function of supply voltage and solution pH. (c) Failure rate at 2000 hours as a function of supply voltage and solution pH.

CONCLUSIONS

Several experimental test techniques have been developed for performing simulation of customer media applications. Separate organic (i.e., fuel, oil, etc.) and aqueous (i.e., acid, base, salt water, etc.) systems with continuous supply voltage, temperature control, *in situ* offset monitoring, and significant safety precautions have been developed. Complementary analytical techniques have been initiated to gain a better understading of the failure mechanisms. Failure mechanisms, including: corrosion, silicon etching, polymer swelling/dissolution, interfacial permeability, and mechanical failures have been observed with barrier coated silicon bulk micromachined, piezoresistive pressure sensors. The ultimate goal of this work is to develop an industry standard analogue to IC reliability for media compatible pressure sensors.

Three specific examples have been presented to illustrate the effects of polymer swelling/dissolution of polymer coatings in organic solutions and corrosion under polymer coatings in aqueous solutions. The polymer solubility parameter was used to begin to predict the swelling/dissolution failure mechanism. The corrosive failure mechanisms were being studied by analyzing failure distributions to allow comparison of media compatibility schemes and to benchmark development products with customer requirements. These environmental exposure simulation experiments require that the customer environment and the failure mechanism resulting from that requirement are very well understood. In the automotive environment, polymer swelling/dissolution from exposure to organic solutions (e.g., fuels) may not result in a catastrophic failure, but the subsequent acid exposure, used to simulate exhaust gases and water, has been observed to cause corrosive failures more rapidly than was observed on devices that have not been exposed to the organic solution.

It was also demonstrated that corrosion under parylene coatings in an alkaline environment for white goods applications is a function of polymer thickness, supply voltage, and solution concentration, as predicted by solving the Fickian diffusion problem for this type of system. This experiment showed that neither parylene nor gel coatings alone provided a barrier that was sufficient to satisfy all lifetime requirements. However, if it can be assumed that the same failure mechanism governs the failure of devices encapsulated with a multiple organic coating scheme, then the results indicated that a combination of barrier coatings will satisfy long term alkaline media exposure requirements.

ACKNOWLEDGMENTS

The authors are grateful to several members of Motorola Semiconductor Products Sector, including: the MEMS1 fabrication facility for providing die, the MKL and SPD Prototype assembly areas for packaging the die, the SPD Characterization Lab for testing devices, the Product Analysis Lab and Chemical and Surface Analysis Labs for helping with the analysis of failed devices, and the SPD Marketing group for helpful discussions concerning the prospects of developing a low-cost, high-volume, media compatible pressure sensor.

REFERENCES

[1] F. Maseeh, "Lack of Manufacturing Know-How is Slowing the MEMS Revolution", *Solid State Tech., 38*, 50, (1995).

[2] F. Maseeh, "The Sensor-Based MEMS Market: An Underachiever?", *Sensor Business News, 2*, (1995).

[3] S. Horntvedt, "Industrialization of an R & D Sensor Company", The 8th International Conference on Solid-State Sensors and Actuators, and Eurosensors IX, Stockholm, Sweden, (1995), 902.

[4] K. Petersen, "MEMS: What Lies Ahead?", The 8th International Conference on Solid-State Sensors and Actuators, and Eurosensors IX, Stockholm, Sweden, (1995), 894.

[5] K. C. Kapur and L. R. Lamberson, "Reliability Estimation: Weibull Distribution", in *Reliability in Engineering Design*. New York: John Wiley & Sons, (1977), 291.

[6] J. M. Hu, "Physics-of-Failure-Based Reliability Qualification of Automotive Electronics", *Communications in RMS, 1*, 21, (1994).

[7] T. Maudie and J. Wertz, "Semiconductor Sensor Performance and Reliability for the Appliance Industry", International Appliance Technical Conference, Champaign/Urbana, IL, (1995),

[8] T. Maudie, "Improving Reliability of Electronic Appliance Pressure Sensors", *Appliance*, 39, (1995).

[9] D. R. Lide, "CRC Handbook of Chemistry and Physics", 75 ed. Boca Raton, FL: CRC Press, (1995), .

[10] A. D. Kurtz and J. R. Mallon, "Media Compatible Pressure Tranducer", U.S. Patent 4,222,277, September 16, 1980.

[11] W. P. Eaton and J. H. Smith, "A CMOS-Compatible, Surface Micromachined Pressure Sensor for Aqueous Ultrasonic Applications", Micromachining and Microfabrication '95, Austin, TX, SPIE (1995), 258.

[12] A. Hanneborg, M. Nese, and P. Øhlckers, "Silicon-to-Silicon Anodic Bonding with a Borosilicate Glass Layer", *J. Micromech. Microeng., 1*, 139, (1991).

[13] A. Hanneborg, M. Nese, H. Jakobsen, and R. Holm, "Silicon-to-Thin Film Anodic Bonding", *J. Micromech. Microeng., 2*, 117, (1992).

[14] M. Nese and A. Hanneborg, "Anodic Bonding of Silicon to Silicon Wafers Coated with Aluminum, Silicon Oxide, Polysilicon or Silicon Nitride", *Sensors and Actuators A, 37-38*, 61, (1993).

[15] M. Noble, "Environmental Concerns for Integrated Circuit Sensors", *Measurement + Control, 19*, 210, (1986).

[16] K. Ryan, J. Bryzek, and R. Grace, "Flexible Miniature Packaging from Lucas NovaSensor Provides Low-Cost Pressure Sensor Solutions to a Wide Variety of Hostile Media Applications", *ECN*, 118, (1995).

[17] S. Shoji, T. Nisase, M. Esashi, and T. Matsuo, "Fabrication of an Implantable Capacitive Type Pressure Sensor", The 4th International Conference on Solid-State Sensors and Actuators (Transducers '87), (1987), 305.

[18] T. Maudie, "Testing Requirements and Reliability Issues Encountered with Micromachined Structures", Proceedings of the Second International Symposium on Microstructures and Microfabricated Systems, Chicago, IL, ECS (1995), 223.

[19] A. Nakladal, K. Sager, and G. Gerlach, "Influences of Humidity and Moisture on the Long-Term Stability of Piezoresistive Pressure Sensors", *Measurement, 16*, 21, (1995).

[20] A. Polak, "Characterization and Use of Polyfluorosiloxanes in Automotive Applications", Materials Research Society Symposium Proceedings: Electronic Packaging Materials Science VI, San Francisco, CA, MRS (1992), 281.

[21] W. Fabianowski, D. Jaffe, and R. J. Jaccodine, "Mechanisms of Microelectronics Protection with Polysiloxane Polymers", 6th International SAMPE Electronics Conference, SAMPE (1992), 256.

[22] C. P. Wong, "Recent Advances in the Application of High Performance Siloxanes Polymers in Electronic Packaging", 6th International SAMPE Electronics Conference, SAMPE (1992), 508.

[23] W. F. Beach, T. M. Austin, and R. Olson, "Parylene Coatings", in *Electronic Materials Handbook™: Volume 1: Packaging, 1*, M. L. Minges, C. A. Dostal, and M. S. Woods, Eds. Materials Park, OH: ASM International, (1989), 789.

[24] W. F. Beach, C. Lee, D. R. Bassett, T. M. Austin, and R. Olson, "Xylylene Polymers", in *Encyclopedia of Polymer Science and Engineering, 17*, 2nd ed: John Wiley & Sons, Inc., (1989), 990.

[25] W. F. Gorham, "A New, General Synthetic Method for the Preparation of Linear Poly-*p*-xylylenes", *J. Polym. Sci., 4*, 3027, (1966).

[26] D. J. Monk and M. Shah, "Thin Film Polymer Stress Measurement Using Piezoresistive Anisotropically Etched Pressure Sensors", Mat. Res. Soc. Symp. Proc., San Francisco, CA, MRS (1995), 103.

[27] S. C. R. Report, "Gasoline/Methanol Mixtures for Material Testing", Motorola Internal Report (SAE: 1990).

[28] F. Rodriguez, *Principles of Polymer Systems*, 2 ed. Washington: Hemisphere Publishing Corporation--McGraw-Hill Book Company, (1982).

[29] L. D. Fiedler, T. L. Knapp, A. W. Norris, and M. S. Virant, "Effect of Methanol/Gasoline Blends at Elevated Temperature on Fluorosilicone Elastomers", Society of Automotive Engineers International Congress and Exposition, SAE (1990).

[30] J. Leidheiser, H., "Mechanisms of De-adhesion of Organic Coatings from Metal Surfaces", Polymeric Materials for Corrosion Control, Chicago,IL, American Chemical Society, Washington,DC (1985), 125.

[31] W. Wang and J. Leidheiser, H., "A Model for the Quantitative Interpretation of Cathodic Delamination", Equilibrium Diagrams Localized Corrosion, ECS (1983), 255.

[32] D. C. Montgomery, *Design and Analysis of Experiments*, 3rd ed. New York: John Wiley & Sons, (1991).

[33] T. Maudie, D. J. Monk, D. Zehrbach, and D. Stanerson, "Sensor Media Compatibility: Issues and Answers", Sensors Expo, Anaheim, CA, (1996).

[34] D. J. Monk, "Pressure Leakage through Material Interfaces in Pressure Sensor Packages", Sensors in Electronic Packaging, San Francisco, CA, ASME (1995), 87.

[35] E. A. Grulke, "Solubility Parameter Values", in *Polymer Handbook*, J. Brandrup and E. H. Immergut, Eds., 3rd ed. New York: John Wiley & Sons, (1989), VII/519.

SELF-ASSEMBLED MONOLAYER FILMS AS DURABLE ANTI-STICTION COATINGS FOR POLYSILICON MICROSTRUCTURES

Michael R. Houston, Roya Maboudian, and Roger T. Howe*

Berkeley Sensor & Actuator Center
Department of Chemical Engineering
Department of Electrical Engineering and Computer Sciences*
University of California at Berkeley, Berkeley, CA 94720 USA

ABSTRACT

We present a self-assembled monolayer (SAM) coating process for high-yield release of compliant polysilicon microstructures – direct from the water rinse. This process has been used to release cantilever beams up to 1000 microns in length and requires no special sublimation or supercritical drying techniques. Once coated and released, these same microstructures have been used to quantify in-use stiction by electrostatically actuating surfaces into contact. Polysilicon structures coated with SAM films exhibit a work of adhesion of 3 $\mu J/m^2$, nearly four orders of magnitude lower than uncoated, hydrophilic structures. These coatings are also thermally stable up to 400 °C for 5 minutes in a nitrogen or vacuum ambient, which falls within the range of typical packaging temperature cycles. The SAM-coating process has been further improved to eliminate the use of carbon tetrachloride, an environmentally unfriendly solvent. Using a combination of chloroform and iso-octane in place of the carbon tetrachloride, films have been produced which exhibit equal or better stiction properties than those formed using the carbon tetrachloride process.

INTRODUCTION

Despite years of advancement in bulk and surface micromachining technologies, stiction of microstructures remains a fundamental problem. Stiction most often manifests itself during the structure release process, where liquid capillary forces deflect compliant structures into contact with adjacent surfaces. Subsequent evaporative drying leads to solid bridging due to precipitates left behind from the liquid phase, which usually results in permanent sticking at this stage [1-3]. This *release-related stiction* has been successfully evaded by several methods, including sublimation drying [1,4], supercritical drying [5], and dry-release [6,7] techniques that avoid the formation of a liquid-vapor interface altogether. In addition, small features have been added to the microstructure perimeter to reduce capillary forces by shaping the liquid meniscus [8].

The second manifestation of stiction is during actual device operation, termed *in-use stiction*. When adjacent microstructure surfaces come into contact (intentionally or unintentionally), capillary, electrostatic, and van der Waals forces may all contribute to adhesion of the surfaces and device failure [2]. None of the techniques mentioned above for the alleviation of release stiction addresses the problem of in-use stiction, which is becoming more important as increasingly compliant microstructures and vacuum packaging are introduced [9,10]. Initial attempts to alleviate in-use stiction focused on reducing the contact area between adjacent surfaces by fabricating micro-dimples [11,12] or by roughening the polysilicon surface on a more microscopic level [13,14]. While some improvements (factor of ~5) were seen due to roughening, attention has turned more recently towards chemical passivation of the polysilicon surface.

One approach to surface passivation introduced self-assembled monolayers (SAMs) into microfabricated devices. Alley *et al.* showed promising results for alleviating release-related stiction and suggested SAMs as a potential post-release anti-stiction lubricant [2]. However, due to the stiffness of the structures investigated, no quantitative measure of the stiction properties of SAM-coated surfaces could be made at that time. Further work involving SAM films by Deng *et al.* showed a marked improvement in the release and wear properties of SAM-coated polysilicon wobble motors. Motors with a SAM coating operated for nearly 80 million cycles over a nine month period, indicating low wear and good durability of these films even on surfaces in continuous moving contact [15].

More recently, a method for measuring the work of adhesion between contacting surfaces using cantilever beam arrays showed that hydrogen passivation of the polysilicon surface reduced in-use stiction by over three orders of magnitude [16]. This technique made use of an ammonium fluoride treatment coupled with supercritical drying to release microstructures with a hydrogen-passivated polysilicon surface. One drawback to this approach, however, was the metastable nature of the hydrogen-terminated silicon surface which limited the long-term usefulness of such a treatment [17]. Hydrogen-terminated silicon surfaces are likely to be further degraded by packaging processes which typically require temperatures of 300-425 °C, a stringent requirement for surface passivation coatings. In another approach recently introduced, fluorocarbon films formed by a plasma polymerization reaction have shown success in reducing post-release adhesion [18].

A low-stiction, chemically-stable surface coating capable of surviving packaging temperature cycles is highly desirable. This work demonstrates that the SAM coating process previously studied by Alley and Deng is effective at reducing both release and in-use stiction. The SAM film is comprised of a close-packed array of alkyl chains which spontaneously form on oxidized silicon surfaces, and have been shown to remain stable indefinitely (up to 18 months in air to date) in numerous environments [19]. The extremely hydrophobic nature of SAM-coated surfaces changes the meniscus shape such that microstructures are not collapsed by capillary forces. As a result, extremely compliant structures may be pulled directly out of water, from which they emerge released and completely dry without post-processing. Furthermore, these films reduce in-use stiction by several orders of magnitude, and their thermal stability is sufficient to withstand various packaging processes. A description of the formation/release process, thermal stability, and in-use stiction properties of SAM-coated polysilicon microstructures will be presented in the following sections.

0-9640024-1-8 $3.00 © 1996 TRF

BASICS OF SAM FORMATION

The films studied in this work were formed from octadecyltrichlorosilane (OTS) precursor molecules having a chemical formula of $C_{18}H_{37}SiCl_3$. When the precursor molecules are placed into a suitable solvent such as a 4:1 hexadecane:carbon tetrachloride (CCl_4) mixture and an oxidized silicon surface is inserted, the trichlorosilane head groups hydrolyze and chemisorb to the surface, forming Si-O-Si links to the surface [20]. The process showing assembly of the film is depicted schematically in Fig. 1. After formation, the film is further stabilized by cross-linking between chains, a process often aided by a low temperature anneal after film formation (~100 °C). When prepared correctly, a SAM-coated Si(100) oxide surface (prepared using OTS) exhibits a water contact angle of 114° ± 2°, which corresponds to a surface energy of about 20 mJ/m^2. The film thickness is about 25 Å, and the film "density" is approximately 5 chains per nm^2 (~90% of the density of crystalline polyethylene), indicative of a true "monolayer" coverage [21].

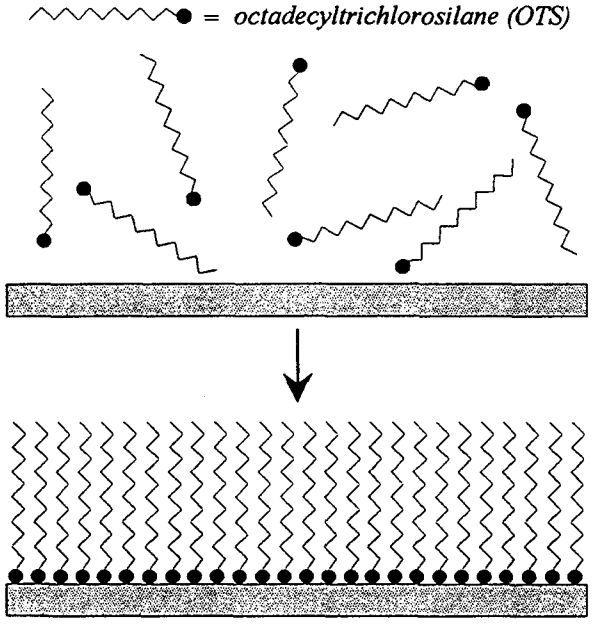


***Figure 1.** Schematic of self-assembled monolayer formation on an oxidized silicon surface using alkyl-trichlorosilane precursor molecules. Film formation is thought to occur within several minutes on smooth surfaces.*

There are two particularly important factors involved in obtaining high-quality SAM films [21]. First, the presence of water in the SAM solution is undesirable because water will cause conglomeration of the OTS precursor molecules. Water is thought to be *necessary*, however, in the form of a thin film on the silicon dioxide surface. This water is responsible for hydrolyzing the trichlorosilane head groups and may promote mobility of the SAM chains, allowing them to move into close-packed structures. Thus, while bulk water must be washed away before placing samples into the SAM solution, hydrophilic silicon surfaces which retain a thin layer of water molecules are desirable for high-quality SAM films formed from these precursor species.

The second important factor in obtaining high-quality films is the temperature at which the films are formed. Specifically, a transition temperature has been observed near room temperature, below which the SAM films display superior properties [22]. In general, this transition temperature drops as the tail chain length becomes shorter. For OTS, the transition temperature is reported to be approximately 28 °C. The existence of such a transition temperature also appears to have important ramifications for forming SAM films from other precursor species, such as fluorinated trichlorosilanes (FTS) [22]. This may be one explanation for the poor performance seen in FTS films studied thus far for micromachining anti-stiction applications [2,15].

One remarkable property of these films is their stability after formation [19]. Specifically, storage times of up to 18 months in air produced indistinguishable thicknesses and contact angles from freshly prepared films. Furthermore, SAM films appear to be impervious to hot organic solvents, acidic media, and boiling water. Even rubbing of the surface with a tissue produced no detectable changes in the film properties, further evidence of film durability. The long-term stability of these film in various ambients, coupled with their low surface-energy, hydrophobic, densely-packed structure, makes self-assembled monolayers an exciting option for anti-stiction coatings in micromachines.

EXPERIMENTAL PROCEDURES

The structures used to characterize the release and in-use stiction properties of the SAM coatings in this work consisted of arrays of 10 μm-wide cantilever beams, which extended from 50 to 1000 microns in length. Between 50 and 260 microns, the beams were incremented in 10 micron lengths; the longer beams were incremented by 20 microns. These structures were fabricated using standard surface micromachining techniques. After thermal oxide growth and low-stress nitride deposition, a 3000 Å in-situ doped polysilicon film was deposited and patterned as the ground-plane layer. The deposition conditions were 590 °C, 500 mtorr, PH_3 mole fraction = $3.2x10^{-3}$, balance SiH_4. A 2 μm PSG layer, deposited and annealed at 1000 °C for one hour, served as the sacrificial layer. After PSG patterning, a 2 μm structural polysilicon layer was deposited under the same conditions listed above, except the deposition temperature for this layer was 585 °C. These polysilicon deposition conditions, followed by an RTA anneal at 900 °C after patterning, have been shown to produce low stress, low strain-gradient polysilicon films [23], which were necessary to obtain the long, *flat* beams (up to 1000 microns) used in this work.

In-use stiction was measured by determining the cantilever beam detachment length and calculating a work of adhesion for the contacting surfaces [3,16,24]. Testing in-use stiction with the cantilever beam arrays required actuation of the beams down into contact with the underlying ground-plane surface. A sticking probability, defined as the ratio of the number of beams stuck to the total number of beams actuated, was calculated for each beam length and surface treatment examined.

For hydrophilic surfaces, tested as a reference for this work, beam actuation was accomplished by pushing the beams down with a sharp probe tip, then releasing them. However, adhesion between the probe tip and the beam surface precluded probe-tip actuation for the hydrophobic SAM-coated structures. A second generation of structures was therefore designed and built which incorporated an electrostatic actuation pad, as shown in Fig. 2. By applying 150 V between the actuation pad and the beams/landing pad, the beams could be pulled into contact with the landing pad without making contact with the actuation pad. By utilizing a single actuation pad, an entire beam array could be tested at the same time; the beam arrays were typically pulled down and released several times before results were recorded. It was also

important to record the results immediately after testing, since transportation of the samples was observed to cause beams to become unstuck in some cases. Using a voltage of 150 V, beams as short as about 400 microns could be pulled into contact; higher voltages resulted in contact and shorting problems with the actuation pad.

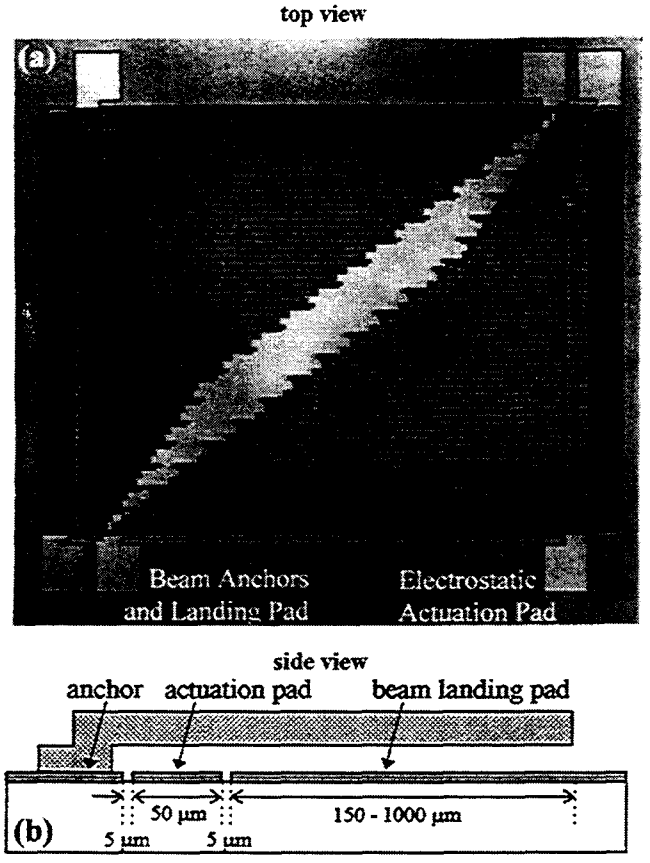


Figure 2. Top view picture (a) and side view schematic (b) of electrostatically-actuated cantilever beam arrays. Polysilicon structures were built with a 3000 Å ground-plane layer and a 2 μm structural layer in a 3-mask process. After fabrication, individual die were annealed in argon for 5 minutes at 900-1000 °C to obtain low stress, low strain-gradient films.

Thermal stability experiments were carried out at atmospheric pressure by heating SAM films in an rapid thermal annealing system [AG Associates Heatpulse 210] to the specified temperature. For films annealed in nitrogen or oxygen, the sample chamber was purged for five minutes with 50 sccm of flowing gas. Vacuum experiments were conducted in a vacuum packaging station with a glass chamber pumped down to below 10^{-4} torr. Specified annealing temperatures were achieved with an IR lamp and were measured by a thermocouple mounted to the back of the sample holder. Contact angle measurements employed the Sessile drop method and used DI water (resistivity ≥ 18 MΩ) as the measurement liquid. Liquid drops were typically 4 μl, and the contact angle was measured with a goniometer [Ramé-Hart, Inc.].

SAM depositions were carried out using standard solvent-grade chemicals, except for hexadecane and iso-octane which were anhydrous grades (Aldrich Chemical Co.). The OTS was purchased from Aldrich at 95% purity, and was used as-received. The mixture used for SAM deposition was a ~1 mM OTS solution in 4:1 hexadecane:CCl_4 or hexadecane:$CHCl_3$ - about 1 drop (~20 μl) of OTS per 50 ml of solvent. The SAM deposition and the peroxide soak were 15 minutes, and rinse times were typically 10 minutes. All steps in the process consisted of dilution rinses in order to keep the structures from leaving the liquid phase. Rinses were accomplished either by simple dilutions or by moving structures from one solution to the next while entraining a liquid dome on the wafer surface. All glassware was dehydrated in a 90 °C oven prior to use, and reactions were carried out in a chemical fume hood.

RESULTS AND DISCUSSION

Release Stiction. The microstructure release process incorporating SAM films is given in Table 1. Once structures were freed in HF (diluted 5:1 with water), the polysilicon surfaces had to be reoxidized for the SAM formation to occur. This was accomplished by forming a chemical oxide with hydrogen peroxide. Before immersion in the OTS solution, the bulk water had to be removed from the samples. Hence, samples were rinsed with iso-propanol (IPA) then carbon tetrachloride (CCl_4) prior to entering the OTS mixture. After SAM formation, the excess precursor molecules were rinsed from the surface with CCl_4, and the CCl_4 was rinsed away with IPA. The release process up to this point was adapted from Alley *et al.* who used methanol instead of iso-propanol, and allowed the samples to dry under an IR lamp after the last solvent rinse [2].

A major concern regarding this work and its industrial application is the use of carbon tetrachloride in the SAM coating process. This is because CCl_4 was banned for most industrial uses under the Montreal Protocol. The role of CCl_4 in the SAM process is to help dehydrate the structures prior to film formation, help solvate the OTS head group in the solvent mixture, and to rinse away excess OTS molecules which may be physisorbed to the surface. Miscibility and wettability issues are also important in choosing a replacement solvent, since liquid must be entrained when moving from one solution to the next. Very recently we found that a combination of chloroform and iso-octane can also be used to produce high-quality SAM films without the use of CCl_4. In this process, iso-octane replaces CCl_4 for the rinses before and after SAM formation, and chloroform is substituted for the CCl_4 in the OTS mixture. Structures coated using these new solvents have shown equal or better performance compared to the films deposited using CCl_4, and the results reported in this work are applicable to both deposition techniques.

Table 1. Summary of the release and SAM-coating process for polysilicon microstructures. Numbers in parentheses indicate that that rinse was performed twice in succession. See text for details. Note: Failure to rinse away all of the excess OTS molecules between steps (iii) and (iv) will lead to residue and/or polymer formation on the structure surfaces! This problem is particularly prevalent if the OTS concentration is too high (>5 mM) during the SAM formation step.

i) Sacrificial Layer Etch	Dilute HF (5:1 v/v H_2O)
ii) Surface Oxidation	H_2O rinse, H_2O_2 soak, H_2O rinse
iii) SAM Formation	IPA rinse (2), CCl_4 rinse (2), OTS solution, CCl_4 rinse (2)
iv) Rinse and Dry	IPA rinse (2), H_2O rinse, pull structures directly from water!

We have refined Alley's release process a step further by taking advantage of the extremely hydrophobic SAM-coated surfaces. At the end of the release process, instead of drying the structures from the IPA (or methanol) rinse, the IPA was instead displaced by water. The structures were then pulled directly out of the water, where they emerged released and completely dry. The structures required no further drying because water dewetted the SAM-coated surfaces immediately as the structures came out of the water. This release technique has been used to successfully release cantilever beam arrays containing beams up to *1 mm* long, as pictured in Fig. 3. Other devices such as Sandia microengines [25] have also been coated and released successfully.

While many surfaces will dewet when pulled from water (e.g., an HF-treated silicon surface, contact angle ≈ 70°), it has been our experience that compliant, uncoated structures pulled from an HF or water solution in the same manner as described above will invariably emerge stuck to the substrate or to adjacent microstructure surfaces. We believe the success of the SAM treatment in this regard results from the water contact angle of ~114° present on these surfaces. Figure 4 shows a simplified picture of the water droplet shape and the corresponding meniscus shape for surfaces with contact angles above and below 90°. On the more hydrophilic surface, the meniscus shape will be concave underneath a compliant structure, creating a capillary force sufficiently strong to pull these surfaces into contact. Alternatively, on a SAM-coated structure, the water meniscus will maintain a convex curvature, thus giving rise to no attractive forces between the two surfaces.

The success of the release process outlined in Table 1 appears to depend on the extreme hydrophobicity (i.e. low surface energy) of the SAM-coated oxidized silicon surfaces. Since the dewetting of the microstructures is a dynamic process (i.e. not an equilibrium process), the receding contact angle will actually determine the meniscus shape during dewetting. Fortunately, there is little hysteresis in advancing and receding contact angles on high-quality SAM films [21], and receding contact angles remain ~114°. Viscous forces due to water streaming will also act on the microstructures when they are pulled from the water rinse; however, these viscous forces have not led to significant stiction problems on the structures released to date.

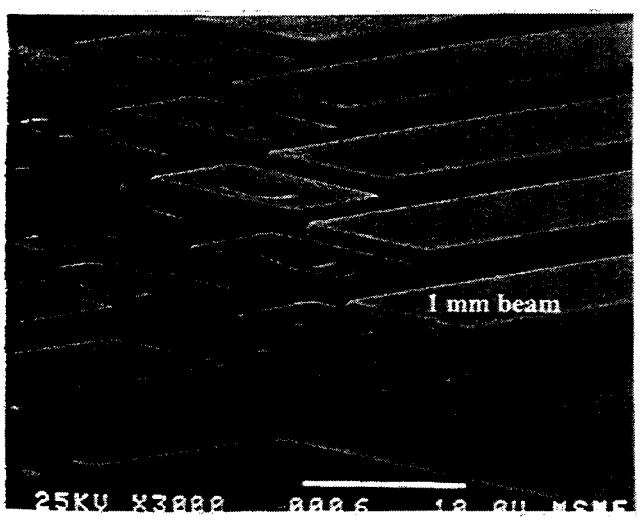


Figure 3. Scanning electron micrograph of SAM-coated cantilever beams released using the process outlined in Table 1. Even the longest beam (1 mm) was successfully released without sublimation or supercritical drying techniques.

In-Use Stiction. After coating and releasing the cantilever beam arrays described above, these structures were then used to quantify the in-use stiction properties of SAM-coated devices. Upon electrostatically actuating the cantilever beams into contact with the underlying substrate, the status of each beam (stuck or unstuck) was recorded. The sticking probability data for three different microstructure surfaces is shown in Fig. 5, where the sticking probability is plotted as a function of beam length. For the H_2O_2 and NH_4F-treated structures [see Ref. 16], the sticking probability curves cross the 50% line at a beam length of approximately 110 and 680 microns, respectively. The sticking probability data for the SAM-coated surfaces is also shown in Fig. 5, and it can be seen that even at a beam length of 950 microns, only about half of the beams remain stuck upon making contact with the substrate.

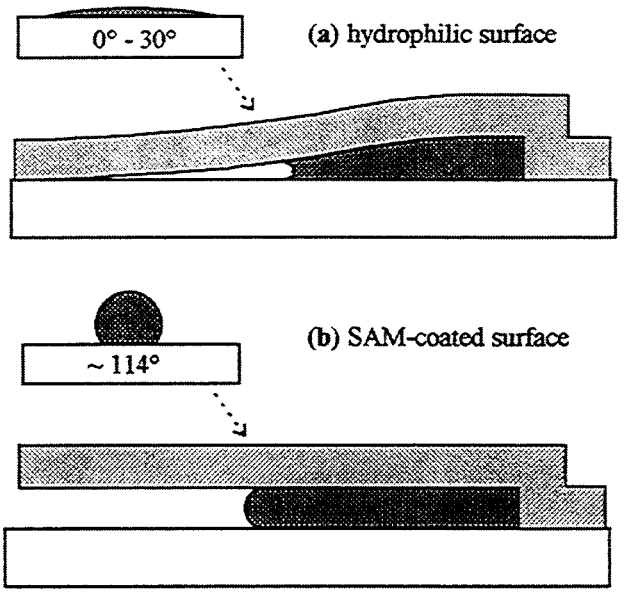


Figure 4. Water contact angles and corresponding droplet shape on hydrophilic oxide (a) and hydrophobic SAM-coated (b) surfaces. Underneath the hydrophilic beam, the concave meniscus creates a capillary force which acts to bring the cantilever beam into contact with the substrate. No such attractive capillary force exists underneath the SAM-coated beam, however, where the meniscus maintains a convex curvature.

If the detachment length is taken as the length at which the sticking probability is equal to 50%, work of adhesion values can then be estimated using the equation [3]

$$W_s = \frac{3}{8}\frac{Et^3h^2}{l_d^4}. \qquad (1)$$

Here, E is the Young's modulus of polysilicon, t is the polysilicon thickness, h is the sacrificial layer thickness, and l_d is the detachment length. Using Eqn. (1), work of adhesion values for each of the treatments are calculated and shown in Table 2, which reveals a dramatic decrease in the work of adhesion on hydrophobic surfaces. Specifically, the SAM-coated structures exhibit a work of adhesion of 3 $\mu J/m^2$, compared to 20,000 $\mu J/m^2$ for structures with conventional hydrophilic surfaces. This corresponds to a reduction of nearly four orders of magnitude in work of adhesion.

The mechanism for stiction reduction is not yet completely understood. However, it seems likely that the adhesion phenomena are quite different in the case of the hydrophilic and hydrophobic surfaces. One would suspect that meniscus forces, driven by capillary condensation, are responsible for the relatively high adhesion seen on the hydrophilic samples. This suggestion is supported by the observation that under conditions of near 100% relative humidity, the detachment length for the hydrophilic surface decreases to about 65 microns, which translates to a work of adhesion of 140 mJ/m^2 - twice the surface tension of water [3,24].

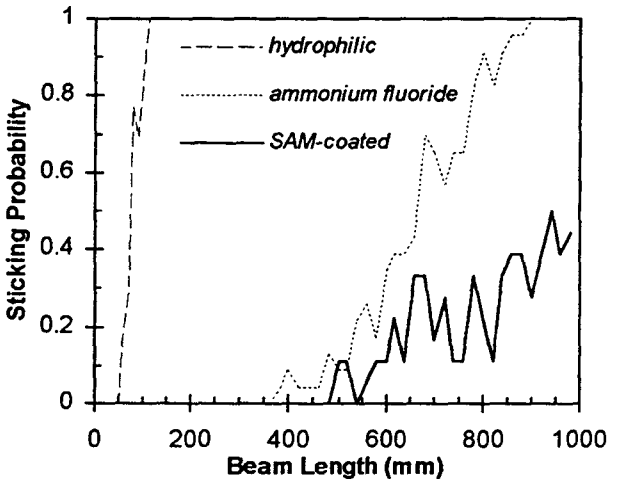


Figure 5. Comparison of in-use stiction properties on hydrophilic, ammonium fluoride-treated, and SAM-coated surfaces. The structures used for all measurements were fabricated during the same processing run.

A contact angle greater than 90°, as seen on SAM-coated surfaces, prevents capillary condensation and the resulting meniscus force altogether. The attractive forces present in SAM-coated structures are therefore expected to originate from an entirely different adhesion mechanism. Adhesion energies on the hydrophobic surfaces appear to be on the order of van der Waals interactions, but surface roughness complicates a more quantitative calculation. Work is in progress to predict these adhesive forces based on the surface topology as measured by atomic force microscopy [24].

Thermal Stability. Figure 6 shows the water contact angles on polysilicon SAM-coated surfaces as a function of anneal temperature. These samples were annealed in either a nitrogen, oxygen, room air, or vacuum environment. For the samples annealed in nitrogen, the contact angles remained constant to 400 °C, above which the contact angle began to decrease. Annealing experiments in vacuum were carried out to mimic an actual vacuum packaging process, and SAM films remained intact up to at least 400 °C (the highest attainable temperature in our packaging system). Thus, it appears that the SAM thermal stability in vacuum or inert atmospheres is adequate for some packaging processes. To our knowledge, the thermal breakdown mechanism in the SAM films is not presently known. However, the film stability is much lower when annealed in oxygen or air, as shown in Fig. 6, indicating a mechanism for chemical reaction at elevated temperatures in oxygen-containing environments. These results appear to be consistent with literature reports of film stability in inert atmospheres up to around 450 °C [26] and film stability in air up to at least 140 °C [27].

Table 2. Approximate detachment lengths and corresponding works of adhesion for each of the treatments listed in Fig. 5. Works of adhesion were calculated using Eqn. (1) with a polysilicon thickness, t, of 2.15 μm, a sacrificial layer thickness, h, of 2.0 μm, and a Young's modulus of 170 GPa.

Surface Treatment		Detachment Length (μm)	Work of Adhesion ($\mu J/m^2$)
H_2O_2	hydrophilic	110	20,000
NH_4F	hydrophobic	700	10
SAMs	**hydrophobic**	**950**	**3**

CONCLUSIONS AND FUTURE DIRECTIONS

This work has systematically investigated the anti-stiction properties of OTS self-assembled monolayer films. By converting the high surface-energy, hydrophilic silicon dioxide surface into a low surface-energy, hydrophobic surface, these films appear to avoid attractive meniscus forces. Not only does this alleviate stiction during the release process, it also eliminates capillary condensation, a dominant cause of in-use stiction. Both release-related and in-use stiction are reduced by almost four orders of magnitude in microstructures incorporating the SAM deposition and release process presented here. SAM films thus represent an anti-stiction surface passivation capable of surviving the packaging temperature cycle. These results, coupled with the well-documented long-term stability of SAM films in various ambients, suggest great promise for use of these films as anti-stiction coatings in micromechanical devices.

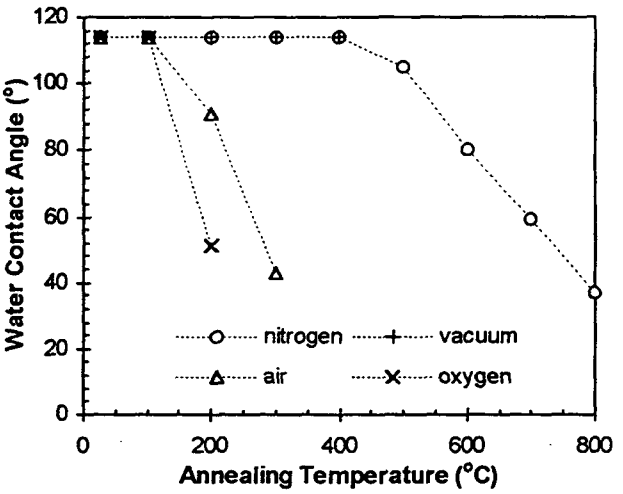


Figure 6. Thermal stability of SAM films on polysilicon in typical packaging environments. Samples were heated to the specified annealing temperature for 5 minutes, then allowed to cool to below 100 °C before removal and measurement.

Self-assembled monolayers also have the advantageous characteristic that the tail group of the precursor molecules can be varied a great deal. For instance, tail lengths for most types of films can be changed to produce thicker or thinner films, and SAM layers might be formed directly on the silicon surface,

eliminating the need for an underlying oxide layer [28]. Other possibilities include changing the deposition method used to form SAMs, possibly by employing a vapor-phase deposition process, which has been demonstrated for some types of films [29]. Finally, recent work involving fluorinated trichlorosilane (FTS) molecules has produced films which exhibit surface energies three times lower than those obtained with the best OTS films [22]. The wide range of options available in investigating and choosing a suitable SAM film makes this a potentially versatile anti-stiction technology.

Acknowledgments

This work has been supported by Sandia National Laboratories, DuPont Central Research and Development, and the Berkeley Sensor & Actuator Center (BSAC). Additional support from the National Science Foundation in the form of a graduate fellowship (MRH) and a young investigator award (RM) is also gratefully acknowledged. The authors would like to thank Norman Tien for his work in designing the test structures, as well as the entire Berkeley Microfabrication Laboratory staff for their assistance.

References

1. H. Guckel, J.J. Sniegowski, T.R. Christenson, and F. Raissi, *Sensors and Actuators*, **A21**, 346 (1990).

2. R.L. Alley, R.T. Howe, and K. Komvopoulos, *Proc. IEEE Solid-State Sensor and Actuator Workshop*, Hilton Head, SC, USA, June 21-24, 1992, p. 202.

3. C.H. Mastrangelo and C.H. Hsu, *Proc. IEEE Solid-State Sensor and Actuator Workshop*, Hilton Head, SC, USA, June 21-24, 1992, p. 208.

4. N. Takeshima, K.J. Gabriel, M. Ozaki, J. Takahashi, H. Horiguchi, and H. Fujita, *Transducers '91*, San Francisco, CA, USA, June 24-27, 1991, p. 63.

5. G.T. Mulhern, D.S. Soane, and R.T. Howe, *Transducers '93*, Yokohama, Japan, June 7-10, 1993, p. 296.

6. T.A. Core and R.T. Howe, "Method for fabricating micro-structures," US Patent # 5,314,572, May 24, 1994.

7. C.H. Mastrangelo and G.S. Saloka, *Proc. IEEE MEMS Workshop*, Ft. Lauderdale, FL, USA, Feb. 7-10, 1993, p. 77.

8. T. Abe, W.C. Messner, and M.L. Reed, *Proc. IEEE MEMS Workshop*, Amsterdam, the Netherlands, Jan. 29 - Feb. 2, 1995, p. 94

9. W.A. Clark, R.T. Howe, and R. Horowitz, this conference (Hilton Head '96).

10. T. Juneau and A.P. Pisano, this conference (Hilton Head '96).

11. L.-S. Fan, Y.-C. Tai, and R.S. Muller, *IEEE Int. Electr. Dev. Meet.*, San Francisco, CA, USA, Dec. 11-14, 1988, p. 666.

12. W.C. Tang, T.-C. H. Nguyen, and R.T. Howe, *Proc. IEEE MEMS Workshop*, Salt Lake City, UT, USA, Feb. 20-22, 1989, p. 53.

13. R.L. Alley, P. Mai, K. Komvopoulos, and R.T. Howe, *Transducers '93*, Yokohama, Japan, June 7-10, 1993, p. 288.

14. Y. Yee, K. Chun, and J.D. Lee, *Transducers '95*, Stockholm, Sweden, June 25-29, 1995, p. 206.

15. K. Deng, R.J. Collins, M. Mehregany, and C.N. Sukkenik, *Proc. IEEE MEMS Workshop*, Amsterdam, the Netherlands, Jan. 29 - Feb. 2, 1995, p. 368.

16. M.R. Houston, R. Maboudian, and R.T. Howe, *Transducers '95*, Stockholm, Sweden, June 25-29, 1995, p. 210.

17. M.R. Houston and R. Maboudian, *J. Appl. Phys.*, **78**, 3801 (1995).

18. P.F. Man, B.P. Gogoi, and C.H. Mastrangelo, *Proc. IEEE MEMS Workshop*, San Diego, CA, USA, Feb. 11-15, 1996, p. 55.

19. S.R. Wasserman, Y.-T. Tao, and G. Whitesides, *Langmuir*, **5**, 1074 (1989).

20. A. Ulman, *Introduction to Ultra-thin Organic Films: From Langmuir Blodgett to Self-Assembly*, Academic Press, San Diego, CA, 1991.

21. A.N. Parikh, D.L. Allara, I.B. Azouz, and F. Rondelez, *J. Phys. Chem.*, **98**, 7577 (1994).

22. J.B. Brzoska, I.B. Azouz, and F. Rondelez, *Langmuir*, **10**, 4367 (1994).

23. M. Biebl, G.T. Mulhern, and R.T. Howe, *Transducers '95*, Stockholm, Sweden, June 25-29, 1995, p. 198.

24. M.R. Houston, R. Maboudian, and R.T. Howe, manuscript in preparation.

25. E.J. Garcia and J.J. Sniegowski, *Sensors and Actuators*, **A48**, 203 (1995).

26. P. Fountaine, D. Goguenheim, D. Deresmes, D. Vuillaume, M. Garet, and F. Rondelez, *Appl. Phys. Lett.*, **62**, 2256 (1993).

27. S.R. Cohen, R. Naaman, and J. Sagiv, *J. Phys. Chem.*, **90**, 3054 (1986).

28. M.R. Linford, P. Fenter, P.M. Eisenberger, and C.E.D. Chidsey, *J. Am. Chem. Soc.*, **117**, 3145 (1995).

29. R.M. Wallace, P.J. Chen, S.A. Henck, and D.A. Webb, *J. Vac. Sci. Technol. A*, **13**, 1345 (1995).

APPLICATION OF CHEMICAL-MECHANICAL POLISHING TO PLANARIZATION OF SURFACE-MICROMACHINED DEVICES

R.D. Nasby, J.J. Sniegowski, J.H. Smith, S. Montague, C.C. Barron, W.P. Eaton, and P.J. McWhorter
Integrated Micromechanics, Microsensors, and CMOS Technology Department
Sandia National Laboratories
Albuquerque, NM 87185-1080

D.L. Hetherington
Microelectronics Patterning and CMP Department
Sandia National Laboratories
Albuquerque, NM 87185-1084

C.A. Apblett and J. G. Fleming
Microelectronics Thin Films Department
Sandia National Laboratories
Albuquerque, NM 87185-1084

ABSTRACT

Chemical-Mechanical Polishing (CMP) has emerged as an enabling technology for the manufacturing of multi-level metal interconnects used in high-density Integrated Circuits (IC). In this work we present the extension of CMP from sub-micron IC manufacturing to the fabrication of complex surface-micromachined Micro-ElectroMechanical Systems (MEMS). This planarization technique alleviates processing problems associated with fabrication of multi-level polysilicon structures, eliminates design constraints linked with non-planar topography, and provides an avenue for integrating different process technologies. We discuss the CMP process and present examples of the use of CMP in fabricating MEMS devices such as microengines, pressure sensors, and proof masses for accelerometers along with its use for monolithically integrating MEMS devices with microelectronics.

INTRODUCTION

Local and global planarization techniques have become key manufacturing technologies for the fabrication of high-density Integrated Circuits (IC). Several methods such as BPSG reflow [1], spin-on-glass [2], and plasma etch [3] have been historically used to accomplish planarization in the IC industry. As increased metal interconnect levels are used and as photolithography requirements are tightened for sub-micron IC fabrication, Chemical-Mechanical Polishing (CMP) has emerged as an enabling technology for fabrication of these structures [4]. CMP processes produce both global and local planarization through relatively simple and quick processing.

In micromachining, the use of polysilicon and silicon dioxide as a structural and sacrificial material, respectively, has been widely employed since its introduction [5]. As these relatively thick (~2 μm) layers of polysilicon and oxide are deposited and etched, considerable surface topography arises which imposes limitations in deposition, patterning, and etching of subsequent layers. It is desirable to planarize specific layers in order to eliminate processing difficulties associated with photoresist step coverage, depth-of-focus of photolithography equipment, and stringer generation during dry etch. Presently, these problems are addressed through careful design of structures, special photoresist processes, and the use of extra mask levels.

In the area of Micro-ElectroMechanical Systems (MEMS), fabrication issues associated with non-planar devices and techniques for overcoming topography problems are just beginning to appear. Researchers at U. Wisconsin have demonstrated locally-planarized surface-micromachined pressure sensors produced in a double LOCOS process [6]; while the use of plasma planarization has recently been demonstrated on MEMS devices by researchers at Delft [7]. Although these planarization techniques have yielded improvement in the manufacturability of MEMS devices through local planarization, CMP provides a higher quality of both local and global planarization in a manufacturing environment. Recently, researchers at Case Western have used CMP to improve the optical quality of mirrored surfaces produced on polysilicon surface-micromachined devices [8] by reducing the roughness of the polysilicon.

Additionally, present non-planar multi-level, surface-micromachining technologies place constraints on designs for upper polysilicon levels to prevent their interference with previously deposited layers. The planarization of sacrificial oxide layers prior to deposition of subsequent, conformal polysilicon layers removes these design constraints and enables the fabrication of new classes of devices. Examples of these new devices are flow channels that contain other micromachined structures within the flow channel and planar covers with applications in packaging of devices. CMP planarization provides an avenue to integrate separate process technologies such as microelectronics and micromechanics or surface micromachining and high-aspect-ratio micromachining. The planarization offered by CMP will also enable the fabrication of devices with four or more structural levels of polysilicon.

CHEMICAL-MECHANICAL POLISHING

Figure 1 illustrates the CMP process in which an oxide surface is planarized by rotating a wafer under pressure against a polishing pad in the presence of a silica-based alkaline slurry. The theory of oxide polishing is not well understood; however, it is generally accepted that the alkaline chemistry hydrolyzes the oxide surface and sub-surface thus weakening the SiO_2 bond structure [9]. The mechanical energy imparted to the abrasive slurry particle through pressure and rotation causes high features to erode at a faster rate than low features, thereby planarizing the surface over time. For the particular results shown here, a colloidal-fumed silica slurry (Cabot SS-12) and a polyurethane pad (Rodel IC1000/Suba IV) are used.

One of the main differences between planarizing inter-level dielectric materials used in ULSI interconnect technology and sacrificial oxide material used in MEMS technology is the large step heights (2-4 μm vs. 0.8 μm) associated with MEMS. These large step heights present a challenge to any planarization strategy. Figures 2 and 3 show the surface topography of single level polysilicon structure prior to CMP (2) and after the CMP process (3). As can be seen from the figure, CMP does an excellent job of removing the ~ 2 μm step height and planarizing the surface.

0-9640024-1-8 $3.00 © 1996 TRF

MEMS APPLICATIONS OF CMP

Improved Manufacturability of MEMS with Large Topography

The deposition and subsequent patterning of thin films conformally deposited over large surface topographies poses significant manufacturing challenges. These challenges include problems with photoresist coverage near large step heights, stringer generation at these steps during dry etch, and photoresist removal. An example of a stringer after release from the edge of a surface-micromachined gear is shown in Figure 4. Planar structures inherently alleviate these problems caused by step heights without the need for additional processing or careful design.

Multi-Level Structures

Multiple levels of structural polysilicon enable fabrication of increasingly complex MEMS devices such as microengines [10] capable of driving sets of gears. Simple flexure structures such as comb drives [11] can be built in a single level of structural polysilicon while a second level enables the production of gears with pin joints. A third level enables linkages between spinning gears, but introduces severe topography problems that must be solved with a combination of careful process and device design. Figure 5 shows the hub and linkage arm of an electrostatic microengine fabricated in a non-planar technology. Intrusion of upper polysilicon levels into gaps left between lower polysilicon levels complicates design and limits layout flexibility in addition to causing undesirable areas of stress concentration. These intrusions are clearly seen in Figure 5, but are completely eliminated in the structure fabricated with CMP shown in Figure 6. Additional levels of polysilicon promise even more complex structures, but the extreme topography of these multi-level structures (four or more active layers) has prohibited their fabrication. A micromachining technology that includes planarization steps such as CMP will enable device fabrication with these additional levels.

CMP also enables the fabrication of structures with flat covers fabricated in upper polysilicon levels that overlap structures fabricated in lower polysilicon levels. Figure 7 shows a fluid pump with a constant cross-section flow channel containing a pumping shaft. Close vertical dimensions between the polysilicon pumping shaft and its cover must be maintained without intrusion of the cover into the pump mechanism. The fabrication of this structure without planarization would have been nearly impossible. These types of covers made from polysilicon or silicon nitride deposited over refilled and planarized regions may also be used for packaging of MEMS.

Pressure Sensors

A planar pressure sensor technology [12] has been developed based upon a silicon nitride layer as the diaphragm material. A trench is etched ~2 μm deep in the surface of a silicon wafer. This trench is refilled with a sacrificial oxide and planarized with CMP. A silicon nitride diaphragm layer is then deposited. The sacrificial oxide underneath this diaphragm layer is removed using HF leaving a cavity beneath the diaphragm. An additional silicon nitride layer is used to seal the cavity in near-vacuum conditions (approx. 200 mTorr). Polysilicon piezoresistors are deposited on the diaphragm to sense the diaphragm strain that results from changes in ambient pressure. A completed, 100-μm-diameter planar pressure sensor is shown in Figure 8. This sensor shows marked improvement in planarity when contrasted against a similar, non-planar sensor shown in Figure 9.

In the non-planar sensor manufactured without CMP the reference pressure cavity is formed above the silicon surface producing large step heights. These steps cause processing difficulties with photoresist coverage and removal, dry etch selectivity to photoresist where the photoresist thins over steps, and metal step coverage. These challenges to manufacturing were eliminated in the CMP-planarized sensor. In the future, the use of an embedded cavity will enable the monolithic integration of this sensor technology with CMOS in a modified version of the integration process described later in this paper and will also enable the use of deeper vacuum cavities to extend the sensing range of the devices.

High-Aspect Ratio Micromachining

In addition to its uses for planarization of surface-micromachined devices, CMP plays a key role in the fabrication of high-aspect ratio micromolded polysilicon and tungsten devices. In the technology presented here[13], deep, narrow trenches lined with oxide are filled with thin films (~2-5 μm) of either polysilicon as shown in Figure 10 or tungsten as shown in Figure 11 and planarized with CMP. These structures can then be integrated with surface micromachined polysilicon structures to form large high-aspect-ratio proof masses with compliant surface-micromachined springs as shown in Figure 12.

Integration of MEMS with Microelectronics

Recently, a great deal of interest has developed in manufacturing processes that allow the monolithic integration of MEMS with driving, controlling, and signal processing electronics [14]. The monolithic integration of micromachines and microelectronics enables the development of wide new classes of small, smart, products with maximum levels of system performance. A new integration scheme [15] overcomes the limitations of traditional integration schemes and enables, for the first time, the integration of micromachines of arbitrary complexity with high performance, state-of-the-art CMOS. Functionality yields in excess of 70% have been achieved with this technology. Examples of devices built in this technology are shown in Figures 13 and 14.

We have developed this embedded MEMS approach to enhance the manufacturability, design flexibility, and performance of microelectronic/micromechanical devices. This process places the micromechanical devices in a shallow (~6-12 μm)trench, planarizes the wafer, and seals the micromechanical devices in the trench. These wafers with the completed, planarized micromechanical devices are then used as starting material for a conventional CMOS process. This technique is equally applicable to other microelectronic device technologies such as bipolar or BiCMOS. Since this integration approach does not modify the CMOS processing flow, the wafers with the subsurface micromechanical devices can also be sent to a foundry for microelectronic processing. Furthermore, the topography of multiple polysilicon layers does not complicate subsequent photolithography. A high-temperature anneal is performed after the devices are embedded in the trench prior to microelectronics processing. This anneal stress-relieves the micromechanical polysilicon and ensures that the subsequent thermal budget of the microelectronic processing does not affect the mechanical properties of the polysilicon structures.

Figure 15 is a schematic cross-section of the integrated technology. Alignment marks are etched onto the surface of wafer in order to provide reference locations for subsequent processing. A shallow trench is etched in (100) silicon wafers using an anisotropic etchant that preferentially etches the (100) crystal plane

and produces a trench with sidewalls having a slope of 54.7° relative to the surface. This slope aids in the subsequent photo patterning within the wells.

The alignment marks from the top surface of the wafer are used as references to generate another set of alignment marks on the bottom surface of the trench. This approach is used to optimize level-to-level registration and resolution of features within the trench. Feature sizes with critical dimensions as small as 0.8 μm were successfully defined within the trench.

A silicon nitride film is deposited to form a dielectric layer on the bottom of the trench. Sacrificial oxide and multiple layers of polysilicon are then deposited and patterned in a standard surface micromachining process. Polysilicon studs provide contact between the micromechanical devices and the CMOS; the depth of the trench is sized so that the top of the polysilicon stud lies just below the top of the planarized trench. The shallow trenches are then filled with a series of oxide depositions optimized to eliminate void formation in high-aspect-ratio structures. The wafer is subsequently planarized with chemical-mechanical polishing (CMP). The entire structure is annealed to relieve stress in the structural polysilicon and sealed with a silicon nitride cap. At this point, conventional CMOS processing is performed. The backend of the process requires an additional step to open the nitride cap over the micromechanical layer prior to release of the micromechanical structures. Photoresist is used as a protection layer over the exposed bond pads during the release process.

This technology does not impose additional limits on the size, thickness, or number of layers of the micromechanical polysilicon structures. The modularity of the process allows changes to be made to either the micromechanical process or the microelectronic process without affecting the other process. A planarized wafer with the embedded MEMS can serve as starting material for a conventional microelectronics foundry service since the technology does not require significant modifications of standard microelectronic fabrication processes.

CONCLUSION

Just as planarization techniques are now important in multi-level-metal IC processes, planar processing is emerging as a crucial step in fabrication of complex MEMS such as devices with multiple structural levels, enclosures, embedded reference pressure cavities, or integrated microelectronics. In this paper, CMP has been demonstrated as a valuable processing technique to add to the toolbox of today's sophisticated micromachinist.

ACKNOWLEDGMENTS

This work, performed at Sandia National Laboratories, was supported by the U.S. Department of Energy under contract DE-AC04-94AL85000. The authors are indebted to the process development engineers, operators, and technicians of Sandia's Microelectronics Development Laboratory for their contributions to the process development, fabrication, and testing of these devices.

REFERENCES

1. S. Robles, K. Russell, M. Galiano, V. Kithcart, V. Sira, and B. C. Nguyen, "Benefits of Sub-Atmospheric CVD BPSG Over the Conventional Pressure Process", *Proc. VLSI Interconnect Conference*, p. 122, (1995).

2. L. Chen, "Aspect-Ratio-Independent-Coating (ARIC) SOG Process for the Double Metal 16M DRAM Planarization", *Proc. VLSI Interconnect Conference*, p. 274 (1995).

3. A. Nagy, J. Helbert, "Planarized Inorganic Interlevel Dielectric for Multilevel Metallization", *Solid State Technology*, p. 53, Jan. (1991).

4. W. Patrick, W. Guthrie, C. Standley, and P. Schiable, "Application of Chemical Mechanical Polishing to the Fabrication of VLSI Circuit Interconnections," *Journal of the Electrochemical Society, 138*, pp. 1778-1784 (1991).

5. R. Howe and R. Muller, "Polycrystalline Silicon Micromechanical Beams," *Journal of the Electrochemical Society, 103*, pp. 1420-1423 (1983).

6. H. Guckel and D. Burns, "Planar Processed Polysilicon Sealed Cavities for Pressure Transducer Arrays," *IEDM*, pp. 223-225 (1984).

7. Y. Li, P. French, and R. Wolffenbuttel, "Plasma Planarization for Sensor Applications," *IEEE JMEMS, 4*, pp. 119-131, (1995).

8. A. Yasseen, S. Smith, M. Mehregany, F. Merat, "Diffraction Grating Scanners Using Polysilicon Micromotors," *Proc. IEEE MEMS '95*, pp. 175-180 (1995).

9. L. Cook, "Chemical Processes in Glass Polishing," *J. Non-Cryst. Solids, 120*, pp. 152-171, (1990).

10. E. Garcia and J. Sniegowski, "Surface Micromachined Microengine," *Sensors and Actuators A, 48*, pp. 203-214 (1995).

11. W. Tang, T-C. Nguyen, and R. Howe, "Laterally Driven Polysilicon Resonant Microstructures," *Sensors and Actuators, 20*, pp. 25-32 (1989).

12. W. Eaton and J. Smith, "Surface-Micromachined Planar Pressure Sensor with a Sub-Surface, Embedded Reference Pressure Cavity," submitted to SPIE Micromachining and Microfabrication (1996).

13. J. Fleming and C. Barron, "Novel Silicon Fabrication Process for High-Aspect-Ratio Micromachined Parts," *Proc. SPIE Micromachining and Microfabrication '95, 2639*, p. 185 (1995).

14. R. Howe, "Polysilicon Integrated Microsystems: Technologies and Applications," *Proc. Transducers* '95, pp. 43-46 (1995).

15. J. Smith, S. Montague, J. Sniegowski, J. Murray, and P. McWhorter, "Embedded Micromechanical Devices for the Monolithic Integration of MEMS with CMOS", *Proc. IEDM ' 95*, pp. 609-612 (1995).

FIGURES

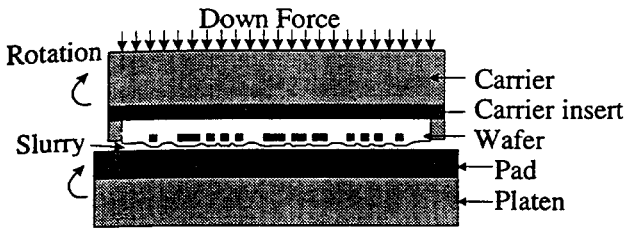


Figure 1. Schematic representation of CMP process in which an oxide surface is planarized by rotating the wafer under pressure against a polishing pad in the presence of a silica-based alkaline slurry.

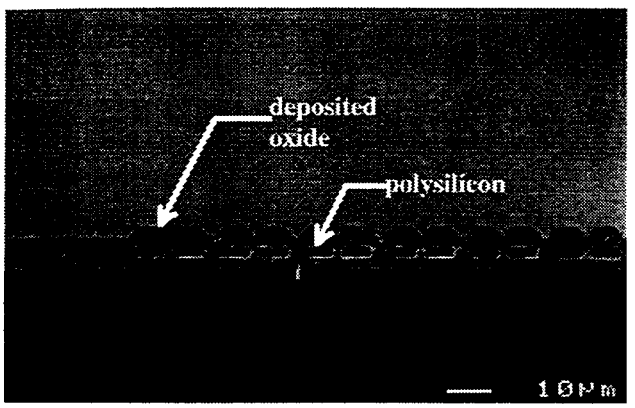


Figure 2. Cross section of a partially-fabricated micromachine showing the uneven topography before CMP planarization.

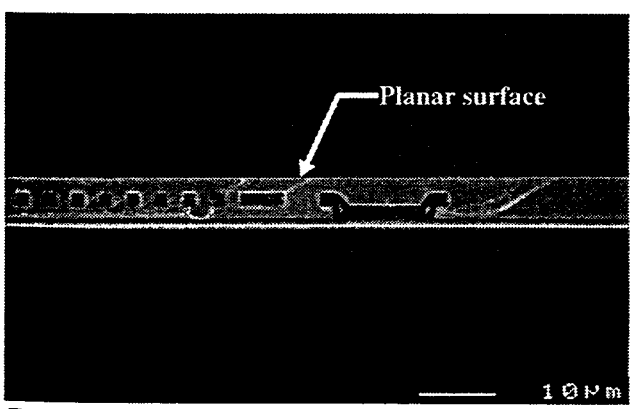


Figure 3. Cross section of a partially-fabricated micromachine after CMP planarization. Note the planar oxide surface.

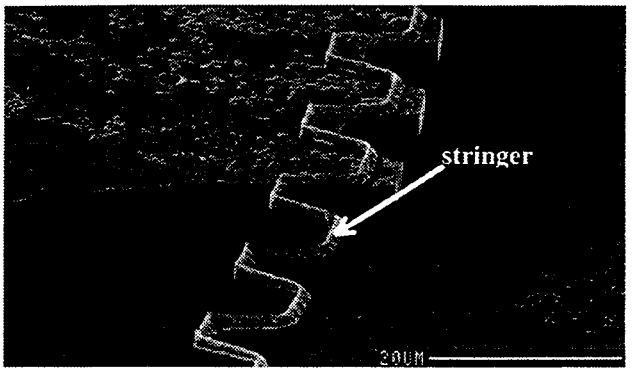


Figure 4. Polysilicon stringer next to a gear edge. The stringer was formed during a later polysilicon deposition and patterning.

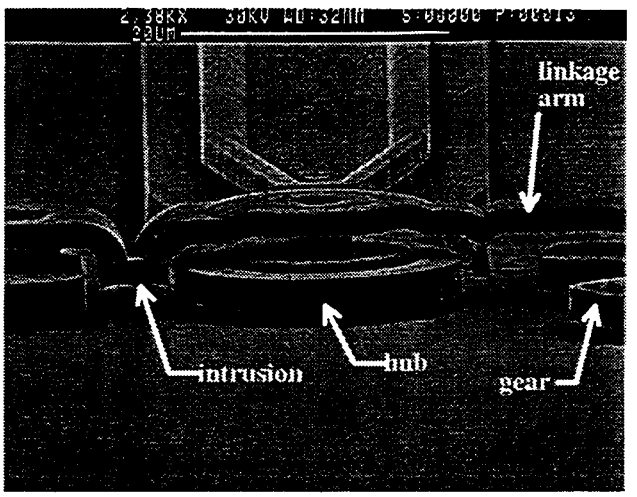


Figure 5. Hub and linkage arm of a microengine produced in a conventional, non-planar micromachining process.

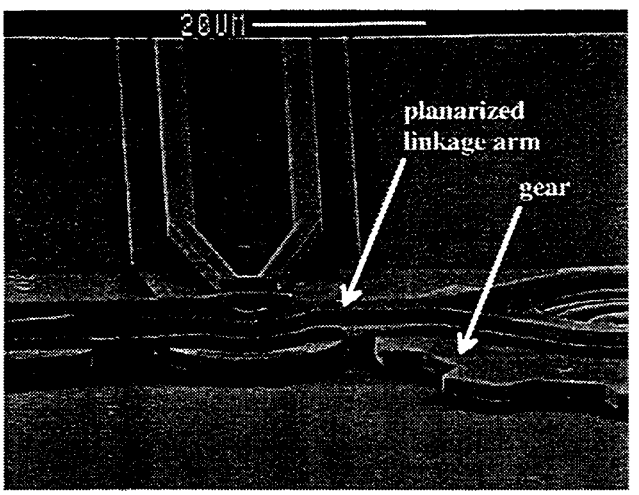


Figure 6. Hub, linkage arm, and gear of a microengine manufactured in a process utilizing CMP planarization.

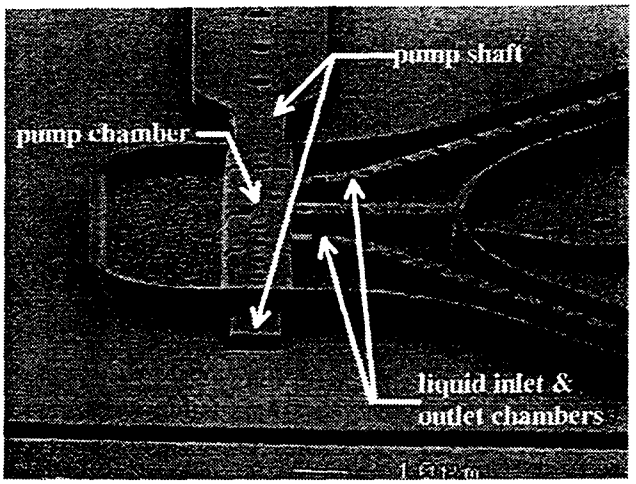


Figure 7. Micromechanical fluid pump fabricated with CMP. The planar top surface of the pump chamber indicates there is no protrusion of the top polysilicon into the enclosed pump mechanism.

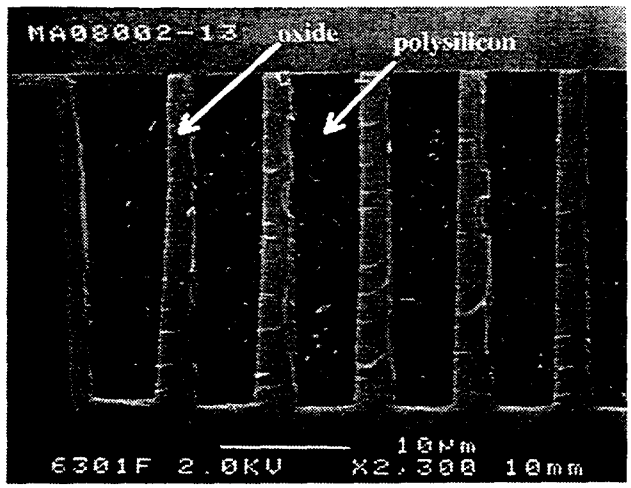


Figure 10. Cross-section of a high-aspect-ratio polysilicon accelerometer proof mass with CMP-planarized top surface.

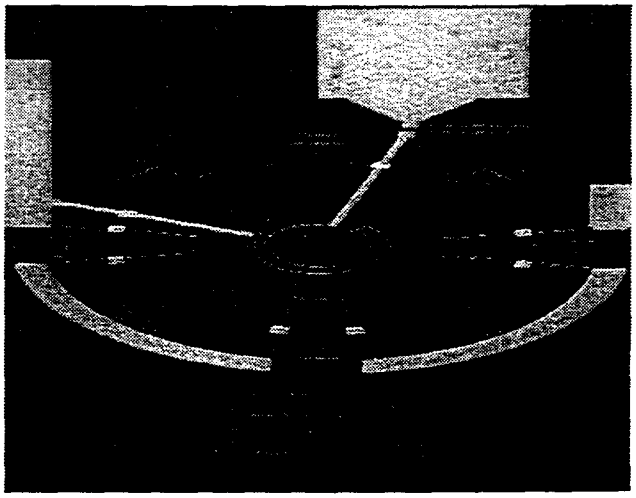


Figure 8. Top view of a planar pressure sensor fabricated in a process that includes CMP. The reference cavity is embedded below the substrate surface, eliminating step coverage problems.

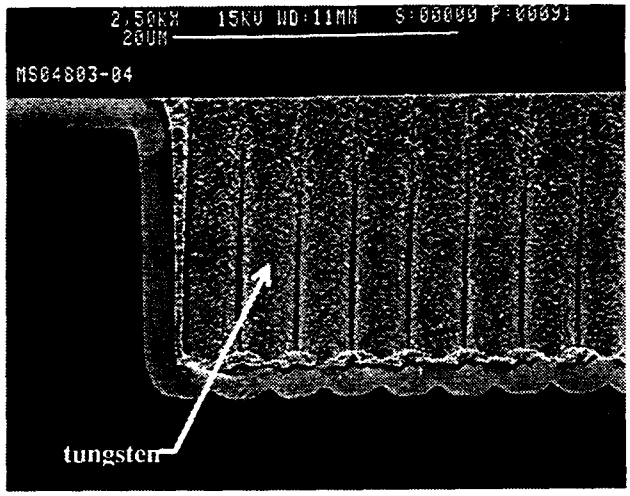


Figure 11. Cross-section of a high-aspect-ratio tungsten accelerometer proof mass with CMP-planarized top surface.

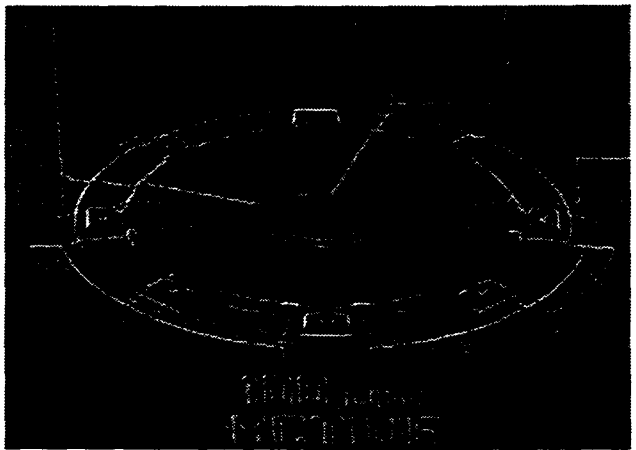


Figure 9. Top view of a non-planar pressure sensor. The circular diaphragm covers a 2μm-high reference vacuum cavity, and is the source of step coverage problems.

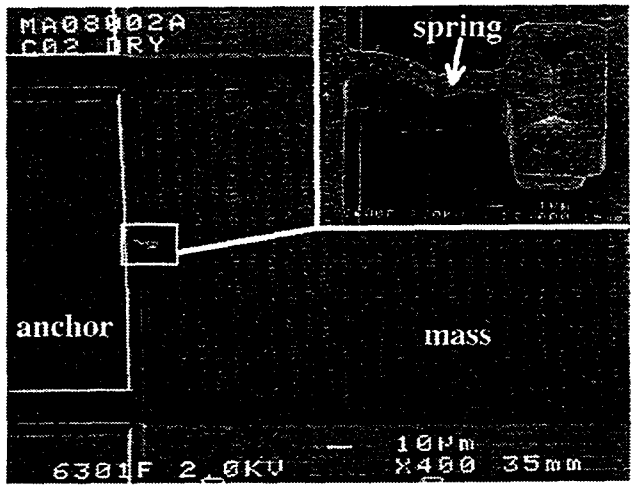


Figure 12. High-aspect-ratio polysilicon accelerometer mass integrated with a surface-micromachined spring

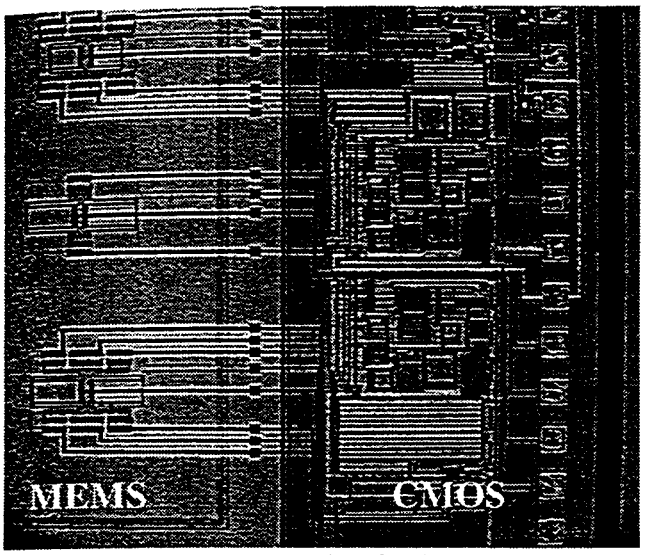


Figure 13. Surface-micromachined polysilicon resonators built in a trench alongside their CMOS sensing electronics.

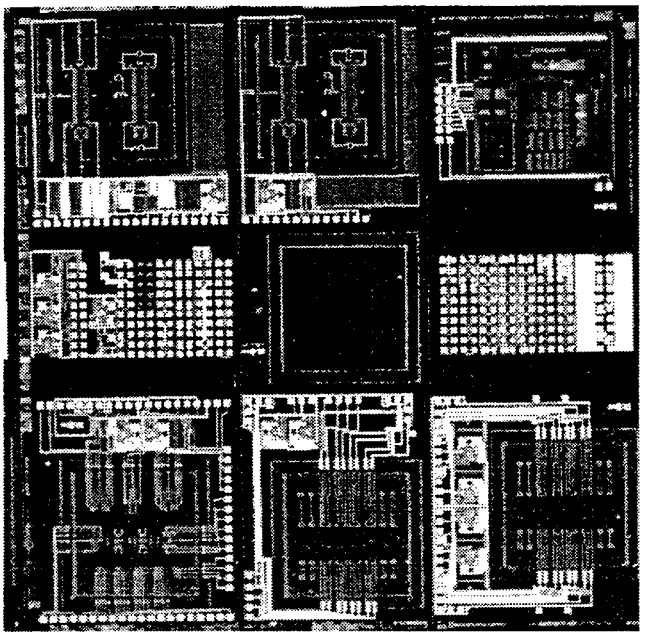


Figure 14. Photograph of the completed integrated MEMS/CMOS die illustrating the layout of the various sensors, resonators, test structures, and electronics.

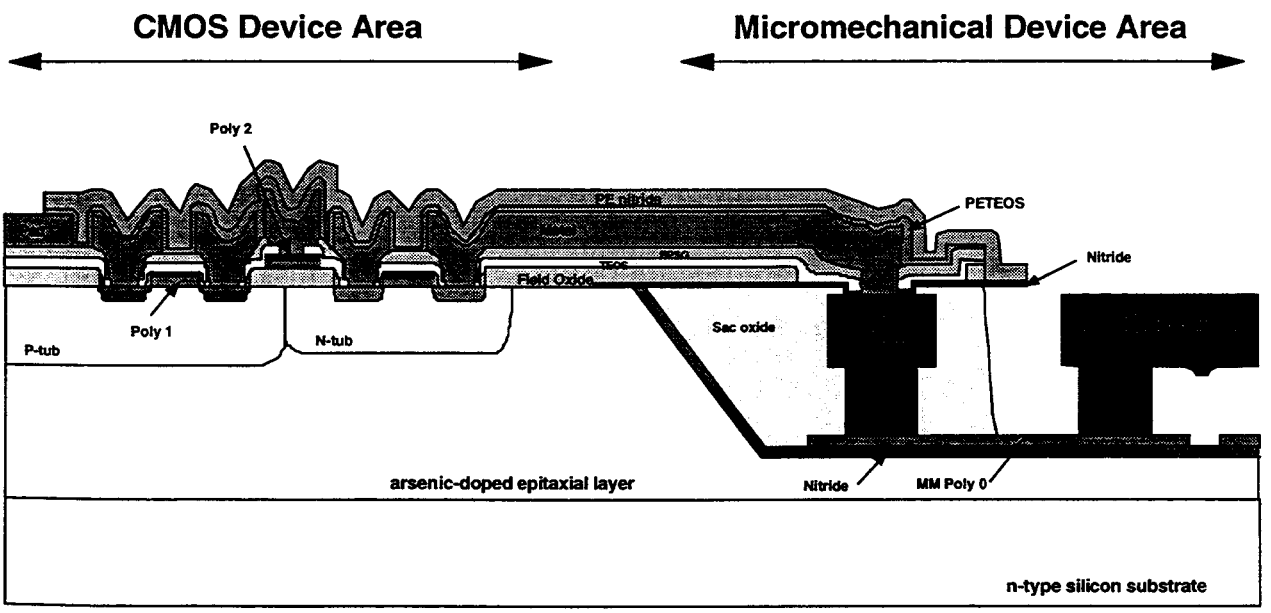


Figure 15. Cross-sectional schematic of the subsurface, embedded MEMS integrated technology.

DIRECT NEWTON FINITE-ELEMENT/BOUNDARY-ELEMENT TECHNIQUE FOR MICRO-ELECTRO-MECHANICAL-ANALYSIS

N. R. Aluru and J. White
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology, Cambridge, MA 02139

ABSTRACT

In this paper we describe a new approach to self-consistent electromechanical simulation. The approach employs a finite-element method for mechanical analysis and a precorrected fast-Fourier transform accelerated boundary-element method for electrostatic analysis. The convergence of the new approach is compared with that of relaxation and surface-Newton methods and is shown to exhibit superior convergence behavior.

I. INTRODUCTION

In order to investigate design alternatives, designers of novel MEMS structures need efficient, robust and easily used computer simulation tools. And since most of the structures of interest are geometrically complicated, electromechanically coupled, and are inherently three-dimensional, Micro-Electro-Mechanical CAD (MEMCAD) tool developers have been focussed on improving the usability, efficiency and robustness of coupled 3-D electromechanical analysis. In particular, finite-element based elastostatic analysis and accelerated boundary-element based electrostatic analysis have been combined using algorithms based on relaxation and a form of surface-Newton method [1], [2], [3]. Although the relaxation algorithm is easy to program, only requiring that data be passed back and forth between "black-box" elastostatic (e.g. ABAQUS [4]) and electrostatic (e.g. FASTCAP [5]) analysis programs, the algorithm diverges if the structure is too flexible or the electric fields are too large. Matrix-free surface-Newton methods preserve the "black-box" nature of the relaxation algorithm and have better convergence properties, but also have perturbation parameters which must be tuned for good performance [3]. In this paper, we present a direct-Newton based approach which is more robust and exhibits superior convergence behavior compared to relaxation and surface-Newton methods.

This paper is organized as follows: In the next section a brief overview is presented on the existing approaches to self-consistent electromechanical simulation. In Section III, the direct-Newton technique is described and numerical results are presented in Section IV. Finally, conclusions are presented in Section V.

II. PREVIOUS APPROACHES

A. Relaxation Technique

Relaxation is the simplest approach to electromechanical simulation as the coupled physical domains are analyzed separately. The surface pressures, which depend on the electrostatic charge, are assumed to be known when computing the structural deformation, and the structural deformations are assumed to be known when computing the electrostatic charge. Figure 1 summarizes the relaxation algorithm for electromechanical analysis. An advantage of the relaxation algorithm is that since the domains are analyzed separately, the numerical schemes can be different. For example, a very efficient relaxation scheme combines Galerkin finite-element elastostatic analysis with multipole/precorrected fast Fourier transform acceleration technique for the electrostatic analysis [1], [2]. The availability of commercial simulators for each physical domain also makes relaxation algorithms appealing as they can be implemented very quickly using a black-box approach. The relaxation algorithm, however, is not robust because it converges very slowly or fails to converge in tightly coupled cases.

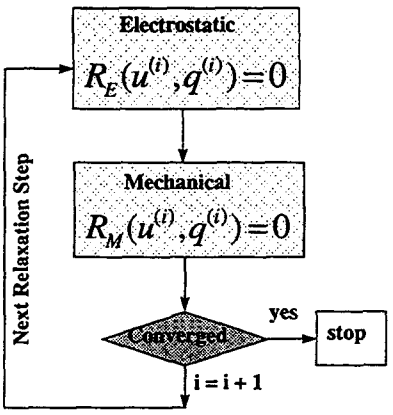


Figure 1. Relaxation algorithm for self-consistent electromechanical analysis. $R_M(u,q)$ and $R_E(u,q)$ denote the elastostatic(mechanical) and electrostatic equations respectively, u is the deformation of the structure, and q is the surface charge that produces the electrostatic force.

B. Surface-Newton Technique

The surface-Newton technique was designed to preserve the black-box nature of the relaxation algorithm but to improve its convergence [3]. The approach presented in [6] is similar to the surface-Newton technique. The key idea in the surface-Newton approach is to reduce the dimensionality of the coupled problem from 3-D to 2-D and to employ a Newton technique to determine the surface node positions. Once the node positions on the structure surface are known, both the surface pressure and the interior node positions can be determined by decoupled electrostatic and elastostatic analysis. Therefore, we can write a surface-Newton iteration equation as

$$R_s({u_s}^k) + \frac{\partial R_s}{\partial u_s}({u_s}^{k+1} - {u_s}^k) = 0 \qquad (1)$$

where u_s is the vector of surface displacements, k is the iteration index, and

$$R_s({u_s}^k) = {u_s}^k - Surf\left[R_M(u_{os}, R_E({u_s}^k + u_{os}, V)\right]. \qquad (2)$$

Here, u_{os} is used to denote the surface of the initial structure and the function $Surf$ extracts u_s from u.

A matrix-free Generalized Conjugate Residual (GCR) iterative method [5] can then be used to solve (2), in which case each GCR

This research was supported by ARPA under ONR contract DABT63-94-C-0053 and FBI contract J-FBI-92-196, by SRC under contract SJ-558, and by grants from IBM and Digital Equipment Corporation.

iteration involves forming a matrix-vector product. The matrix-vector product can be computed using finite-differences i.e.

$$\frac{\partial R}{\partial u} * r \approx \frac{1}{\theta}(R(u + \theta * r) - R(u)) \tag{3}$$

where

$$\theta = sign(u * r) * min(1, \frac{a\|u\|}{\|r\|}, \frac{b\|R(u)\|}{\|r\|})$$
$$a \in (0.01, 0.5) \qquad b \in (0.1, 1)$$

The surface-Newton technique suffers from several drawbacks. First, the algorithm's robustness is very sensitive to the precise choice of θ. If θ is too large, the nonlinearities in R will corrupt the derivative estimate. If θ is too small, the small numerical errors in solving the electrostatic and elastostatic problems will corrupt the derivative estimate. Second, each surface Newton iteration requires a complete solution of the nonlinear elastostatic and electrostatic equations. Thus the approach may not be very efficient. Third, the Jacobian employed in the surface-Newton technique is not exactly identical to the Jacobian of the fully-coupled method presented in the next section. Hence, the convergence of the surface-Newton technique may not be optimal for all cases.

III. New Approach: Direct-Newton Technique

The key idea in the direct-Newton or coupled approach is to account for all the coupling terms between electrical and mechanical systems and to compute the coupling efficiently and exactly. Figure 2 summarizes the coupled algorithm for electromechanical analysis. A coupled approach has been attempted before [7] by employing a finite element method for both the electrostatic and elastostatic equations. This approach may not be computationally very efficient as a finite-element method, in comparison to a multipole-accelerated boundary-element method, would require the construction of an exterior mesh for the electrostatic analysis. A computationally efficient approach of a hybrid finite-element/boundary-element method is employed in the direct-Newton technique.

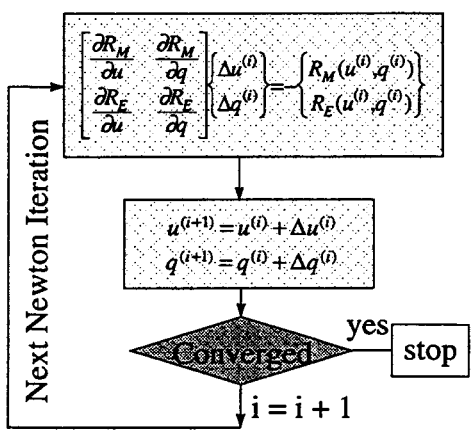


Figure 2. A coupled algorithm for self-consistent electromechanical analysis.

As shown in Figure 2, the outer Newton iteration solves the nonlinear coupled system, and the linear system within each Newton iteration is solved by employing a generalized-minimal-residual (GMRES) algorithm [8]. The coupled system Jacobian can be divided into four parts: the entirely elastostatic part, often referred to as the stiffness or deformation-coefficient matrix, which determines the change in force due to geometric perturbations; the entirely electrostatic part, which determines the change in potential due to perturbations in surface charge; the mechanical to electrical part, which determines the change in force due to perturbation in surface charge; and the electrical to mechanical part, which determines the change in potentials due to geometric perturbations. The deformation-coefficient matrix is computed by employing a Galerkin finite-element formulation. The formulation accounts for both material and geometric nonlinearities and is summarized in Figure 3(a). A sparse storage scheme [9] is employed to store the deformation-coefficient matrix. The mechanical to electrical coupling, $\frac{\partial R_M}{\partial q}$, is computed by calculating the variation of the elastostatic residual equation with respect to the charge and integrating over the boundary elements on the mechanical domain.

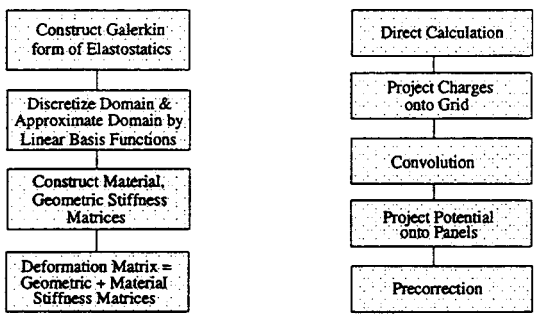


Figure 3. (a) Computation of the deformation coefficient matrix (b) Computation of the matrix-vector product employing precorrected-FFT technique.

The electrostatic or potential coefficient matrix is not needed explicitly as the GMRES algorithm, within each Newton iteration, requires only a matrix-vector product involving the matrix, $\frac{\partial R_E}{\partial q}$, and the charge increment vector, Δq. In earlier approaches to electromechanical simulations [1], [2], [3], a multipole algorithm is employed to efficiently compute the matrix-vector product. In [10], a precorrected Fast Fourier Transform (FFT) algorithm for computing the matrix-vector product was presented, and as compared to multipole algorithm, has been shown to be faster and utilizes less memory. The precorrected-FFT algorithm is employed in our direct-Newton technique and the computation of the matrix-vector product with this approach is summarized in Figure 3(b).

The electrical to mechanical coupling term can be computed by employing a matrix-free approach. The residual equation for the electrostatic system is given as

$$R_E = Pq - \overline{p} = 0 \tag{4}$$

where P is the potential-coefficient matrix, q is the charge vector and $\overline{p}$ is the vector of applied potentials. An approach to compute the electrical to mechanical coupling term is given as

$$\frac{\partial R_E}{\partial u}\Delta u^{(i)} \approx \frac{P(u + \epsilon\Delta u^{(i)})q - P(u)q}{\epsilon} \tag{5}$$

where ϵ is a small parameter and plays a similar role as θ does in the surface-Newton technique. However, we will show in the results section that the choice of ϵ is not critical for the robustness of the method. In order to compute the matrix-vector product, $\frac{\partial R_E}{\partial u}\Delta u^{(i)}$,

accurately a small value of ϵ is desired and is determined through an optimization problem [11]. For well-scaled residuals, an optimal value of ϵ is $O(\epsilon_m^{1/2})$, where ϵ_m is the machine precision.

According to Equation (5), the electrical to mechanical coupling can be computed by performing two matrix-vector products. The first matrix-vector product, $P(u)q$, is straight forward. The second matrix-vector product, $P(u+\epsilon\Delta u^{(i)})q$, can be obtained by perturbing the panel/conductor geometry from u to $u+\epsilon\Delta u^{(i)}$ and applying the precorrected-FFT algorithm.

IV. Results

The direct-Newton technique is compared with the relaxation technique for two examples: a beam over a ground plane structure and two silicon bars with one end of one silicon bar and the other silicon bar held fixed. To make the comparison accurate, a new relaxation code is developed that also employs a precorrected-FFT accelerated boundary-element method for the electrostatic equations, and the finite-element formulation for the elastostatic equations is similar in both the direct-Newton/coupled and relaxation codes.

Figure 4 compares the peak deflection obtained from the relaxation and coupled Newton algorithms for a 500 μm $\times$ 50 μm $\times$ 14.35 μm beam positioned 1 μm above a ground plane. The results are identical verifying the accuracy of the coupled solver. The deflection of the beam for an applied bias of 17.23 V is shown in Figure 5.

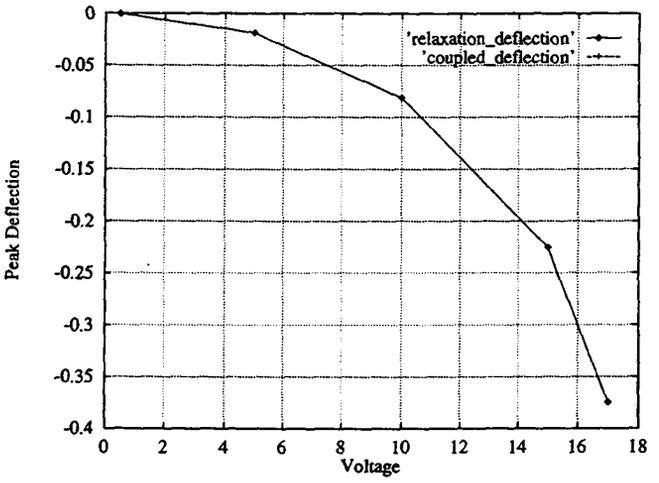


Figure 4. Comparison of peak deflections from relaxation and coupled algorithms for a beam over a ground plane structure. Note that the two curves overlap.

Figures 6 and 7 compare the convergence of the relaxation and coupled algorithms for the beam and ground plane example. The pull-in for the structure is 17.24 V and note that closer to pull-in the relaxation algorithm converges slowly, but the coupled algorithm converges rapidly. The slow convergence of the relaxation algorithm, near pull-in, is due to the increased coupling between elastostatic and electrostatic systems. As the direct-Newton technique accurately accounts for all the coupling it exhibits rapid convergence behavior.

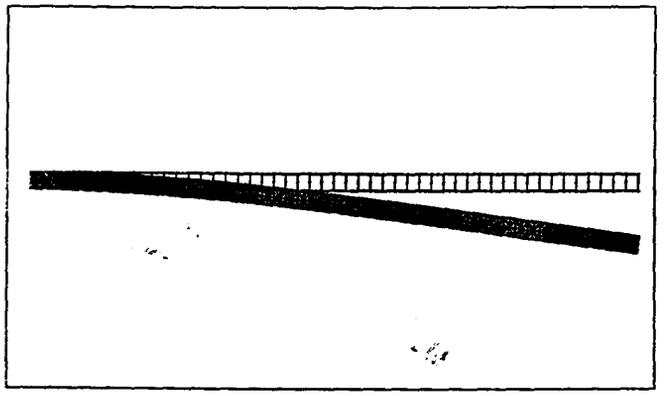


Figure 5. Deflection of the beam for an applied bias of 17.23 V.

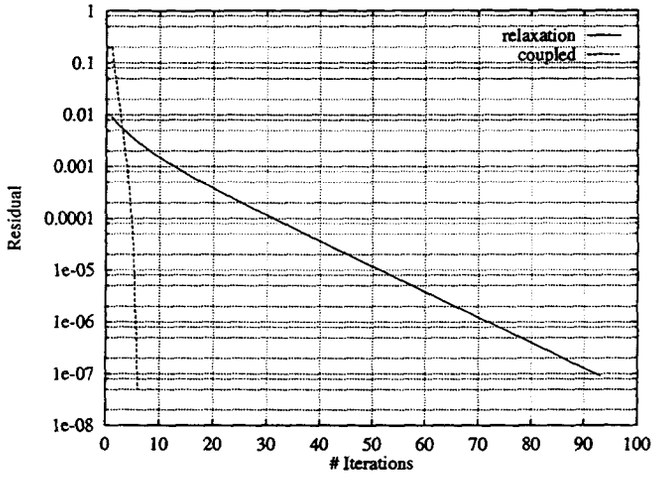


Figure 6. Convergence of relaxation and coupled algorithms for a beam and ground plane structure at 17.20 V.

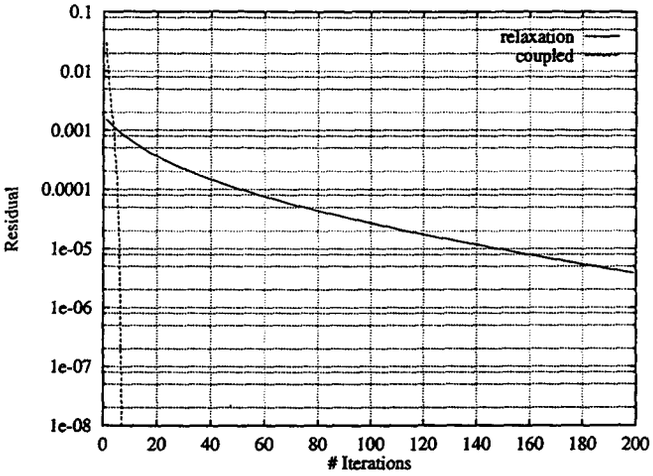


Figure 7. Convergence of relaxation and coupled algorithms for a beam and ground plane structure at 17.23 V.

The direct-Newton technique employs a matrix-free approach to compute the electrical to mechanical computing. The matrix-free parameter, ϵ, is a concern as the method may not work for all values of ϵ. Figure 8 compares the convergence of the coupled algorithm for three different values of ϵ. A value of $\epsilon = 1.0$ requires only one more iteration for convergence as compared to using an optimal ϵ. No noticeable convergence rate differences are observed between employing an optimal ϵ and an ϵ larger than optimal by a factor of 100. Hence, the choice of ϵ is not critical for the robustness of the method, unlike the surface-Newton technique.

Figure 9 compares the convergence of the relaxation and coupled algorithms for two silicon bars with a potential difference of 850 V. The relaxation algorithm fails to converge for this example because of the strong coupling between mechanical and electrical domains at

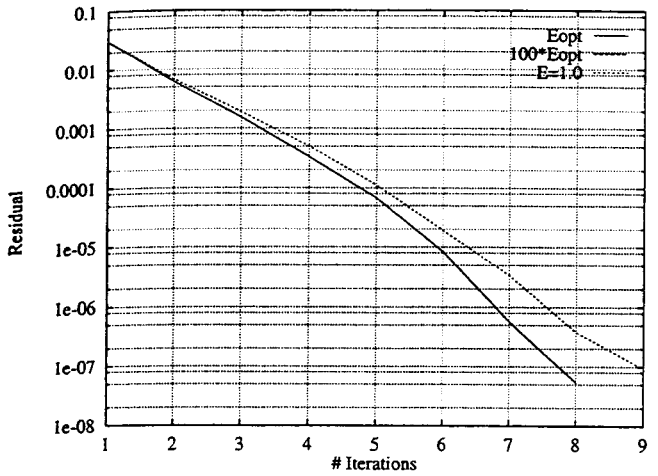


Figure 8. Comparison of convergence for different values of ϵ. Eopt = $sqrt(\epsilon_m)$, 100*Eopt = $100 \cdot sqrt(\epsilon_m)$, and E = 1 corresponds to $\epsilon = 1.0$.

850 V. The coupled algorithm not only converges but converges very rapidly.

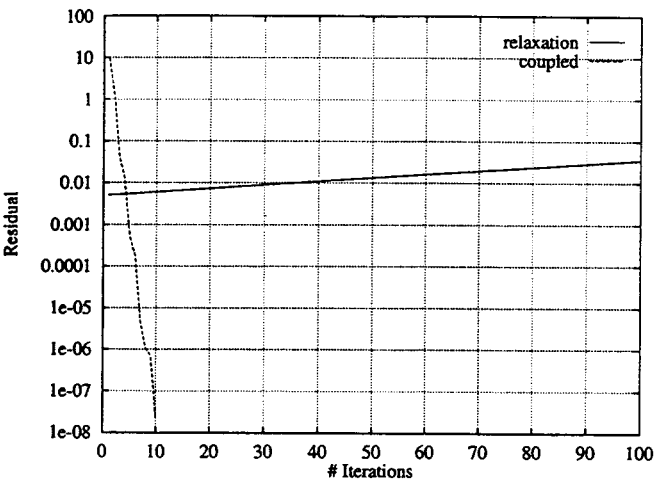


Figure 9. Convergence of relaxation and coupled algorithms for two silicon bars with a potential difference of 850 V.

V. Conclusion and Acknowledgements

The direct-Newton approach is shown to exhibit superior convergence behavior as compared to relaxation method and does not have the parameter sensitivity of the surface-Newton method. In addition, unlike the surface-Newton approach, the direct-Newton method does not require solving nonlinear elastostatic and a linear electrostatic problem in each Newton iteration. The approach presented in this paper to compute the electrical to mechanical coupling term may not be very efficient. We are presently exploring other approaches and they will be reported in a future publication. The efficiency of the direct-Newton technique can be further improved by employing acceleration techniques, such as preconditioners, and this is also being studied.

The authors would like to thank Dr. John R. Gilbert, Dr. Peter M. Osterberg, Joel R. Phillips and Professor Steven D. Senturia for many valuable discussions.

References

[1] X. Cai, H. Yie, P. Osterberg, J. Gilbert, S. Senturia and J. White, "A relaxation/multipole-accelerated scheme for self-consistent electromechanical analysis of complex 3-D microelectromechanical structures," *Proc. ICCAD 1993.*

[2] J. R. Gilbert, R. Legtenberg and S. D. Senturia, "3D coupled electromechanics for MEMS: Applications of CoSolve-EM," *Proc. MEMS 1995.*

[3] H. Yie, X. Cai and J. White, "Convergence properties of relaxation versus the surface-Newton generalized-conjugate residual algorithm for self-consistent electromechanical analysis of 3-D micro-electro-mechanical structures" *Proc. NUPAD V 1994.*

[4] Hibbit, Karlsson and Sorenson, Inc., Providence, R.I.

[5] K. Nabors and J. White, "FastCap: A multipole-accelerated 3-D capacitance extraction program," *IEEE Trans. CAD*, vol. 10, pp. 1447-1459, 1991.

[6] M. Bachtold, J. G. Korvink, J. Funk and H. Baltes, "New convergence scheme for self-consistent electromechanical analysis of iMEMS," *Proc. IEDM 1995.*

[7] H. U. Schwarzenbach, J. G. Korvink, M. Roos, G. Sartoris, and E. Anderheggen, "A microelectromechanical CAD extension for SESES," *JMEMS*, vol. 3, pp. 162-171, 1994.

[8] Y. Saad and M. H. Schultz, "GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems," *SIAM J. Sci. Stat. Comput.*, vol. 7, pp. 856-869, 1986.

[9] K. S. Kundert and A. Vincentelli, "Sparse User's Guide: A sparse linear equation solver," *Univ. of California, Berkeley*, 1988.

[10] J. R. Phillips and J. White, "A Precorrected-FFT method for capacitance extraction of complicated 3-D structures," *Proc. ICCAD 1994.*

[11] P. E. Gill, W. Murray and M. H. Wright, *Practical Optimization*, Academic Press, London, 1981.

THE NANOSCILLOSCOPE: COMBINED TOPOGRAPHY AND AC FIELD PROBING WITH A MICROMACHINED TIP

D.W. van der Weide
Dept. of Electrical Engineering
University of Delaware
140 Evans Hall
Newark, DE 19716-3130
dan@eecis.udel.edu

P. Neuzil
Microfabrication Applications Lab
University of Illinois at Chicago
1120 SEO, 851 S. Morgan St.
Chicago, IL 60607-7053
pavel@casimir.eecs.uic.edu

ABSTRACT

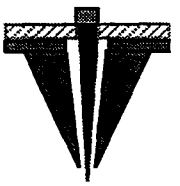
We introduce a new concept for combined probing of topography and local AC fields based on a micromachined silicon coaxial tip, and we report preliminary results of scale-model measurements and tip fabrication.

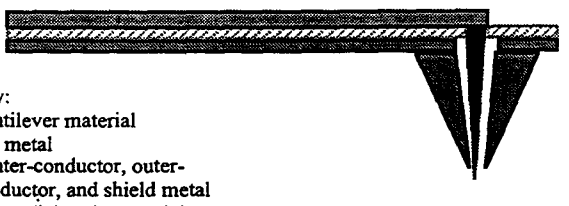
INTRODUCTION

The scanning force microscope (SFM) has been used extensively for measuring topography of samples ranging from soft biological materials to integrated circuits; it has recently been used in measuring AC fields via a frequency-mixing technique that relies on the nonlinear response of the cantilever to voltage differences between the sample and the scanning tip[1, 2]. While this approach is useful for measuring local voltages on integrated circuits, it requires injecting a signal onto the SFM cantilever which is synchronized with the IC signal to be measured, limiting its range of application.

Here, by contrast, we are extending near-field scanning optical microscopy (NSOM) down in frequency from the visible—where cylindrical waveguide, a tapered and metal-coated optical fiber, is used—by employing a *coaxial* micromachined waveguide/tip[3] while maintaining the topographic and tip-sample distance control capabilities of the SFM[4]. Because this instrument can be used to measure both AC fields and topography simultaneously with nanometer-level localization, we call it a NANOSCILLOSCOPE.

Frontal and side-view cross-sections of scanning-force/ evanescent-wave microscope tip and cantilever (from U.S. patent application)





Key:
- Cantilever material
- Tip metal
- Center-conductor, outer-conductor, and shield metal
- Air or dielectric material

Figure 1. Schematic layout of combined SFM/coaxial tip with a coplanar waveguide running along the cantilever.

Coaxial waveguide exhibits no cutoff frequency, unlike the cylindrical tapered waveguide used in NSOM, so if coax is tapered to a tip, a near-field antenna results. Essentially a shielded SFM tip with a planar transmission line integrated along the cantilever, this new structure allows simultaneous measurement of topography and local AC electric fields. With appropriate instrumentation at the root of the cantilever, the frequencies of these fields could be in the several-hundred GHz regime[5, 6].

SCALE-MODEL MEASUREMENTS

We have performed scale-model experiments using available coaxial transmission line tapered to a ~ 20 μm tip radius, and have measured field amplitude and phase contrast in transmission between two tips with wavelength-relative resolution > 10^{-6}, indicating the essentially electrostatic nature of the fields at the tapered tip[3], also shown in the calculation of Fig. 2. These measurements were made over a frequency range of 1–1000 MHz, limited only by the equipment available.

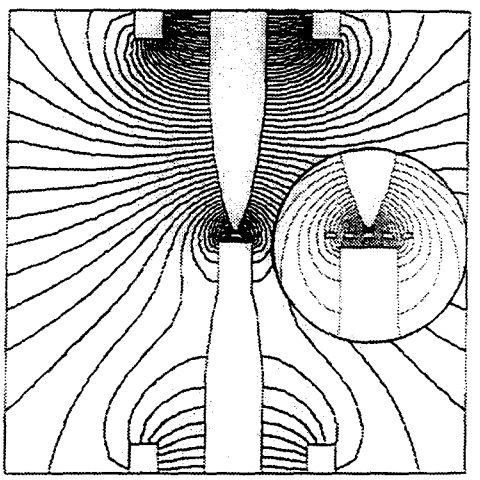


Figure 2. Quasistatic calculation of electric field lines between scale-model coaxial tips in transmission mode. A metallic grid sample is being scanned between the two tips, as shown in the close-up.

Using just a tapered coaxial tip in reflection has already given measurable contrast in the phase of the electric field at microwave frequencies[7], so this mode of the principle is well-established. A final operating mode to be established is that of detection, whereby fields from the sample itself are measured locally. While again a larger-scale example of this principle is regularly employed in many microwave laboratories to measure local electric (magnetic) field strengths by using an open-ended (loop-terminated) coaxial cable, we can also make a rough estimate of the signal voltage expected in this configuration at the scale of the present tip: As depicted in Fig. 3, voltage of 1 V across two conductors spaced by 1 μm results in a ~ 10^4 V/cm electric field. Immersing into this field a coaxial tip with ~ 50 nm tip-shield spacing will result in a ~ 50 mV signal—

0-9640024-1-8 $3.00 © 1996 TRF

fully measurable. This simple argument has neglected losses, but also accounts only for the in-plane components of the field; we expect a contribution from the normal components, as well.

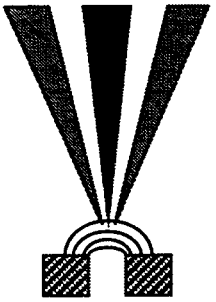


Figure 3. Coaxial tip geometry probing the electric field between two conductors.

MICROMACHINED STRUCTURE

To bring the scale-model experiments into the microscopic regime, we have fabricated a silicon/polysilicon coaxial tip using a method similar to that in [8] . This tip (shown in Figs. 4 and 5), when integrated with a modified coplanar waveguide running along the cantilever, will result in a complete NANOSCILLOSCOPE. While this realization does not use the hybrid metallic/dielectric structure described in Fig. 1, it can achieve a sharper tip radius and can be integrated more readily with existing VLSI processes.

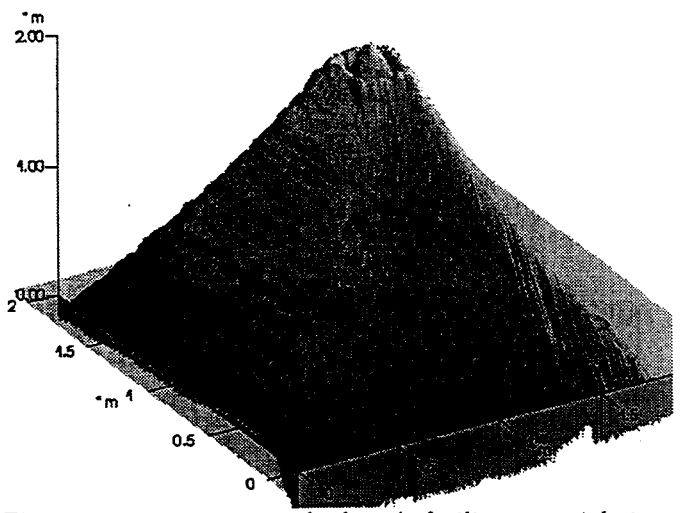


Figure 4. SFM image of silicon/polysilicon coaxial tip; actual tip height is ~ 5 μm.

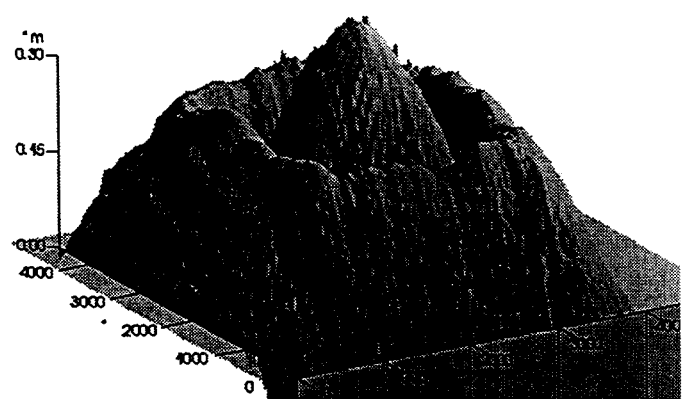


Figure 5. Close-up of coaxial structure; tip radius is ~ 10 nm; horizontal scale is in 1000 Å increments.

CONCLUSIONS

By integrating a micromachined coaxial tip with a planar transmission line/cantilever structure, we achieve an instrument capable of simultaneous localized AC field measurements and sample topography. We expect that the NANOSCILLOSCOPE will be useful in IC failure analysis, local probing of surface chemistry, and electrical characterization of neural function on the nanometer scale.

REFERENCES

[1] A. S. Hou, F. Ho, and D. M. Bloom, "Picosecond electrical sampling using a scanning force microscope," *Electronics Letters*, vol. 28, pp. 2302-2303, 1992.

[2] A. Leyk, C. Böhm, D. W. van der Weide, and E. Kubalek, "104 GHz signals measured by high frequency scanning force microscope test system," *Electronics Letters*, vol. 31, pp. 1046-1047, 1995.

[3] F. Keilmann, D. W. van der Weide, T. Eickelkamp, R. Merz, and D. Stöckle, "Extreme sub-wavelength resolution with a scanning radio-frequency transmission microscope," *To appear in Optics Communications*, 1995.

[4] D. W. van der Weide, "Combined scanning-force/evanescent-mode microscope," United States Patent Pending, 1995.

[5] D. W. van der Weide, "Delta-doped Schottky diode nonlinear transmission lines for 480-fs, 3.5-V transients," *Applied Physics Letters*, vol. 65, pp. 881-883, 1994.

[6] B. Nechay, F. Ho, A. S. Hou, and D. M. Bloom, "Applications of an atomic force microscope voltage probe with ultrafast time resolution," *Journal of Vacuum Science and Technology B*, vol. 13, pp. 1369-1374, 1995.

[7] M. Fee, S. Chu, and T. W. Hänsch, "Scanning electromagnetic transmission line microscope with sub-wavelength resolution," *Optics Communications*, vol. 69, pp. 219-224, 1989.

[8] R. C. Davis, C. C. Williams, and P. Neuzil, "Micromachined submicrometer photodiode for scanning probe microscopy," *Applied Physics Letters*, vol. 66, pp. 2309-2311, 1995.

ADVANCES IN PHOTORESIST BASED PROCESSING TOOLS FOR 3-DIMENSIONAL PRECISION AND MICRO MECHANICS

H. Guckel, P. S. Mangat, H, Emmerich, S. Massoud-Ansari, J. Klein, T. Earles
Department of Electrical and Computer Engineering
University of Wisconsin
Madison, WI 53706

J. D. Zook, T. Ohnstein
Sensor and System Development Center
Honeywell Technology Center
12001 State Highway 55
Plymouth, MN 55441

E. D. Johnson, D. P. Siddons
Brookhaven National Laboratory
National Synchrotron Light Source Department
Upton, NY 11973

T. R. Christenson
Sandia National Laboratory
Albuquerque, NM 87185

ABSTRACT

This report deals with the recent advances aimed at processing tools for microactuators with force outputs which can approach one Newton and precision manufacturing techniques for larger parts via deep x-ray lithography at the University of Wisconsin - Madison. The work is based on the *a priori* decision to use photoresist in a cost effective manner. The cost and performance issues move the tool into the thick photoresist, to 2000 μm, arena in which additive processing such as electroplating is much preferred over subtractive techniques. Application potential for the tool is very much enhanced if multiple sequential masks can be accommodated. This involves re-planarization and mask alignment issues which are unique to thick photoresist tools.

INTRODUCTION

Recent progress towards finding a processing tool for 3-dimensional structures for use in microactuators and other precision engineering applications has concentrated on high aspect ratio fabrication techniques. These techniques increase the 3-dimensionality of structures by allowing them to be very tall but consequently result in prismatic shapes. LIGA and LIGA-like processes are one such group of high aspect ratio processes which allow tall prismatic structures to be fabricated in a variety of materials.[1,2] The basic LIGA process starts with the exposure of a layer of polymethyl methacrylate, PMMA, attached to a conductive substrate. The exposure is performed by a large flux of highly collimated photons (synchrotron radiation) to allow for deep penetration and sidewall run-outs of 0.1 microns over 100 microns of structure height. The exposed PMMA is then developed down to the conductive substrate and the cavities are electroplated. The remaining PMMA may be removed and the metal structures may be used as is or mass produced by injection molding.

Several issues must be addressed to extend this process to microactuators with one Newton output forces and precision machined parts larger than 500 micron in thickness. These issues include: (i) a photoresist technology and an exposure tool which can exceed 500 microns in height, (ii) re-planarization of the upper electroplated surface for both height control and the addition of a second layer, and (iii) a multi-layer process which allows alignment tolerances within the resolution of the process. The basic LIGA process is limited to 500 microns of height due to photoresist strain from volume contraction during polymerization and thermal annealing of the PMMA when constrained to a substrate. Strain can cause photoresist adhesion loss and pattern distortion. Electrodeposition rates are diffusion and field dependent. The photoresist patterns will concentrate the field resulting in non-uniform plating rates. Small structures can lead to electrolyte diffusion limitations. Both of these effects result in the inability to control the upper surface of the electroplated structures making multiple levels without re-planarization impossible. A multiple, thick level process requires alignment methods which can align with mask to substrate gaps in the millimeter range, but hold micron tolerances. These tools extend the basic LIGA process to large precision engineered parts and new classes of microactuators which are detailed in the following sections, respectively.

PROCESS TOOLS

THICK PHOTORESIST

Thick photoresist layers can fail because of built-in strain. The strain typically originates from process induced volume changes and adhesion requirements. This adverse effect can be eliminated by using strain-free photoresist sheets and solvent bonding them to the substrate.[2] A final adjustment of the photoresist height via diamond fly cutting produces excellent results. Using this technique, photoresist thicknesses from 50 microns to several centimeters can be attached to a substrate. Because this process requires no thermal cycles, a variety of substrates can be used including ceramics, metals, plastics, and circuit covered wafers.

Exposure of these thick films to submicron tolerances require a photon source with high collimation, large flux, and deep penetration. Synchrotron radiation proves to meet these requirements. Exposure depth depends on the absorption length at the photon energy used. Figure 1 shows the PMMA absorption length for the appropriately filtered photon

0-9640024-1-8 $3.00 © 1996 TRF

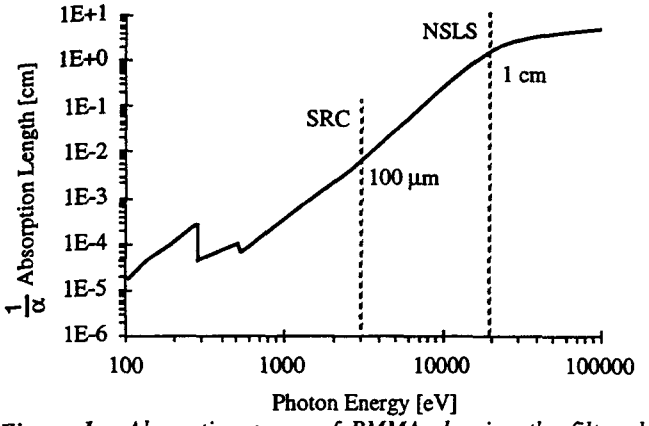


Figure 1: Absorption curve of PMMA showing the filtered spectrum peaks of SRC - Madison and NSLS - Brookhaven.

spectrum of the 1.0 GeV Synchrotron Radiation Center, University of Wisconsin - Madison and the 2.6 GeV National Synchrotron Light Source, Brookhaven National Laboratory.

As a consequenc, two options exist with the use of high energy exposures. Very thick photoresist samples may be exposed[3], or a multiple number of thin photoresist exposures may be performed at the same time through stacking. The achievable throughput with high energy photons is large enough to make direct manufacturing via deep x-ray lithography and electroplating without injection molding cost effective[4].

RE-PLANARIZATION

Recesses in the photoresist can be filled with metals via electroplating. Plating uniformity can be improved by cell design and operating procedures.[5] However, a perfect deposit is impossible for masks with large, field controlled, and small, diffusion controlled, feature sizes. The difficulty can be removed by overplating followed by lapping and polishing.

A lapping and polishing technique has been developed to control the uniformity and final height of the electroplated structures. Lapping involves removal of material in a controlled condition using fine sized diamond slurry, which is embedded into a lapping plate. This is followed by polishing of the lapped structures. Thus, re-planarization via lapping and polishing produces a photoresist height which is identical to the metal height with mirror like surfaces. The height uniformity of these re-planarized parts can be controlled to ±2 µm over an area of 50 mm × 50 mm. The surfaces of the micro

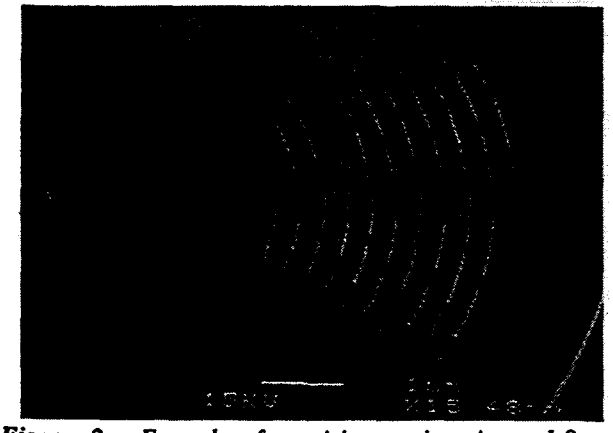


Figure 2: Example of precision engineering: 1.8 cm electrode disk with vertical wall mesh. All tolerances are submicron.

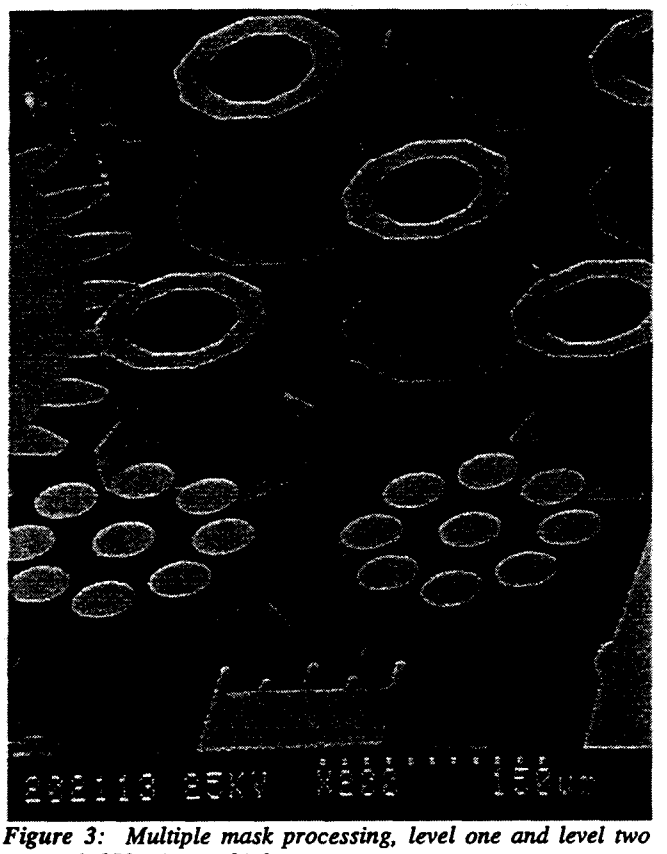


Figure 3: Multiple mask processing, level one and level two are each 150 microns high.

parts have been measured to a flatness of 26 Å rms. Figure 2 shows a large precision machined part which was produced by this procedure.

In addition, this re-planarization technique allows the extension of the basic LIGA process to multiple electroplated layers. Figure 3 shows an example of two layer nickel structures fabricated using this technique. Each layer is 150 microns high.

ALIGNMENT

Multiple mask processes require mask to wafer alignment. For thick photoresists, the alignment proceduremarks have to be viewed through several hundred microns of thick photoresist. "Alignment by assembly" is a concept which was developed in this research. It currently produces alignment errors of 4 microns and is independent of the photoresist thickness and does not require any expensive optical instrumentation (see figure 3).

DEVICES

Thick photoresist in conjuction with new alignment techniques with controlled z-dimensionality, have allowed new classes of actuators to be considered. These actuators take advantage of the increased 3-dimensionality and higher precision of the process tools.

VERTICAL LINEAR ACTUATOR

Figure 4 shows a first generation electrostatic vertical linear actuator with expected force outputs in the milliNewton range and a vertical displacement of hundreds of microns.[6] The actuator consists of three parts. The upper plate is held up by four springs and contains holes which fit around posts in the lower plate, not visible, attached to the substrate. The lower plate has an electrical connection which is shown to extend out

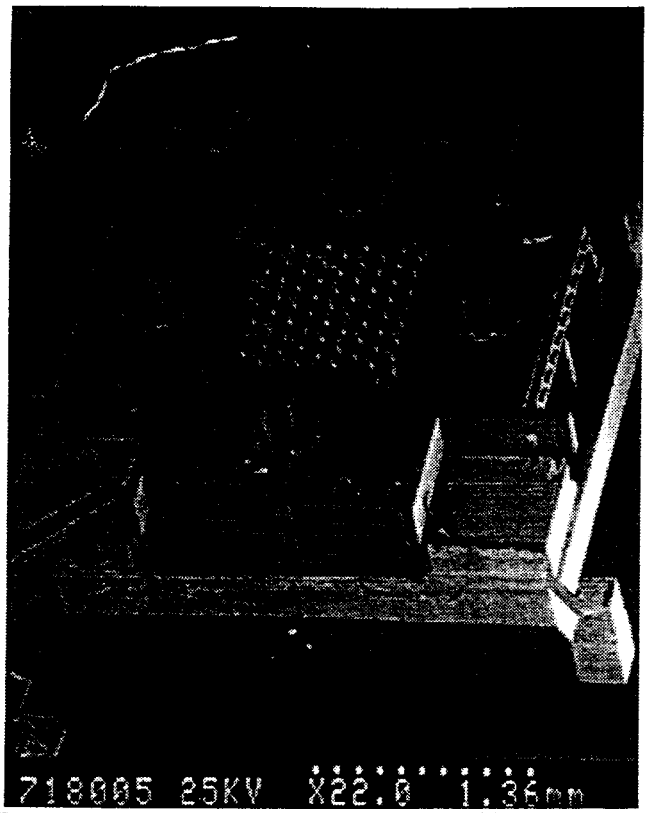


Figure 4: Assembled vertical electrostatic table-top actuator. This device cannot be constructed without polishing.

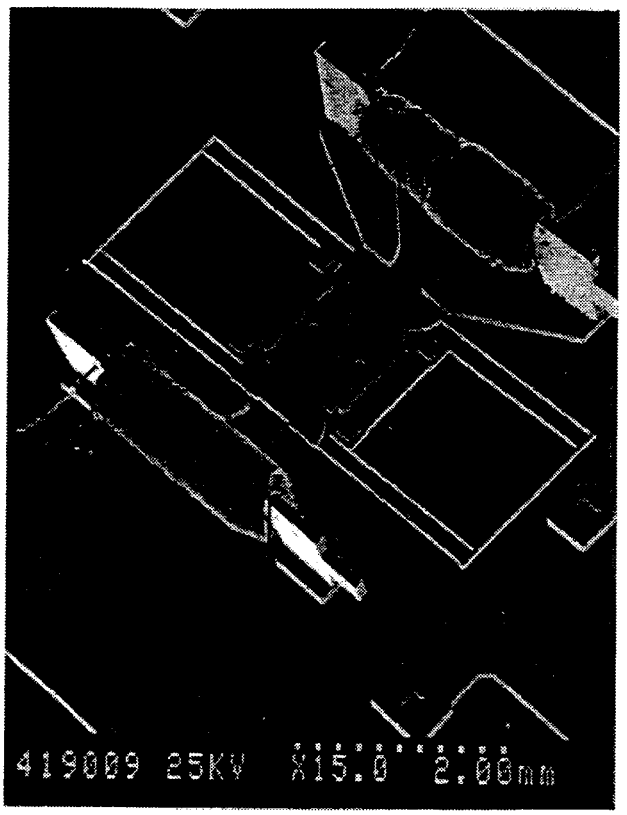


Figure 5: Magnetic linear actuator using assembled micro coils which are inserted into the magnetic path.

of the lower left corner of the device along the substrate. This lower plate is electrically isolated from the upper plate which is attached mechanically and electronically to the four peripheral springs and the lower outer connection ring. A voltage may be applied between the upper and lower plate resulting in electrostatic attraction.

A spring constant of 3.5 Newton/meter was extracted from a force versus vertical displacement measurement. The prototype device shown in figure 4 has a 10 micron post to hole air gap requiring an estimated 100 volts for every 20 microns of vertical displacement. The situation further improves with smaller posts, holes, and air gaps. Fabrication of this device at NSLS can also increase the vertical displacement of the device.

MAGNETIC ACTUATORS

Lapping and polishing has established thickness control of free and fixed metal structures. This has allowed tighter tolerances for assembly. Assembly allows small tolerances through the subtraction of precisely held structures. Assembly also simplifies and increases the 3-dimensionality of the fabrication process. Some examples of recent advances in micro assembly are the assembled micro coil and increasing structural height through stacking.

Figure 5 shows a linear actuator using a single assembled micro coil for each direction of motion.[7] The micro coil has significantly reduced the power required to operate the device, the fabrication time, and actuator size. Powers are in the milliwatt range due to the higher coil packing density. The fabrication time has decreased from several weeks to a few days because an integrated coil technology requires multiple levels and alignment. Since the coil is in the vertical direction, the overall device requires less substrate area.

The linear actuator of figure 5 has a $1/d^2$, non-linear, force due to the plunger's complete closure with the magnetic path. This simple actuator can move 100 microns, and generates forces approaching 35 milliNewton at closure. The actuator consists of a shuttle constrained between two electromagnets. The components are all fabricated out of 78 permalloy and are polished to 100 microns in height. The device uses a 500 turn electromagnet coil that can handle up to 100 mA of current. The 1 T saturation flux density of the electroplated 78 permalloy, reached at about 21 mA, allows the maximum predicted force to approach 35 milliNewton. Coil inductance measured against shuttle position was used to determine the force behavior of the actuator. Larger force are possible at larger structural heights.

Other linear actuator versions with more linear force outputs have resonant travel of up to 300 microns at 400 cps with a quality factor in air between 300 to 400. Operation at resonance with 200 micron peak to peak motion is achieved with a power dissipation of 200×10^{-6} Watts.[7]

Figure 6 shows a microdynamometer using the assembled micro coil.[8] This device is a complete microsystems containing both a drive and sense mechanism. At the top is a three phase variable reluctance micro motor which contains a 1 mm diameter 150 micron thick rotor, which drives into a center idling gear of 250 micron diameter. At the bottom is a magnetic breaking gear which can be used to apply a load via a current in an additional micro coil to measure the torque of the rotor. The motor requires only 1 milliwatt of power to keep the three gears spinning. Measured tangential forces at the rotor tip are in the 1 milliNewton range with torque outputs of 1×10^{-6} Newton-meter for 20 mA of applied current.

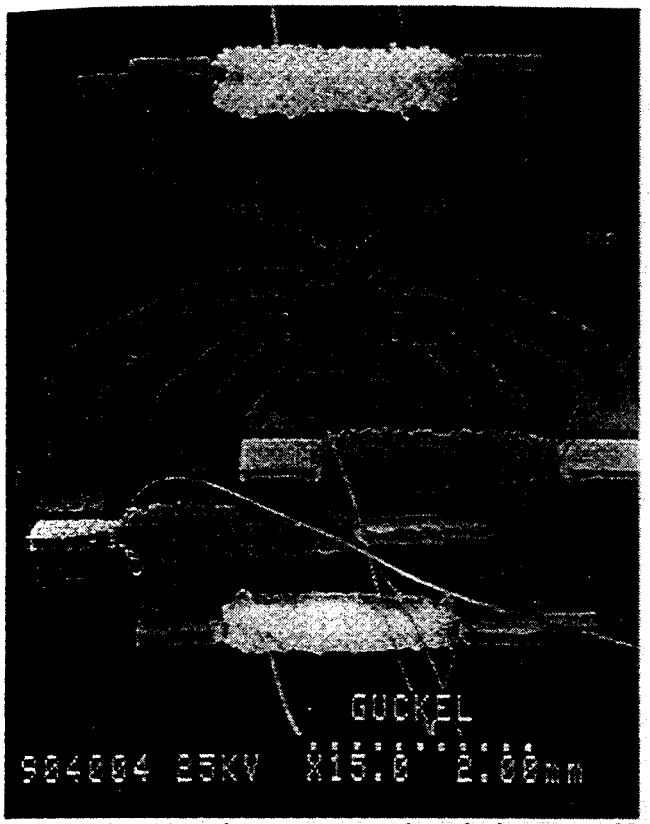


Figure 6: Microdynamometer with a 3-phase variable reluctance motor, idling gear, and magnetic break.

Figure 7: Peg assembly to increase structure height of a variable reluctance motor. Each layer is 250 μm thick.

Figure 7 shows a micro rotationary device using the assembled micro coil technology and stacking the magnetic path of the device. Micro assembly has allowed to increase structure height without increasing electroplating times. Structure height can be increased with the use of very high energy photons but electrodeposition durations can become increasingly long. An alternative is the peg assembly of stacked electroplated structures. Since this stacking uses assembly, submicron tolerances have been achieved. Each permalloy layer of figure 7 is 250 μm thick and have pegs of a different thickness dropped into the holes visable at the magnetic path edges. The resulting structure is 500 microns high but can be extended with more layers to significantly larger heights. The center post and coils (not shown in this example) have also been stacked to acheive thicker structures.

CONCLUSIONS

Recent extensions to the LIGA and LIGA-like processes have resulted in more efficient large output force actuators and precision machined parts. These new tools have not only simplified the design, fabrication, and testing process, but has also allowed new classes of microactuators and manufacturing techniques. Current work is focusing on the use of these tools to create microactuators, microsensors, and other microsystems.

REFERENCES

1. E.W. Becker, W. Ehrfeld, P. Hagmann, A. Maner, and D. Münchmeyer, "Fabrication of Microstructures with High Aspect Ratios and Great Structural Heights by Synchrotron Radiation Lithography, Galvanoforming, and Plastic Moulding (LIGA Process)," Microelectronic Engineering, vol. 4, pp. 35-56, 1986.

2. H. Guckel, K.J. Skrobis, J. Klein, T.R. Christenson, and T. Wiegele, "Deep X-ray Lithography for Micromechanics," Proceedings of *Québec '93 - SPIE* , Québec City, Canada, August 1993, pp. 15-21 .

3. D.P. Siddons, E.D. Johnson, H. Guckel, "Precision Machining with Hard X-rays," Synchrotron Radiation News, Vol 7, No 2, 1994, pp. 16-18.

4. H. Guckel, "Deep X-Ray Lithography for Micromechanics and Precision Engineering," Synchrotron Radiation Instrumentation, Advanced Photon Source, Argonne, October 19, 1995.

5. S. Mehdizadeh, J.O. Dukovic, P.C. Andricacos, L.T. Romankiw, "The Influence of Lithographic Patterning on Current Distribution: A Model for Microfabrication by Electrodeposition," Journal of Electrochemical Society, Vol 139, No 1, January 1992, pp. 78-91.

6. S. Massoud-Ansari, P.S. Mangat, J. Klein, H. Guckel, "A Multi-Level, LIGA-Like Process for Three Dimensional Actuators," IEEE MEMS 96, February 11-15, 1996, pp. 285-289.

7. H. Guckel, T. Earles, J. Klein, D. Zook, T. Ohnstein, "High Aspect Ratio Electromagnetic Linear Actuators for Micro Relay, Micro Valve and Precision Positioning Applications," HARMST '95, Karlsruhe, Germany, 3-5 July 1995.

8. T.R. Christenson, H. Guckel, J. Klein, "A Variable Reluctance Stepping Microdynamometer," HARMST '95, Karlsruhe, Germany, 3-5 July 1995.

MINIATURIZED SPECTROMETERS FOR BIOCHEMICAL ANALYSIS

Gaylin M. Yee[1], Paul A. Hing[2], Nadim I. Maluf[1,3], Gregory T. A. Kovacs[1]

[1]Stanford University
Center for Integrated Systems, CISX-205
Stanford, CA 94305-4075

[2]Applied Biosystems Division,
Perkin-Elmer Corporation
Foster City, CA 94404

[3]Lucas Novasensor
1055 Misson Court
Fremont, CA 94539

ABSTRACT

Miniaturized spectrometers were demonstrated by mounting micromachined diffraction gratings onto CCD imaging devices. Two implementations were tested: one for a high dispersion and sensitivity application, and the other for a low-cost consumer application. The first system showed a dispersion of 1.7 nm/pixel and a resolution of 74.4 nm/nm for the bandwidth of interest. The free spectral range of the device was designed to be 400-800 nm for this particular application. The diffraction efficiency of the system is 63%. The second, low-cost system demonstrated a dispersion and resolution of 2.55 nm/pixel and 69.8 nm/nm respectively. Results are shown for their applications in biochemical analysis. In addition, further optimization was sought by adding micromachined lenses and creating specialized gratings to compress and shape the spectral signal.

BACKGROUND

Spectrometers detect the chemical composition of a substance by measuring its absorption or fluorescence spectra. Light from the sample passes through a diffraction grating and disperses into a continuum of wavelengths. The longer the wavelength, the greater its angle of diffraction. By examining the intensities of the various frequencies, the substances can be identified by matching to the known spectral signatures of the chemical components.

Conventional benchtop spectrometers involve a complex system of lenses and moving parts, and are thus bulky and costly. By applying micromachining techniques it is possible to produce a spectrometer drastically reduced in size and intricacy. In addition, these systems are less expensive and compact, allowing them to be easily arrayed for improved spectral range and/or resolution.

Previously integrated spectrometers involve those formed with planar waveguides [1] or Fabry-Perot etalons [2]. Those that involve a diffraction grating have implemented bulk micromachined channels [3], amorphous silicon detectors on a transparent substrate [4], or free-standing gratings on a substrate [5]. However, many of these implementations are relatively involved and difficult to implement in volume.

Two implementations of a more manufacturable miniature spectrometer were demonstrated. Both devices are formed by mounting a transmission grating directly over a charge-coupled detector (CCD) imager, as shown in Fig. 1. The fine grating pitch creates high dispersion angles, thus providing sufficient spectral image dispersion without a large projection distance.

The first implementation was for applications in the detection of fluorescent dyes used in DNA separation [6]. Thus, both high sensitivity and dispersion were important, and so necessitated the use of a Peltier cooled CCD detector of 512x64 pixels. The second implementation was designed for use in low-cost applications such as colorimetry. A grating with wider pitch was mounted above a low-cost, easily available consumer CCD. In both applications, an external lens was used for light collimation and higher resolution and efficiency.

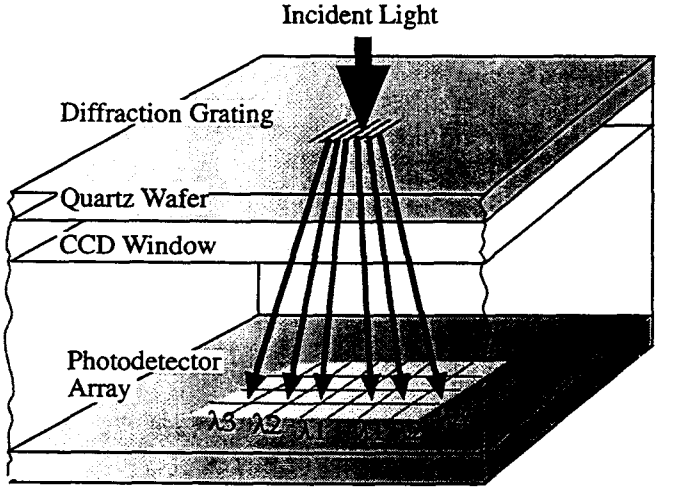


Figure 1: Cross-section of the spectrometer device. Light passes through the diffraction grating and disperses according to the wavelengths of their components, where $\lambda 3 > \lambda 2 > \lambda 1$. The spectral signals are recorded by the photodetector array for the determination of chemical substances.

0-9640024-1-8 $3.00 © 1996 TRF

Solid-State Sensor and Actuator Workshop
Hilton Head, South Carolina, June 2-6, 1996

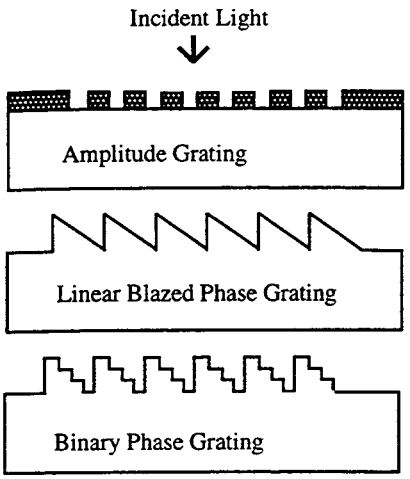


Figure 2: A schematic of basic grating types. Instead of blocking out sections of the light with opaque sections, phase gratings act to modulate its phase with variations in the grating shape. A binary step phase grating is a good approximation to the linear blaze, and is easily implemented with micromachining techniques.

THEORY

First order descriptions of diffraction gratings have long been available [7]. An amplitude grating consists of a simple array of slits, but has low efficiency in that a small percentage of light diffracts into the first peak, or spectral order. This can be remedied by using phase gratings or even blazed gratings, which have a theoretical efficiency of 100%. A blazed grating can easily be approximated in micromachining technologies by etching a series of steps, as shown in Fig. 2.

Dispersion, or the amount of lateral distance change with respect to wavelength, for an amplitude grating with pitch a is

$$D \equiv \left.\frac{dx}{d\lambda}\right|_{\theta=\theta_o} = \frac{m \cdot h}{a \cdot \cos^3 \theta_o} \qquad (1)$$

where m is the spectral order, θ_o is the angle of dispersion, and h is the optical distance between the grating and the detector. Resolution, or the sharpness of the resolved peaks, is simply

$$R \equiv \frac{\lambda}{\Delta\lambda} = \frac{N \cdot a \cdot \sin\theta_o}{\lambda} \qquad (2)$$

where N is the number of lines illuminated. Also of importance is the free spectral range, or the range of wavelengths which the spectrometer can measure. Finally, the diffraction efficiency may be defined as the amount of light incident on the grating which is converted into useful signal. Efficiency of a phase modulated grating shaped into a number I of binary steps is

$$\eta = \left| \operatorname{sinc}\left[\frac{m}{I}\right] \frac{\operatorname{sinc}\left[m - \frac{\phi_o}{2\pi}\right]}{\operatorname{sinc}\left[\frac{m - \frac{\phi_o}{2\pi}}{I}\right]} \right|^2 \qquad (3)$$

where λ_o is the wavelength designed for and λ is the actual wavelength used [7].

It has been shown [8] that ideal Fourier response of the overall system can be achieved if the light source is imaged onto the detector array. This was accomplished by placing the detector in the focal plane of a small lens. As a result, the resolution and the free spectral range were improved.

FABRICATION

The gratings were fabricated using chrome on quartz wafers. Because of the high resolution 0.4-0.5 μm step dimensions, the PMMA resist was exposed using direct write on a Hitachi HL700 electron beam system. After wet etching the chrome, the underlying quartz was RIE etched in oxygen. These processes were repeated for multiple steps in blazed gratings.

For the first implementation, a grating with 0.4 μm slit width and 0.8 μm pitch was mounted on the CCD top window, approximately 5.7 mm above the detector. The CCD consists of 512x64 pixels, each 27x27 μm in size. In the low-cost implementation, a grating of 0.5 μm slits and 1.0 μm pitch was mounted over a consumer grade CCD camera, QuickCam™ (from Connectix Corp., San Mateo, CA). The detector consisted of 320x240 pixels, each 11x10.4 μm.

EXPERIMENTAL RESULTS

The setup for the detection of DNA labeling dyes is shown in Fig. 3. The sample is excited using a laser source and the resulting fluorescence is focused onto the spectrometer. The spectra are shown in Fig. 4 showing the response of the system to the He-Ne laser and to a cyanine dye designed to emit at 657 nm. Using these two wavelengths for calibration, the measured dispersion and resolution are 1.7 nm/pixel and 74.4 nm/nm respectively. The free spectral range is 400-800 nm. At short wavelengths, the response is limited by the CCD quantum efficiency; at longer wavelengths, it is limited by total internal reflection at the lower interface of the CCD window. The range could easily be extended to longer wavelengths by altering the dimensions of the device. The diffraction efficiency of a phase grating for the system was measured to be 63%.

In addition, further sensitivity was obtained by adding a micromachined cylindrical lens on top of the grating. The lens was a binary 3-phase Fresnel zone plate which focused light in the direction perpendicular to that of the dispersion. As shown in Fig. 5, this increased the signal/noise ratio of the signal while requiring a narrower and less expensive photodetector

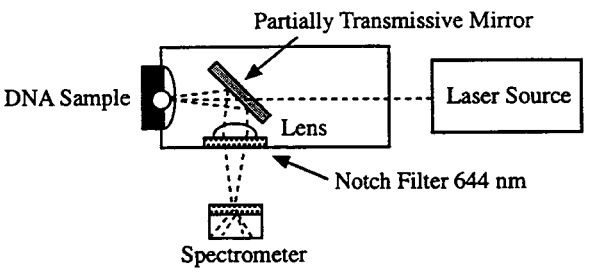


Figure 3: Detection setup involving sample, light source and spectrometer. A laser excites the sample and fluoresces its dyes. The resulting light is reflected by the mirror into a lens which focuses the light onto the spectrometer.

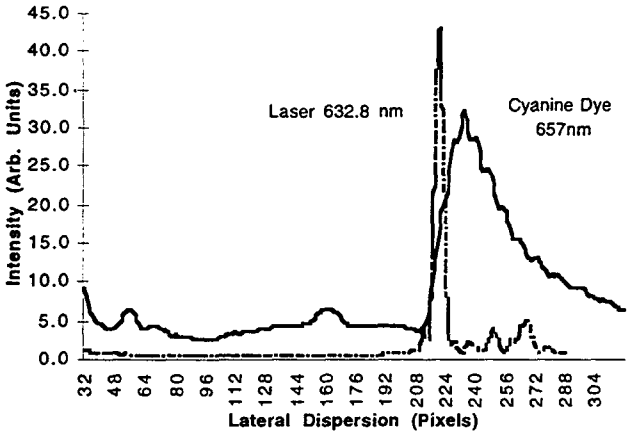


Figure 4: The response of the spectrometer showing the presence and intensity of the cyanine dye. A He-Ne laser spectra is provided for comparison.

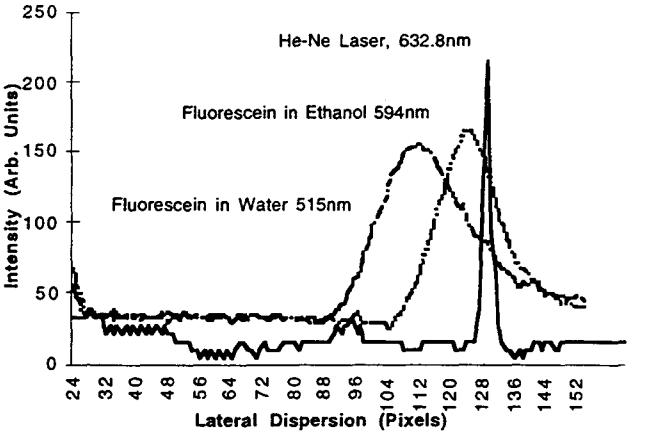


Figure 6: Response of the low cost spectrometer to the dye fluorescein. The spectrometer clearly distinguishes the peaks of the dye in water versus ethanol. A He-Ne laser spectrum is provided for comparison.

array. This is especially critical to increase the response and perhaps allow the use of a less expensive, non-cooled device.

The low-cost implementation was tested with the biochemical fluorescent dye known as fluorescein. The responses in two different solvents are shown as in Fig. 6. The dispersion was measured as 2.55 nm/pixel and the resolution as 69.8 nm/nm for the bandwidth of interest. The free spectral range is 500-800 nm, demonstrating less sensitivity at low frequencies and a cutoff due to CCD dimensions for the high frequencies. The results for both spectrometers are summarized in Table 1.

In the assembly of these devices, little attention was given to grating alignment, interface imperfections, and glass irregularities. These normally result in a loss of resolution. This problem will be remedied in the future by aligning the glass to the detector at the wafer level and using anodic bonding. Loss of efficiency was also observed (Table 1) and is attributed to decreased transmission through the glass at high diffraction angles and to lack of depth control in the etch of the gratings. These undesirable effects will be alleviated by improving process monitoring and characterization.

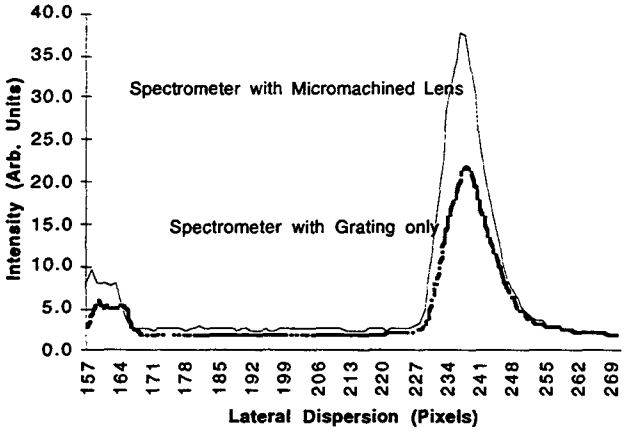


Figure 5: The improved signal from the more sensitive spectrometer with a micromachined cylindrical Fresnel lens. The lens was shaped using a three phase step blaze in order to focus the dispersing light into a single line.

SPECTRAL DETECTION USING CUSTOMIZED GRATINGS

The overall spectrometer system was further optimized by using custom gratings to shape the spectral signal. Earlier research in such masks has mainly focused on improving the imaging in optical lithography [9]. Here, the process was extended to provide expanded resolution in the spectrometry of biochemicals in a specific bandwidth of 550 nm-660 nm. Since only the first order spectral peaks on each side provide information, these gratings were designed to compress central "dead space" between the first order peaks. This was done by simulating the diffraction process operating in reverse and dithering the mask pixels to provide the desired response. Shown in Fig. 7 is a SEM micrograph of the resultant mask of 512 x 512 pixels of 0.4 μm width.

Results obtained using the nonlinear grating are shown in Fig. 8. With a source at 635 nm, the diffraction peaks for the conventional grating are 10.64 mm apart while those of the nonlinear grating are only 3.54 mm apart. This is nearly a third of that of the conventional grating, providing substantial savings in area. However, this reduction is at the cost of a loss of both the efficiency and local dispersion.

CONCLUSIONS AND FUTURE WORK

An easily manufacturable spectrometer has been demonstrated for use in biochemical analysis. The obtained dispersions and resolutions were sufficient to detect the

	Theory	Actual	
Results for the High-Dispersion Spectrometer			
Dispersion	1.5	1.7	nm/pixel
Resolution	113	74.4	nm/nm
Efficiency	80.2	63	%
Results for the Low-Cost Spectrometer			
Dispersion	2.59	2.55	nm/pixel
Resolution	100	69.8	nm/nm

Table 1: Summary of Results for both implementations of the spectrometer.

difference of several biochemical dyes. In addition, we have optimized the response by using either additional micromachined cylindrical lenses or custom, nonlinear gratings to shape the resulting spectral signal.

Future work will concentrate on an entirely integrated system, which will allow for scaling and manufacturability. The device will consist of a micromachined grating anodically bonded onto a CMOS imager. The alignment would be at the wafer level, providing improved alignment and tilt which in turn, will enhance resolution and dispersion.

ACKNOWLEDGMENTS

The authors would like to thank the Applied Biosystems Division of Perkin-Elmer and the National Science Foundation for supporting this research (Kovacs' NYI Award, Grant # ECS9358289, and Yee's Graduate NSF Fellowship). In addition, thanks go to Mark McCord, Chris Storment, Mike Albin, Kristen Pol, and Hao Xu for technical assistance and support. The e-beam system is also supported by a grant from the National Science Foundation.

REFERENCES

1. D. Goldman, P. White, N. Anheier, "Miniaturized Spectrometer Employing Planar Waveguides and grating Couplers for Chemical Analysis", *Applied Optics*, Vol.29, No. 31, (November 1990), pp. 4583-4589.

2. J. Holm-Kennedy, K. Tsang, W. Sze, F. Jiang, D. Yang, "A Novel Monolithic, Chip Integrated, Color Spectrometer", *SPIE*, vol . 1527, (1991), pp. 322-331.

3. T. A. Kwa and R. F. Wolffenbuttel, "Integrated Grating/Detector Array Fabricated in Silicon Using Micromachining Techniques", *Sensors and Actuators A*, vol 31 (1992), pp. 259-266.

4. Z.E. Smith, III and R. Street, "Spectral Resolving and Sensing Apparatus", *U.S. Patent No. 5,037,201*, 1991.

5. S.S. Lee, L.Y. Lin, and M. C. Wu, "Surface-Micromachined Free-space Micro-Optical Systems Containing Three-Dimensional Micrograting". *Applied Physics Letters*, vol.65, no. 15, (1995), pp. 2135-2137.

6. M. O'Neill, "Sequencers Benefit from Solid-State Detectors", *Laser Focus World*, (October 1995), pp. 135-142.

7. M. Bass, *Handbook of Optics, vol. II*, (New York: Optical Society of America, 1994), pp. 8.9-8.10.

8. J. W. Goodman, *Introduction to Fourier Optics*, (New York: McGraw-Hill, 1996), pp. 98-114, 210-214.

9. T. Terasawa, N. Hasegawa, H. Fukuda, S. Katagiri, "Imaging Characteristics of Multi-Phase-Shifting and Halftone Phase-Shifting Masks", *Journal of Applied Physics*, (November 1991), pp. 2991-2996,

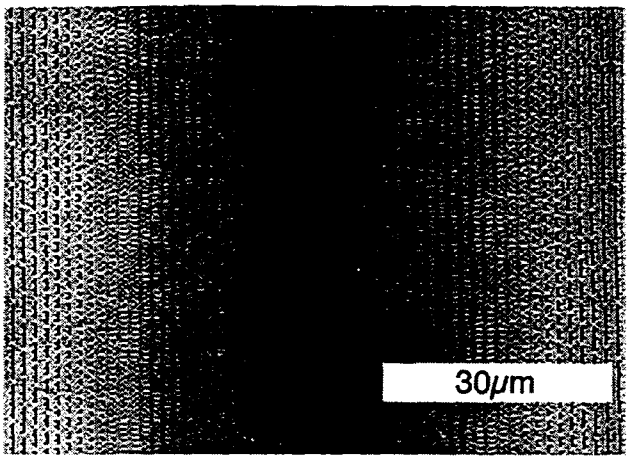


Figure 7: SEM photo of the computed custom grating, consisting of a 512x512 array of mask grating pixels 0.4×0.4 μm in size. The pixels are dithered vertically into those which are clear, chrome-covered, and etched. The diffracting light is dispersed horizontally.

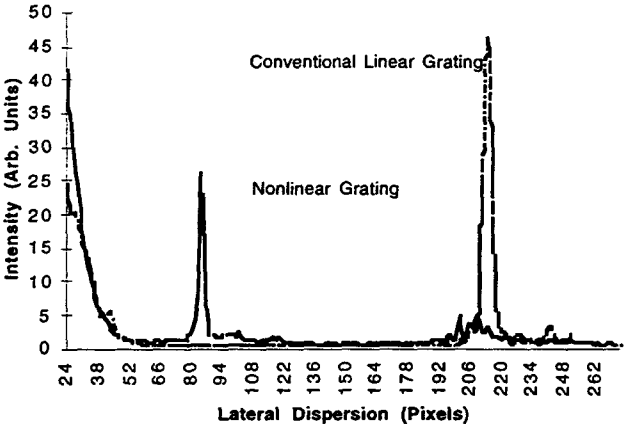


Figure 8: Comparison of the results of the customized nonlinear grating versus that of a standard linear grating. The first order peak has moved drastically towards the center, where a part of the incident beam is shown. Note the decreased amplitude, which denotes a loss of diffraction efficiency due to the large amount of mask pixels covered with chrome.

DESIGN, FABRICATION AND MEASUREMENT OF A TUNNELING TIP ACCELEROMETER

Jianchao Wang, Bob McClelland, and Paul M. Zavracky
Northeastern University, Department of Electrical and Computer Engineering
360 Huntington Avenue, Boston, MA 02115

Frank Hartley and Ben Dolgin
Jet Propulsion Laboratory
4800 Oak Grove Drive, Pasadena, CA 91190

ABSTRACT

We have designed and fabricated a bulk micromachined single axis accelerometer. The accelerometer is designed to have a full scale range of one milli-g and a sensitivity of 10^{-8}g. This is accomplished by using an ultrahigh sensitive tunneling tip as a displacement sensor. The device consists of four silicon dice which are assembled together using eutectic bonding technique. Each die is processed separately before bonding. The fabrication procedures for the three different dice are described Electrical contacts are made between layers during the bonding operation. The force feedback control approach is used for the accelerometer system. The clamping function for the parking mode is fully tested. The preliminary testing on the tip control function is performed. Initial results of acceleration measurement including the time domain response and power spectra are presented

INTRODUCTION

After the first report of a silicon accelerometer by Roylance and Angell [1] in 1979, many kinds of micromachined silicon accelerometers with different operating principles and structures have been developed. Among these sensors, piezoresistive [2] and capacitive [3, 4] accelerometers are the most common and well developed. These accelerometers have demonstrated good performance with sensitivities down to a micro-g. However, for applications below a micro-g, as in microgravity research in the space shuttle; measurement of orbital drag; and active aiming and isolation systems, the traditional position sensing methods have their limitations and a new position transducer must be used.

In this paper, we describe the development of a micromachined accelerometer which is designed to be capable of measuring 10^{-8} g. Our approach to ultrahigh resolution devices is to incorporate weak springs and a relatively large mass with an ultra-sensitive tunneling tip position detector. Electron tunneling through a vacuum barrier between a tip electrode and a plane electrode has been successfully used in Scanning Tunneling Microscopy (STM)[5] to achieve real space images on an atomic scale. With the development of micromachining technology, tunneling tips have also found applications in sensors with reported resolution as low as hundredths of an Angstrom [6-8]. This is due to the fact that the electron tunneling current is exponentially proportional to the electrode separation. Theoretical analysis has shown that for a typical tunneling spacing of 10 Å and a barrier height of 4 eV, the tunneling current increases by one order of magnitude for each 1 Å decrease in electron separation.

We also use a force feedback control in which the position of the proof mass is held constant. This is accomplished by feeding back position information to the control electrodes. Using force feedback can improve the linearity, increase the dynamic range, and stabilize the system. Key elements of our design include the optimized placement of the tip, the minimization of off-axis sensitivity and tip protection for over loading.

DESIGN FEATURES

An accelerometer usually consists of a proof mass to convert an acceleration to a force, a spring to balance the acceleration force and a position detector to sense the position changed due to acceleration. When the system is experiencing an acceleration, the proof mass will move from its original position to a new position which is determined by the balance between inertial force on the proof mass and the spring force. With a simple mechanical spring, the acceleration will be directly proportional to the distance traversed by the proof mass from its equilibrium position. In order to build an micro-gravity accelerometer, a large proof mass, weak springs and an ultra-sensitive position detector are needed. In this design, the use of an electron tunneling tip is fundamental to its superior performance. The basic design features and specifications for the accelerometer are listed in Table 1 and Table 2, respectively.

Table 1
Design Features

Electron Tunneling Tip Position Detection
Electrostatic Force Feedback
Pitch and Roll Control
High Off-axis Insensitivity

Table 2
Specifications

Parameter	Specification
Full-Scale Sensitivity	10mg
Resolution	1×10^{-8}g
Off-Axis Sensitivity	1×10^{-9}FS
Bandwidth	20Hz
Built in Tunneling Tip Protection	>32g

The required proof mass is achieved by producing a relatively large (~2 cm^2) bulk micromachined device in two separate wafers that are eutectically bonded together. The proof mass is held to the surrounding frame by a set of springs referred to as 'crab legs' [9]. These weak springs are fabricated by etching back the silicon to form 25 μm thick diaphragms which are further patterned to form 2 cm long folded spring elements. The resulting spring constant is only 1.7 N/m. Since the tunneling tip must be held in near proximity (10 Å) to the proof mass during operation, electrostatic force feedback has been employed. This is coupled to the proof mass on one side through a matrix of four plates that surround the tunneling tip and a single force plate on the other.

0-9640024-1-8 $3.00 © 1996 TRF

Solid-State Sensor and Actuator Workshop
Hilton Head, South Carolina, June 2-6, 1996

An important design issue is the minimization of off-axis sensitivity. With the unique design of the crab leg springs, calculation shows that the lateral spring constant is nearly five orders of magnitude higher than the normal spring constant. Further minimization of the off-axis sensitivity can be accomplished by ensuring that the proof mass and spring design is symmetric, and the spacing between the neutral position of the proof mass and the top of the tip is zero. As a consequence, thermal sensitivity is also reduced. To further reduce thermal sensitivity, all dice used in the device are fabricated in silicon. This choice minimizes stresses that can arise from thermal mismatch. Another issue considered in our design is that of tip protection. The single force plate doubles as a parking electrode designed to attract and hold the proof mass in place using a low voltage(<15 V).

FABRICATION

The accelerometer [10] is fabricated using bulk micromachining technology. The cross sectional view of the finished device is shown in Fig.1. It is constructed from four dice, with the proof mass assembled from two identical dice. Three different processes for the proof mass, force plate and tip plate are described below.

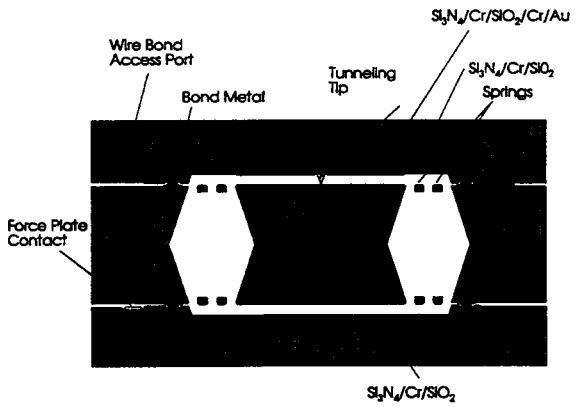


Figure 1. Cross-section of an assembled accelerometer.

The proof mass wafers (Fig. 2) are back etched in 25% KOH to define the proof mass and separation channel. At the same time, vias are etched to permit wire bonding through the assembled structure. The back side etch continues until the remaining diaphragm is 25 microns thick. This is a timed etch and no dopant stop is used in order to avoid strain related problems. The wafer is stripped back to silicon and recoated with silicon nitride. TiW/Au is deposited on the front side and patterned appropriately. This metallization provides the electrical contacts to the surface of the proof mass. Additional electrical leads are used to align and route electrical connections through the structure. The wafers are repatterned to the spring mask and etched anisotropically in a plasma etcher. This deep plasma etch releases both the springs and the dice. Two proof mass dice are then cleaned and bonded together eutectically.

The force plate wafer is deep etched in a manner similar to the proof masses to define the separation channel. Each wafer contains a grounded bond ring that surrounds the central well and is electrically isolated from the other electrical features. The force plate wafer contains a single 5 μm deep well in the center of the die. A two metal system is employed. The first metal is Chromium which is deposited by sputtering and patterned lithographically. It is then covered with an LTO layer which is patterned to open vias to the first metal. The second metal is TiW/Au which is patterned to form the bond regions and the wire bond pads. When completed, the force plate die is bonded to one of the assembled proof mass die in such a way that the side of the proof mass die with the springs is the side that is bonded.

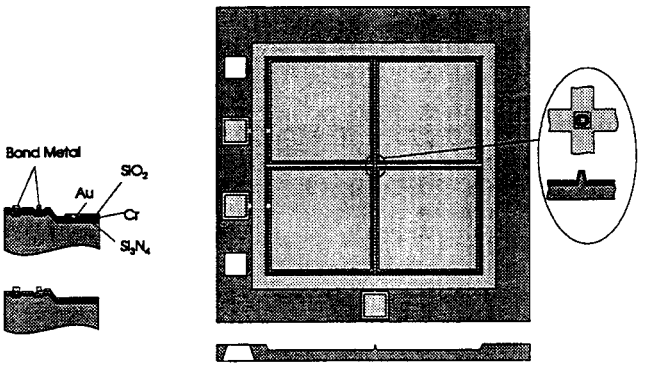


Figure 2. Various views of the tip die.

The tip dice (Fig. 2) undergo a similar initial process. The backside is deep etched in KOH for die separation purposes. The

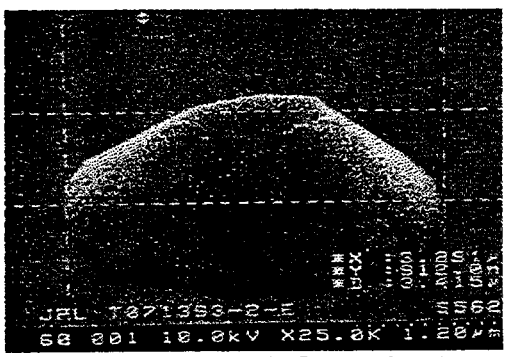


Figure 3. An SEM micrograph of a flat tunneling tip.

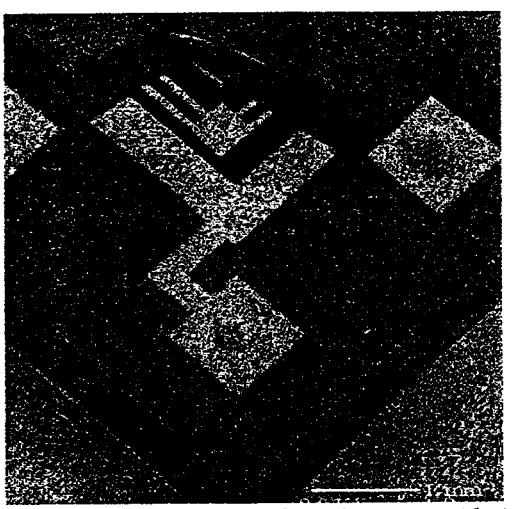


Figure 4. Top view of completed accelerometer with tip die cut away to reveal the internal structure using a SEM.

tip is created by a sequence of dry and four anisotropic wet etch steps which result in a flat topped tip. The tip is surrounded by four electrodes and sits in a well. Grooves are also provided for a eutectic bond ring. Again, a two metal process similar to that of the force plate is used. Electrical leads to the force plates and tip

connect to contact pads at the periphery of the die. A final plasma etch process is used to separate individual die. The completed tip die is eutectically bonded to the proof mass. Since the tip die is placed above the proof mass during assembly, the proof mass is held away from the tip by gravity. This prevents damage to the tip during bonding. The SEM pictures for the tunneling tip and part of the completed accelerometer are shown in Fig. 3, and Fig. 4, respectively.

EXPERIMENTAL RESULTS

The accelerometer has four modes of operation: parking, transition to operation mode under capacitance control, operational mode (tunneling control), and overload protection. The device has been fully tested in the parking mode. Tunneling tip control has also been tested.

Parking of the proof mass protects the fragile structures (tunneling tip and the flexures) from damage during shipment, handling, and spacecraft launch. The ability of the parking mechanism to withstand the high load was tested by applying a large force directly to a parked proof mass (the quadrant plate was removed). Fig. 5 shows the value of the force required to dislodge the proof mass as a function of the applied voltage. The force is expressed in units of acceleration of the proof mass (F = ma). The theoretical curves correspond to zero and 1μm average separation between the proof mass and the force plate. No stiction of the proof mass was observed after parking for as long as four days.

The feasibility of the tunneling tip control was tested by using a simple control circuit in a 1g field. The block diagram for the control system is shown in Fig.6. The closed-loop control system is linearized at the operating point where the tunneling spacing is 10Å. It should be noted that the bias voltage, V, used to

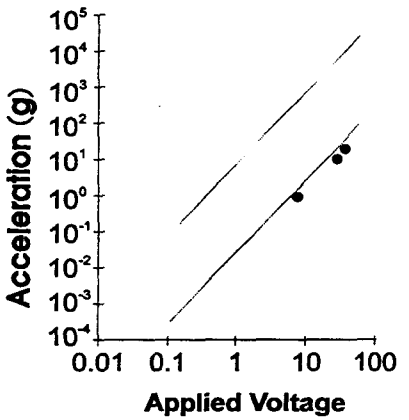


Figure 5. Experimental and theoretical data showing the maximum achievable acceleration against which the force plate can hold as a function of the voltage difference between the force plate and proof mass.

offset the background 1 g gravitational field, influences the magnitude of the effective spring constant for the mechanical system and the return ratio for the system. It can be shown that the system is stable as long as the damping term for the accelerometer is large. This is easily accomplished by squeeze film damping when operating at one atmosphere. A circuit board was fabricated and included a special zero insertion force connector for the packaged accelerometer and the control circuit. The unit was slowly turned to a position where the proof mass is resting on the tunneling tip. At this point, the tunneling current electronics would register a high current. The feedback loop was turned on next. A noise spectrum as shown in Fig. 7 is obtained using a HP dynamic signal analyzer with the device operated at

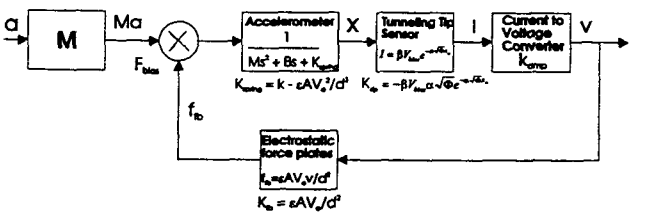


Figure 6. A block diagram of the force feedback control system used to measure the tunneling tip accelerometer.

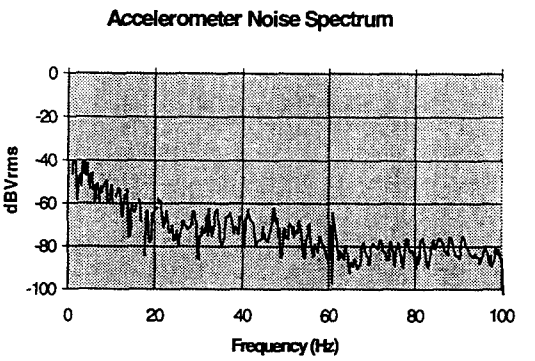


Figure 7. Accelerometer noise spectrum for a frequency range from 1Hz to 100Hz.

atmosphere. It was found that the flicker noise dominates the low frequency region, where the highest value corresponds to 2 milli-g/Hz$^{1/2}$. Comparing the result from the reference accelerometer and the noise from the circuit, we believe that this 1/f noise comes from the electron tunneling process. It may result from the migration of atoms on the active electrode surface, surface diffusion of adsorbed molecules and the adsorption-desorption process[11].

The acceleration response of the tunneling tip accelerometer was tested on a vibration isolation stage. The stage is constructed using a 100 lb platform suspended by 64 6-feet long bungee cords. The tunneling tip accelerometer is mounted on a thrust stand sitting on the platform. A reference accelerometer (Sunstrand QA3000) with sub-μg sensitivity was also installed on the stand for calibration purpose. The motion of the unit including both accelerometers was controlled by a piezo-driver which was able to generate sinusoidal accelerations at frequencies from DC to about 200 Hz. Fig. 9 shows the responses of both the tunneling tip accelerometer and the Sunstrand QA3000 at 53 Hz with an amplitude of 80 mg. Data from the QA3000 was collected with a 2nd order Butterworth filter during the measurement whereas the data for the tunneling tip device was unfiltered. The data for the tunneling tip device shows a higher frequency (628 Hz) signal overlaid on the 53 Hz response. This may be a resonance mode of the spring system. For the measurement at much lower frequencies, the tunneling tip accelerometer is put into a vacuum holder in order to keep the device clean and reduce the noise. The bandwidth of the measurement system is set to 10 Hz. The data is then digitized and converted to power spectra using Fast Fourier transform (FFT) techniques. One of the results is shown in Fig. 9, where the exciting frequency is 0.2 Hz and the magnitude is 1 mg. In order to obtain the power spectrum for a 0.2 Hz signal, the measurement time would be at least 8 minutes. It can be seen that the two responses from the tunneling tip accelerometer and the QA 3000 are consistent in the vicinity of the highest peak (0.2 Hz).

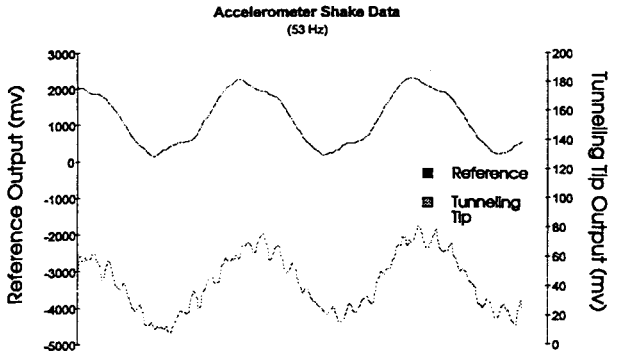


Figure 8. Accelerometer data comparing a Sunstrand QA3000 (reference) with the tunneling tip accelerometer.

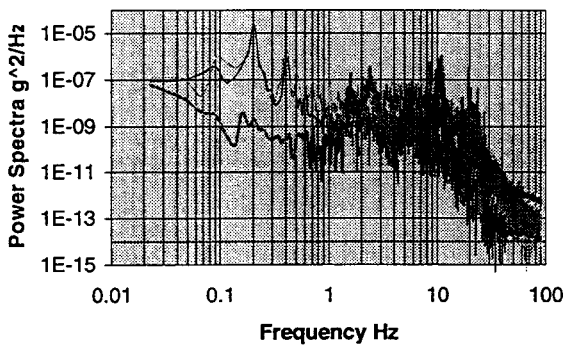


Figure 9. Power spectra for the responses to a 0.2 Hz 1 mg excitation. The dark one of the upper two curves is the response of the tunneling tip accelerometer. The light one is from QA3000. The lowest curve is the background noise.

CONCLUSIONS

We have reported the design considerations, fabrication procedures and some experimental results for an ultra-sensitive accelerometer. The accelerometer is fabricated using a bulk micromachining approach. Three distinctly different dice are processed separately and then eutectically bonded together to form the final device. The bonding process also provides hermetic sealing and electrical interconnection between dice. The accelerometer is designed with a force plate that can be used to hold the proof mass and spring assembly electrostatically during transportation. The capability of the electrostatic clamping mechanism has been demonstrated. A linear closed-loop control model for the force feedback accelerometer system is presented in a block diagram. From preliminary measurement, we found that the device suffers from 1/f noise for operations in atmosphere. The response to a sinusoidal excitation in time domain is demonstrated and compared with a calibration accelerometer. The acceleration measurement at very low frequencies was performed. The resulted power spectrum shows good agreement between the tunneling tip accelerometer and the reference accelerometer in the vicinity of the signal frequency.

ACKNOWLEDGMENTS

The authors would like to acknowledge the help and support of JPL. We also appreciate the assistance of Keith Warner in reviewing this document.

REFERENCES

1. L. M. Roylance and J. B. Angell, "A Batch-Fabricated silicon accelerometer", *IEEE Trans. Electron Devices*, Vol. ED-26, No. 12 December 1979, pp. 1911-1917.

2. H. Seidel, U. Fritsch, R. Gottinger, J. Schalk, J. Walter and K. Ambaum, "A Piezoresistive Silicon Accelerometer with Monolithically Integrated CMOS-Circuitry", *The 8th Int. Conf. on Solid-State Sensors and Actuators, and Eurosensors IX,* Stockholm, Sweden, pp. 150-C4, June 25-29, 1995.

3. F. Rudolf, A. Jornod, J. Bergqvist and H. Leuthold, "Precision Accelerometers With μg Resolution", *Sensors and Actuators,* A 21-A23, 1990, pp. 297-302.

4. E. Peeters, S. Vergote, B. Puers and W. Sansen, "A highly symmetrical Capacitive Micro-Accelerometer With Single Degree of Freedom Response", *J. Micromech. Microeng.* Vol. 2, pp. 104-112, 1992.

5. Y. Kulk and P. J. Silverman, "Scanning Tunneling Microscope Instrumentation", Rev. Sci. Instrum. 60(2), February 1989, pp. 165-180.

6. T. W. Kenny, W. J. Kaiser, H. K. Rockstadm J. K. Reynold, J. A. Podosek, and E. C. Vote, "Wide-Bandwidth Electromechanical Actuators for Tunneling Displacement Transducers", *J. MEMS.* Vol. 3, No. 3, pp. 97-104, September 1994.

7. T. W. Kenny, W. J. Kaiser, S. B. Waltman, and J. K. Reynold, "A Novel Infrared Detector Based on a Tunneling Displacement Transducer", *Appl. Phys. Lett*, Vol. 59, pp. 1820-1822, 1991.

8. H. K. Rockstad, J. K. Reynold, T. K. Tang, T. W. Kenny, W. J. Kaiser, and T. B. Gabrielson, "A miniature, high-sensitivity, electron tunneling accelerometer", *Transducers'95*, pp407 (PC 12).

9. P. M. Zavracky, F. Hartley, and D. Atkins, "New Spring Design and Processes for an Electron Tunneling Tip Accelerometer", to be published.

10. P. M. Zavracky, F. Hartley, N. Sherman, T. Hansen, and K. Warner, "A New Force Balanced Accelerometer Using Tunneling Tip Position Sensing", *7th Int. Conf. on Sensors and Actuators*, Yokahama, Japan, June 7-10, 1993, pp. 50-51.

11. T. Tiedje, J. Varon, H. Deckman and J. Stokes, "Tip Contamination Effects in Ambient Pressure Scanning Tunneling Microscopy Imaging of Graphite", J. Vac. Sci. Technol. A 6(2), Mar/Apr., 1988, pp. 372.

DAMPING CONTROL OF MEMS DEVICES USING STRUCTURAL DESIGN APPROACH

C.-L. Chen and J. Jason Yao
Rockwell Science Center
1049 Camino Dos Rios
Thousand Oaks, CA 91360

ABSTRACT

The transient responses or dynamics of MEMS parallel-plate, squeeze type devices have been simulated numerically in the present work. Emphasis is given to the study of the micro-fluidynamic damping effect by solving the Reynolds equation with a slip boundary condition at the structure surfaces. This approach couples the micro-fluidynamics into the dynamic simulation of MEMS devices in a CPU efficient way, and allows a damping effect simulation at every time step. The computed results indicate that the damping behavior can be altered using a structural design approach which employs structural modification by creating various "holes" in the paddle-like plate structure and by varying the hole distributions on the paddle. The transient time for one of the test structures, for example, is reduced from 9 ms in the without-hole case to 70 μs in the with-hole case. These simulation results have been qualitatively supported by experimental data.

INTRODUCTION

For physical and cost reasons, MEMS devices are seldom operated in high vacuum. The performance of MEMS devices is thus strongly affected by the surrounding fluids, e.g., air. Consequently, understanding the fluid behavior at micro scale, modeling the fluid effects, coupling it into the dynamic system, and tracing the dynamic response are crucial steps in the design of MEMS devices for practical applications. Several works have addressed fluid effects on various MEMS devices with various numerical approaches [1-4]. The Direct Simulation Monte Carlo (DSMC) type approach is still very CPU intensive. In this work, we present a method of computing micro-fluidynamic behavior for MEMS parallel-plate, squeeze type devices to obtain the damping load by solving the Reynolds equation with slip-flow consideration. The damping force acting on the cantilevered paddle structure is a function of the paddle velocity, the gap, the Knudsen number $Kn = \lambda/h$ (where λ is the mean free path and h is the fluid film thickness), the surrounding fluid properties, and the paddle configuration. Our approach couples the micro-fluidynamics into the dynamic simulation of MEMS devices in a CPU efficient way, and is able to provide a damping effect simulation at every time step using a single workstation. Here we focus on examining the resulting effects of various structural changes on the dynamics of electrostatically driven plate structures. The change of structural design is made by creating "holes" on the moving paddle and by changing the hole configuration on the paddle.

ANALYSIS

The model is shown in Fig. 1. The gap h between the paddle and the substrate is assumed to be much smaller than the paddle dimensions in the x and y directions. The paddle is supported by a spring (not shown in Fig. 1) and can move vertically in the z direction. The gas flow is assumed laminar and primarily viscous. The relationship between pressure and density at any point in the gas film is described by a polytropic process with exponent n: p/ρ^n = constant.

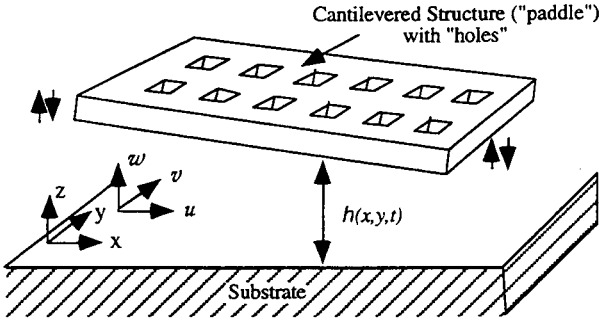


Figure 1. A schematic illustration showing the model to be simulated.

Suppose that function $F(x,y,z,t) = 0$ represents the equation for a moving boundary surface, that is

$$F = z - h(x,y,t) = 0 \tag{1}$$

When the solid boundary is in motion, the fluid velocity component normal to the boundary surface equals the velocity component of the surface normal to itself, but the flow is still allowed to slip in the tangential direction along the boundary surface. That is,

$$\mathbf{V} \cdot \mathbf{n} = \mathbf{U_b} \cdot \mathbf{n} \tag{2}$$

where $\mathbf{n}$ is the unit vector normal to the surface, $\mathbf{V}(u,v,w)$ is the velocity of the fluid particle at the paddle boundary and $\mathbf{U_b}$ denotes the velocity of the points on the boundary with which the fluid particle is in contact. Using the above equations, we can obtain the vertical component, w, of the velocity vector at the moving boundary $h(x,y,t)$ as

$$w = u\frac{\partial h}{\partial x} + v\frac{\partial h}{\partial y} + \frac{\partial h}{\partial t} \tag{3}$$

which is the general kinematic boundary condition equation. Using this equation and integrating compressible continuity equation with respect to z and applying Leibnitz's theorem, we obtain

$$\frac{\partial \rho h}{\partial t} + \frac{\partial}{\partial x}\int_0^h \rho u\,dz + \frac{\partial}{\partial y}\int_0^h \rho v\,dz = 0 \tag{4}$$

For x momentum (also similarly for y momentum), under the assumptions of incompressibility, constant pressure across the fluid film in the squeeze direction, and small Reynolds number and fluid inertia, the above equation may be approximated to

$$\frac{dp}{dx} = \mu\frac{d^2u}{dz^2} \tag{5}$$

where μ is the viscosity. Using slip flow boundary conditions[5], $u = U_s$ at $z = 0$, and $u = -U_s$ at $z = h$, and

$$U_s = \lambda\alpha\frac{du}{dz}\Big|_{wall},\quad \alpha = \frac{2-d}{d} \tag{6}$$

0-9640024-1-8 $3.00 © 1996 TRF

where λ is the mean free path, and d is the diffuse reflection coefficient, the velocity profile becomes

$$u = \frac{h^2}{2\mu}\frac{dp}{dx}(\bar{z}^2 - \bar{z} - \alpha K_n),\ \bar{z} = \frac{z}{h} \quad (7)$$

Similar procedure can be repeated for v in the y direction. Therefore, Eq. 4 becomes

$$12\mu\frac{\partial p^{\frac{1}{n}}h}{\partial t} - \frac{n}{n+1}\left(\frac{\partial}{\partial x}\left(h^3(1+6\alpha K_n)\frac{\partial}{\partial x}p^{\frac{n+1}{n}}\right)\right. + \left.\frac{\partial}{\partial y}\left(h^3(1+6\alpha K_n)\frac{\partial}{\partial y}p^{\frac{n+1}{n}}\right)\right) = 0 \quad (8)$$

If $h = h(t)$ only in the *parallel* plate case (Special Case #1), the equation can be further simplified to

$$h\frac{\partial p^{\frac{1}{n}}}{\partial t} - \frac{h^3(1+6\alpha K_n)}{12\mu}\frac{n}{n+1}\left(\frac{\partial^2}{\partial x^2}p^{\frac{n+1}{n}} + \frac{\partial^2}{\partial y^2}p^{\frac{n+1}{n}}\right) = -p^{\frac{1}{n}}\frac{\partial h}{\partial t} \quad (9)$$

For small pressure fluctuation (Special Case #2), pressure may be approximated as $p = p_a + p'$, in which p_a is the ambient pressure and p' is the pressure perturbation. In the isothermal process (n = 1), the equation can be written as

$$p_a\frac{\partial h}{\partial t} + h\frac{\partial p'}{\partial t} - \frac{h^3(1+6\alpha K_n)}{12\mu}p_a\left(\frac{\partial^2 p'}{\partial x^2} + \frac{\partial^2 p'}{\partial y^2}\right) = 0 \quad (10)$$

Under the condition that $h\,\partial p'/\partial t$ is negligible (Special Case #3), the equation is a linear Poisson's equation in which the forcing term strength is proportional to $(\partial h/\partial t)(1/h^3)$. Therefore, for each configuration, Eq. 10 needs to be solved only once for a unit forcing term. The rest of the solutions can be obtained by superposition.

NUMERICAL APPROACH

Here we define an unknown variable q for $p^{(1/n)}$ and solve for q numerically. The general case Eq. 8 can be rewritten as

$$\frac{\partial q}{\partial t} = \mu'\left(\frac{\partial}{\partial x}\left(\alpha'\frac{\partial F}{\partial x}\right) + \frac{\partial}{\partial y}\left(\alpha'\frac{\partial F}{\partial y}\right)\right) - \gamma q \quad (11)$$

where $F = q^{n+1}$. Note that Eqs. 9-10 are special cases of Eq. 8. Thus, the numerical approach described here to solve Eq. 8 is also suitable to solve Eqs. 9-10. The Runge Kutta 4th order scheme for 1-D diffusion equation (2nd order central spacial difference) suffers CFL limit of 0.695, whereas the trapezium formula is unconditionally stable. The above equation, discretized using trapezium formula (except for the source term) and coupled with central spacial difference, becomes:

$$q_{i,j}^{m+1} - q_{i,j}^{m} = \frac{\Delta t}{2}\mu'\left(\frac{\delta_x^-\alpha'\delta_x^+F_{i,j}^{m+1}}{\Delta x^2} + \frac{\delta_y^-\alpha'\delta_y^+F_{i,j}^{m+1}}{\Delta y^2} + \frac{\delta_x^-\alpha'\delta_x^+F_{i,j}^{m}}{\Delta x^2} + \frac{\delta_y^-\alpha'\delta_y^+F_{i,j}^{m}}{\Delta y^2}\right) - \Delta t\gamma q_{i,j}^{m} \quad (12)$$

where m is the time step index, Δx and Δy are the grid spaces and $\delta_x^+ F_{i,j} = F_{i+1,j} - F_{i,j}$, $\delta_x^- F_{i,j} = F_{i,j} - F_{i-1,j}$. Through linearlization and factorization, the equation can be written in Δ-form:

$$\left(1 - \frac{\Delta t\mu'}{2\Delta x^2}\delta_x^-\alpha' A_{i,j}^m\delta_x^+\right)\left(1 - \frac{\Delta t\mu'}{2\Delta y^2}\delta_y^-\alpha' A_{i,j}^m\delta_y^+\right)\Delta q = \Delta t\left(\mu'\left(\frac{\delta_x^-\alpha'\delta_x^+F_{i,j}^m}{\Delta x^2} + \frac{\delta_y^-\alpha'\delta_y^+F_{i,j}^m}{\Delta y^2}\right) - \gamma q_{i,j}^m\right), \quad (13)$$

where $A = dF/dq = (n+1)\,q^n$ and $\Delta q = q^{(m+1)} - q^m$. The codes have been written for generalized coordinates which allow nonuniform grid stretching and have been validated by solving model equations such as Poisson's equation, heat equation and $\partial^2 u/\partial x^2 = u$. The accuracy of the algorithm has been verified. For the paddle structure with holes, the pressure within holes is assumed ambient pressure. The simple one-zonal grid approach is used. In the hole regions, the right-hand side of Eq. 13 is masked to be zero; the corresponding diagonal term is set to be one, and off-diagonal terms are zeros. Once q is obtained, the damping force F_d can be integrated to couple with the dynamic model.

The one-dimensional dynamic model for the electrostatically driven paddle structure above the substrate can be expressed as

$$m\frac{d^2h}{dt^2} + k(h - h_i) = F_e + F_d \quad (14)$$

where m is mass, k is spring constant, and h_i is the initial paddle-to-substrate gap with no spring force. The electrostatic force $F_e = \varepsilon A\,V^2/2h^2$, where ε is the permittivity and V is the applied voltage. The 4th order Runge-Kutta method is used in the time integration.

RESULTS AND DISCUSSION

In the present work, the initial displacement and velocity of the paddle are set to zero. The voltage is a step function applied at $t = 0$. The permittivity $\varepsilon = 8.9\times10^{-12}$ F/m; the area of the paddle, A, is approximated with 50x100 μm^2 regardless of the hole area; k is set to be 0.1 N/m; the mass of the paddle, m, is 1.15×10^{-11} kg; the gap between paddle and substrate is 5 μm. The grid point used here is 101x51 in the x-y plane.

The effects of slip and no-slip boundary conditions are compared in Fig. 2. The diffusion reflection coefficient, d, used here is 0.9 and the mean free path, λ, of air is 0.0635 μm. Applying slip boundary condition effectively reduces the strength of the forcing term in the Poisson's equation, and consequently reduces the damping force in the dynamic system. When the applied voltage is 8.7 V, two stable steady-state solutions exist. These two stable solutions are qualitatively illustrated in Fig. 3(a), labeled "A" and "B", with the critically stable solutions illustrated in Fig. 3(b). In the stable solution "A", as one reduces the gap between the paddle and the substrate, i.e., moving towards the right on the x axis in Fig. 3(a) away from point "A", the spring back force becomes greater than the electrostatic force F_e for a fixed applied voltage, and thus pulling back the paddle structure to point "A", making point "A" a stable solution. With the same analogy, it is obvious that point "C" is an unstable solution. As one further increases the applied voltage, the electrostatic force curve in Fig. 3 shifts upward, forcing point "A" and "C" move closer. When reaching a critical voltage, point "A" merges with point "C", and makes that a critically stable solution as illustrated in Fig. 3(b). After passing this critical voltage, the cantilevered paddle structure will converge to the only remaining stable steady-state solution labeled "B" in Fig. 3. This jump from the critically stable solution (A/C) to the stable solution "B" is commonly observed as the "snapping" motion in the parallel plate system, and occurs at one-third the distance of the total, unforced separation gap.

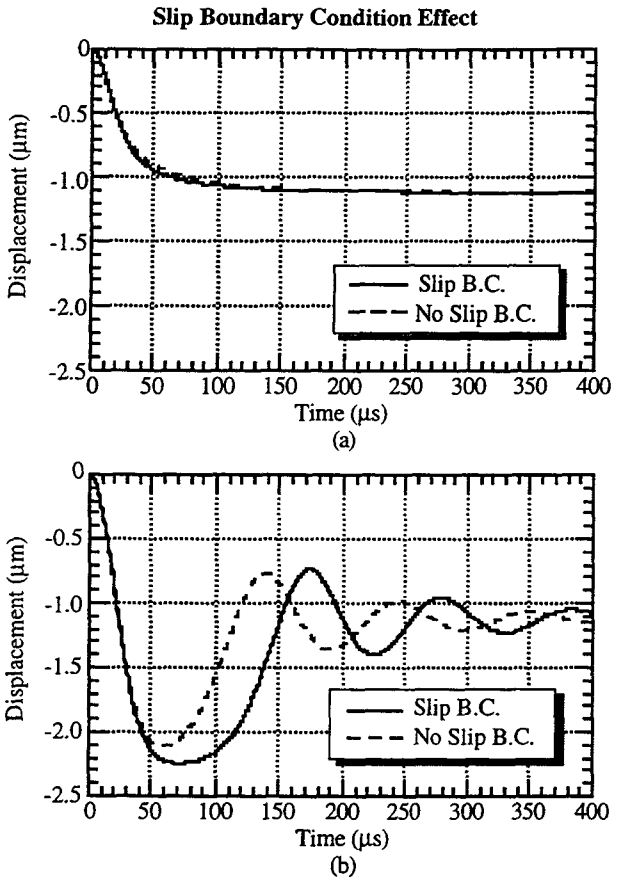


Figure 2. Simulation results showing the effect of slip boundary condition for the over-damped (no-hole paddle) case in (a) and for the under-damped (8-hole paddle) case in (b).

Using current initial condition, the results in Fig. 2 show that the paddle converges to the solution with smaller displacement (Point "A" in Fig. 3). By further reducing the diffuse reflection coefficient, e.g., from 0.9 to 0.8, or increasing λ, the damping force is reduced, and the paddle motion in response to the step input has been observed to converge to the other branch solution "B" and hit the substrate.

When the same voltage of 8.7 V is applied to various structural designs (no-hole, two-hole and eight-hole), the paddle can move with no-oscillation, one-oscillation, and under-damping oscillation as shown in Fig. 4. The corresponding contours of the pressure load when the transient damping force is at its maximum are shown in Fig. 5. The lighter contour color denotes lower pressure. The maximum pressure always occurs remote to the hole and to the paddle fringe edges. The maximum pressure is $1.000186p_a$ for the no-hole paddle, $1.000168p_a$ for the 2-hole paddle, and $1.00008p_a$ for the 8-hole paddle. The results indicate that the holes effectively relieve the pressure buildup in the gap, and thus reduce the damping force.

For the example system described in this paper, the critical voltage occurs at 9.123 V, at which one of the two stable steady-state solutions (point "A") merges with the unstable solution (point "C"), and only one stable steady solution (point "B") remains. Therefore, the paddle structure moves across the separating gap and comes in contact with the substrate. The transient behaviors of the no-hole design and of the 8-hole design are compared in Fig. 6 with an applied voltage of 9.124 V. The slow contact of the no-hole paddle is dramatically improved with the 8-hole design. The contact time is reduced from 8.9 ms to 67 μs, a more than 100 times reduction. The experimental data has qualitatively verified these results [6], and Fig. 7 shows two SEM images of a test structure.

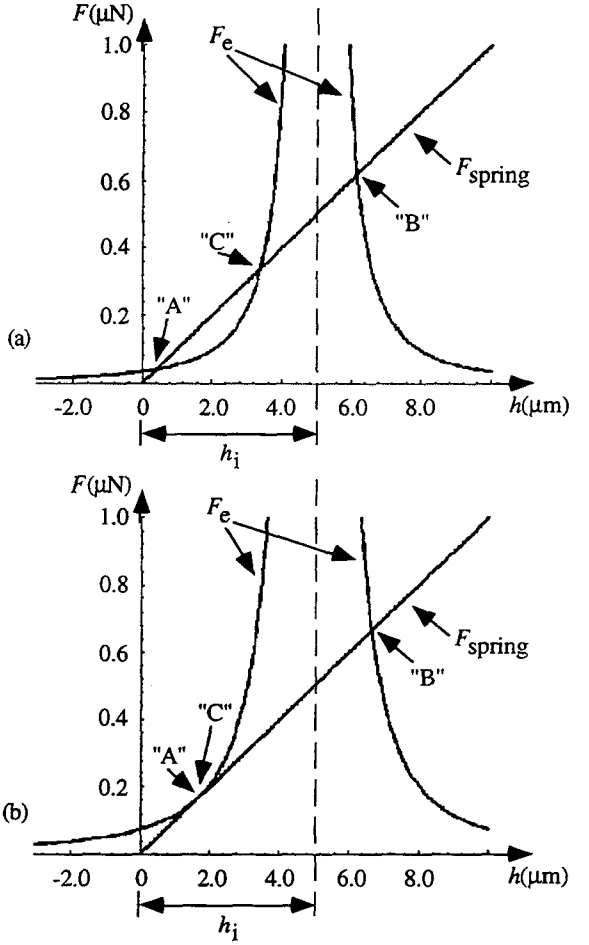


Figure 3. Plots of the electrostatic force and the spring-back force as a function of the separating gap between two parallel electrodes showing the two stable solutions "A" and "B", and one unstable solution "C". As the points "A" and "C" merge, shown in (b), the solution becomes critically stable.

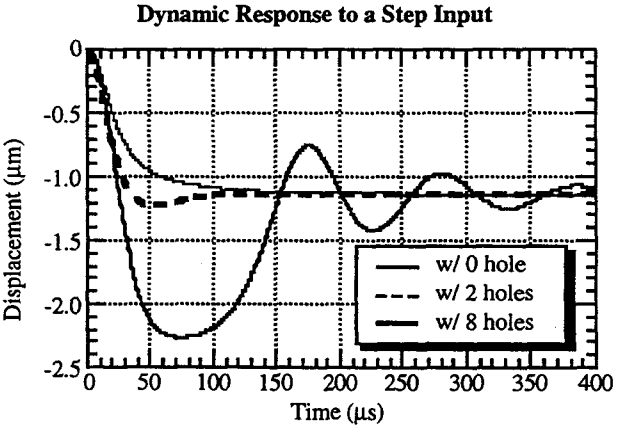


Figure 4. Effect of structural design in terms of the number of holes used in the cantilevered paddle structure on the paddle's dynamic response to a step voltage input.

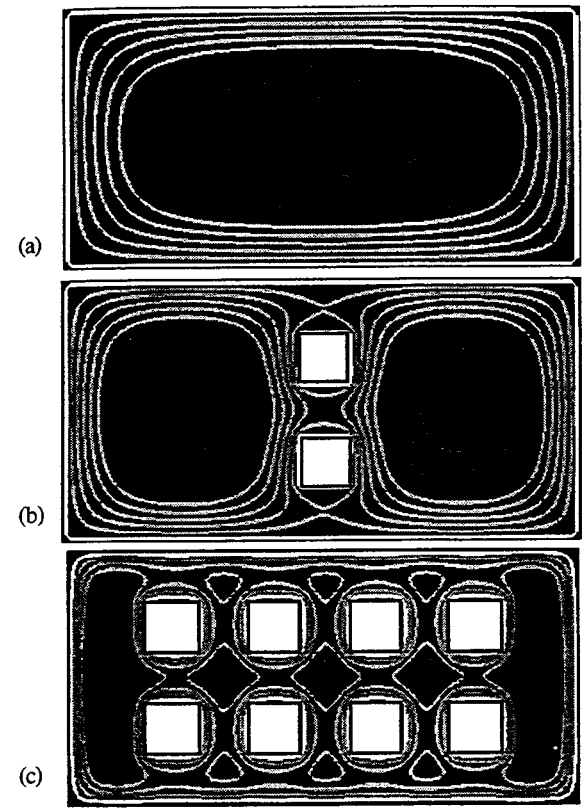


Figure 5. Contours of pressure loads at maximum transient damping force for designs with (a) no-hole, (b) 2-hole, and (c) 8-hole. The dynamic responses of these designs are in Fig. 4.

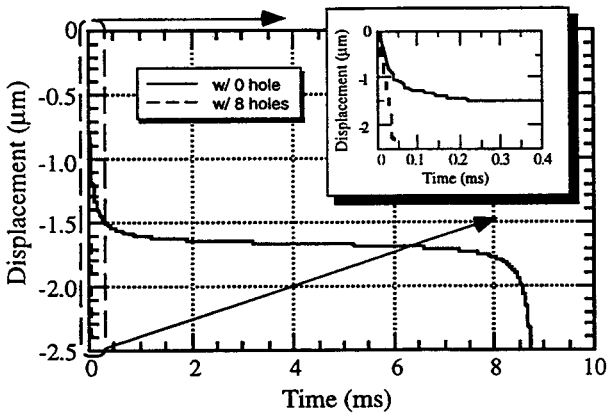


Figure 6. Simulation results showing the transient responses of the no-hole design and the 8-hole design with a step function of 9.124 V applied. The contact time has been reduced from 8.9 ms (in the no-hole design) to 67 μs (in the 8-hole design), a more than 100 times reduction.

CONCLUSION

Software for simulating the dynamics and the damping effects of MEMS squeeze type devices has been developed. The micro-fluidynamic damping effect is modeled to cover the slip flow by solving the Reynolds equation coupled with slip boundary condition at the structure surfaces. The analysis and the numerical procedure are presented. The software is very well-suited for modeling squeeze type devices with various holes, and it can be very effective for designing new MEMS squeeze type devices with desired damping characteristics.

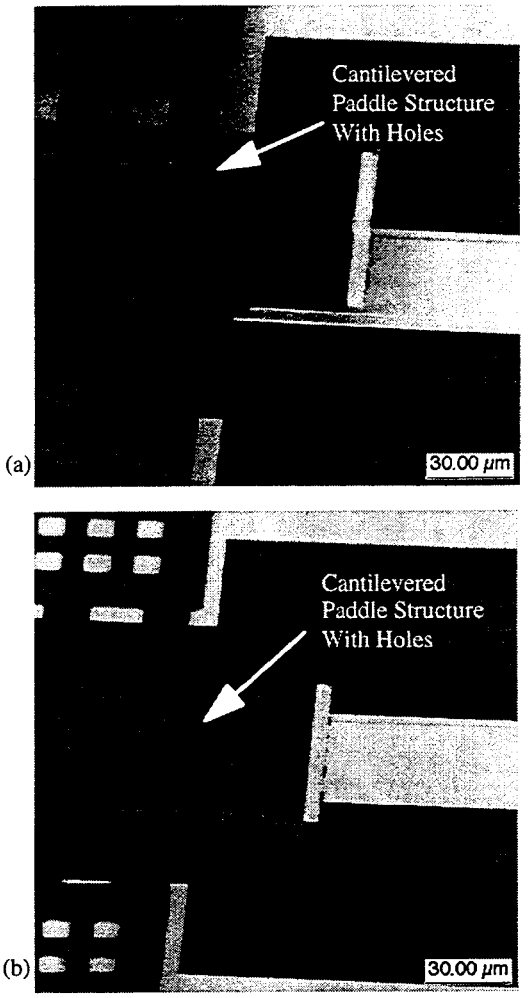


Figure 7. SEM back scattering images showing a close-up, tilted view of a test paddle structure with holes in (a) Up and (b) Down positions relative to the substrate.

REFERENCES

1. A. Beskok and G. E. Karniadakis, "Simulation of Slip-Flows in Complex Micro-Geometries," ASME, DSC-Vol. 40 (1992), pp. 355-370.

2. E. S. Piekos and K. S. Breuer, "DSMC Modeling of Micromechanical Devices," AIAA-95-2089 (1995).

3. W. S. Griffin, H. H. Richardson, and S. Yamanami, "A study of Fluid Squeeze-Film Damping," Journal of Basic Engineering (1966), pp. 451-456.

4. L. Zhang, D. Cho, H. Shiraishi, and W. Trimmer, "Squeeze Film Damping in Microelectromechanical Systems," ASME, DSC- Vol. 40 (1992), pp. 149-160.

5. S. A. Schaaf and P. L. Chambre, "Flow of Rarefied Gases," Section H in *Fundamentals of Gas Dynamics*, H. W. Emmons, ed., Princeton University Press (1958), p. 718.

6. J. J. Yao and M. F. Chang, "A Surface Micromachined Miniature Switch for Telecommunications Applications With Signal Frequencies From DC Up To 4 GHz," in Tech. Digest, The 8th International Conference on Solid-State Sensors and Actuators, Transducers-95, June 25-29, 1995, Stockholm, Sweden, pp. 384-387.

NUMERICAL SIMULATION OF COMPRESSIBLE SQUEEZED-FILM DAMPING

Yao-Joe Yang and Stephen D. Senturia
Microsystems Technology Laboratory, Massachusetts Institute of Technology
Cambridge, MA 02139

ABSTRACT

This paper presents the simulation of isothermal compressible squeezed-film damping effects for MEMS devices. The squeezed-film governing equation, the isothermal Reynold's equation, is widely used in lubrication theory. Both damping and compressibility effects of the squeezed film are modeled numerically. Small and large amplitude motions are simulated by using finite-element and finite-difference codes respectively. Macro-model examples with this squeeze-film effect are also demonstrated. We find that the air compressibility effect is significant in high frequency operation for small amplitude oscillations, and that the nonlinearity of the governing equation causes a nonlinear pressure response for large amplitude motion.

INTRODUCTION

MEMS devices are often operated in ambient pressure, so that air functions as an important working fluid. An air film between two closely spaced plates moving in normal relative motion gives rise to a pressure distribution over the plate. The total pressure force, which opposes the motion of the plate, is known as squeeze-film damping. Fig. 1 shows the schematic view of squeeze-film damping. Theory [1,2] shows that the viscous damping effect of the air film dominates at low frequencies, but spring-like behavior takes over at higher frequencies. The experimental studies [3-7] are in general agreement with the theory, but most of these studies focused only on the behavior of small amplitude oscillations. In this paper, we demonstrate that these effects can be successfully modeled numerically under small-amplitude sinusoidal conditions by a finite element package, as well as under large amplitude motion by finite difference codes. A simulated macromodel including the squeeze-film damping effect for a micro-accelerometer [8] with small amplitude motions is presented. The frequency response shows that the spring effect of the squeeze film due to air compression shifts the resonance. Simulations of large amplitude sinusoidal motions show that the nonlinear effects of the pressure force are dramatic.

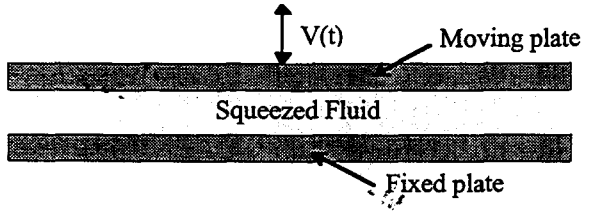


Fig. 1 Typical squeeze film geometry.

ISOTHERMAL REYNOLD'S EQUATION

Parallel rigid plates in relative normal motion experience a back-force from the time-dependent pressure distribution in the gas film. The product of the plate velocity and this back-force is the power transferred into the gas film by the plate motion, some of which appears as damping, the rest as stored energy in the compressed gas. The governing equation, the compressible isothermal Reynold's equation, is a time-dependent non-linear parabolic partial differential equation. The non-dimensional form of this equation for normal relative motion can be written as

$$\nabla\cdot(PH^{3}\nabla P)=\sigma^{*}\frac{\partial(PH)}{\partial t}$$

where

$$H=\frac{h}{h_0}\ ;\ P=\frac{p}{P_a}\ ;\ \sigma^{*}=\frac{12\mu L^2}{P_a h_0^2}$$

h is gap thickness, h_0 is the initial gap thickness, p is pressure, P_a is ambient pressure, L is plate length, μ is viscosity, and σ^* is a characteristic time [1].

The Reynold's equation is obtained from the Navier-Stokes equation, the continuity equation and the equation of state by assuming that (1) the inertia effect in the Navier-Stokes equation is negligible, (2) the pressure distribution across the gap is uniform and (3) the fluid velocity component perpendicular to plate surface is negligible.

LINEARIZED MODEL

Since the isothermal Reynold's equation is a highly non-linear partial differential equation, in general it has to be solved numerically. However, this equation can be linearized into a diffusion-like partial differential equation assuming (1) small gap thickness, (2) low Reynold's number, (3) isothermal conditions, (4) small motion amplitude, and (5) small pressure variation. The non-dimensional form of the linearized Reynold's Equation is

$$\nabla^2\Theta-\sigma\frac{\partial\Theta}{\partial\tau}=\sigma\frac{\partial e}{\partial\tau}$$

where

$$\tau=\omega t\ ;\ e=\frac{\delta}{h}\cos\omega t$$

$$\Theta=\frac{\Delta p}{P_a}\ ;\ \sigma=\frac{12\mu L^2\omega}{P_a h_0^2}$$

0-9640024-1-8 $3.00 © 1996 TRF

ω is oscillation frequency, τ is non-dimensional time, δ is amplitude of oscillation, Δp is small variation of pressure, Θ is the linearized non-dimensional pressure, and σ is the *squeeze number* [1].

The key parameter in the linearized Reynold's Equation is the squeeze number σ, which measures the compression of the fluid in the gap. If σ is close to 0 (low speed or frequency), the air film obeys nearly incompressible viscous flow; if σ goes to infinity (at very high speed or oscillation frequency), the fluid is essentially trapped in the gap and behaves like a spring. The linearized analytical solution for sinusoidal motion is available for rectangular and circular plates [2]. Furthermore, since the linearized Reynold's equation is in the same form of the transient energy-flow equation, we can use the finite element fluids-modeling package FIDAP [9] to generate the film pressure distribution for complex plate geometries by mapping the temperature variable in the energy equation to the pressure variable in the linearized Reynold's equation. Fig 2 shows analytical and numerical results of the non-dimensional pressure forces vs. squeeze number for an air film between two square plates with one plate in small-amplitude sinusoidal motion. The sinusoidal back-force on the plate has two components: one is in phase with plate velocity (the viscous damping force), the other is in phase with plate displacement (the spring force due to the compressibility of air). At small squeeze number (low frequency), the viscous force dominates because the air can squeeze out of the gap without being compressed. At large squeeze number (high frequency), the spring force increases because of compression of the air film. The numerical results for the spring force component agree with the analytical results, but the damping force component is about 8% lower than the analytical results at higher squeeze number. The reason for this small discrepancy is not yet understood, which places a limit of confidence of about 10% on the numerical results which follow. The normalized transient response for the total back-force is shown in Fig. 3. The forces have been normalized to their steady state maximum values. The force lags behind the sinusoidal plate velocity, and the phase lag increases with squeeze number because of the air compressibility effect. Seidel observed this lag experimentally, but chose to model it with a fixed delay instead of as a compression effect [4].

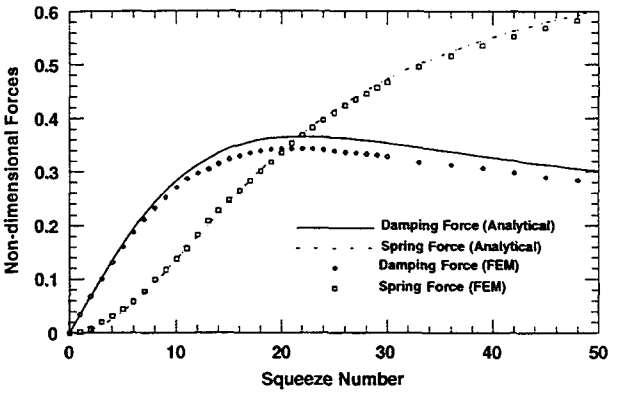


Fig. 2 Analytical and numerical results of non-dimensional damping and spring forces vs. squeeze number for a square plate.

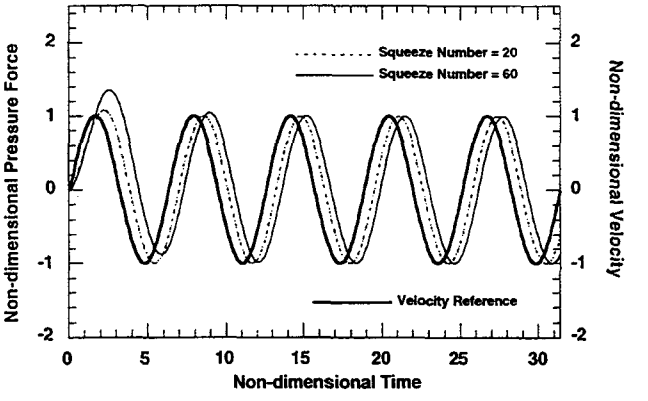


Fig. 3 Transient responses of the total back-force for two different squeeze numbers (20 and 60).

Fig. 4 is the schematic of a micro-accelerometer fabricated by the MIT MEMS group [10]. The proof mass is a 500 mm square plate supported 1 μm above the substrate by flexible tethers. The simulated small-amplitude dynamical responses of the accelerometer are illustrated in Fig. 5. Curve 1 is a typical response without any air damping. Curve 2, which drops very fast, is the response with the air damping. In order to reduce damping, a perforated proof mass can be used. Fig. 6 is a contour plot of the transient pressure distribution for a perforated mass; Curve 3 of

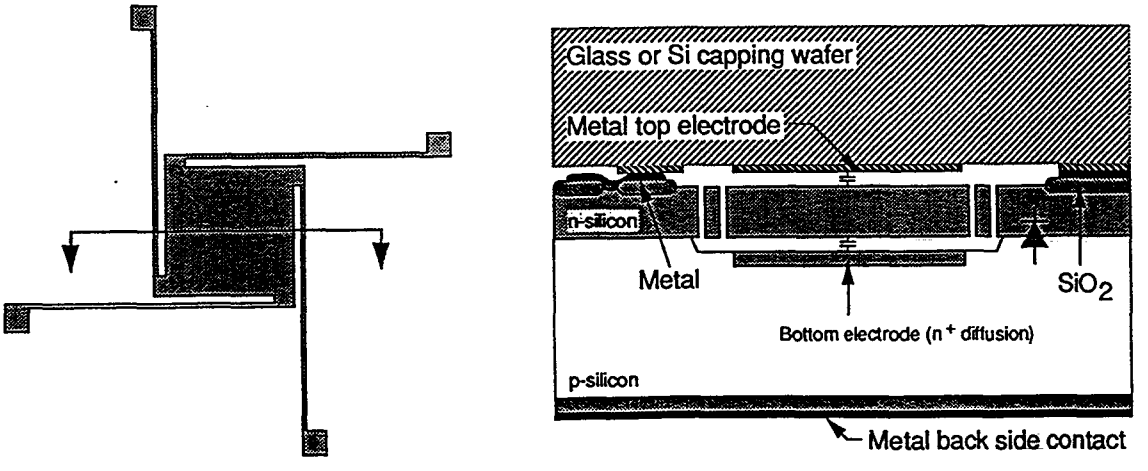


Fig. 4 Schematic view of a micro-accelerometer fabricated by the MIT MEMS group [10].

Fig. 5 is the corresponding response. Curves 2 and 3 show that the spring effect of the gas film results in a resonance shift. At low frequencies (low squeeze number), the system is overdamped because of highly viscous damping. At higher frequencies, the viscous damping decreases while the spring force increases (see Fig. 2). The resonance occurs at a high enough frequencies to make the resonance itself be underdamped.

Note that in many cases the gap thickness is too small to consider the air film as a continuum because of slip-flow condition at the boundaries. As a result, the damping effect is reduced. The viscosity can be modified to a smaller number to account for the slip-conditions at least to first order [5,7,11].

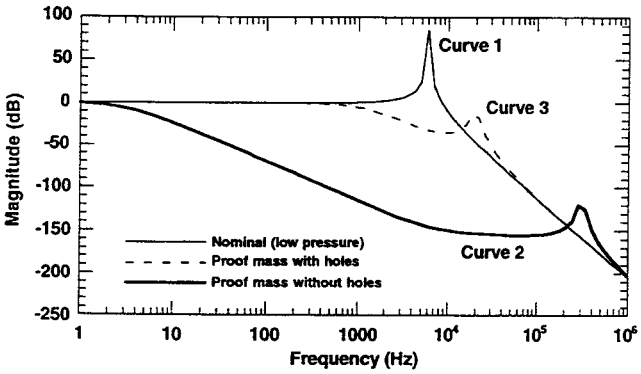


Fig. 5 Simulated dynamical responses of the micro-accelerometer with and without air film damping effects.

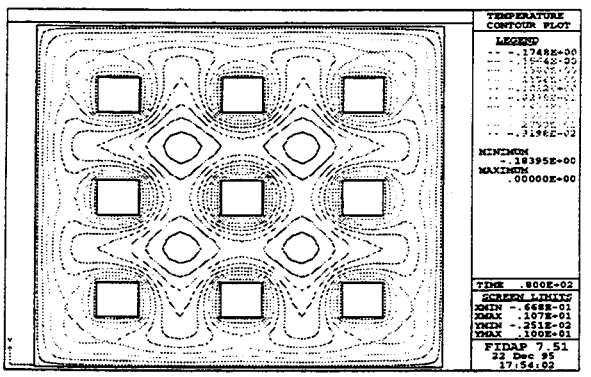


Fig. 6 Contour plot of the transient pressure distribution on a perforated proof mass of the accelerometer.

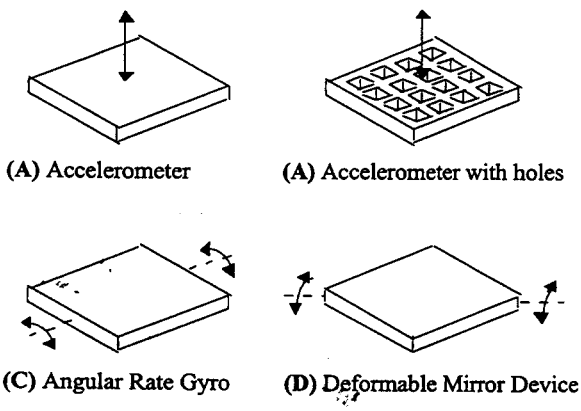


Fig. 7 Four examples of plate motion related to squeeze-film damping.

LARGE AMPLITUDE MOTION

In MEMS devices, the motion of the plate is not always a small amplitude oscillation. For example, the proof mass of the accelerometer shown in Fig. 4 will move a significant distance when undergoing a sudden change of acceleration. Therefore, the small amplitude oscillation model will not be valid in such a circumstance. Furthermore, some devices such as an angular rate gyro [12] use torsional motion, which is also significantly affected by squeezed-film damping, but which cannot be modeled using small amplitude motion. In order to deal with those problems, we have used the finite difference method to solve the isothermal Reynold's equation numerically. Fig. 7 shows four cases of squeeze-film damping problems with either large amplitude motions or non-parallel relative normal motions. The 3D surface and contour plots of pressure distributions for case (A) are shown in Fig. 8. These plots are the results at the time in the cycle at which the pressure reaches its maximum value. Fig. 9 is the dimensionless time response of the total back force for case (A), obtained by integrating the pressure on the plate over the plate area. The oscillation amplitude in this case is equal to 30 % of the gap thickness. Because of the nonlinearity of Reynold's equation, the back force curve, which is normalized to the ambient pressure force (i.e., ambient pressure times the plate area), is not symmetric about the ambient pressure force (P=1 in the figure). When the plate moves down, the maximum back-force is about 1.2 times of the ambient pressure force; when plate moves up, the minimum back-force is about 0.9 of the ambient pressure force. Therefore

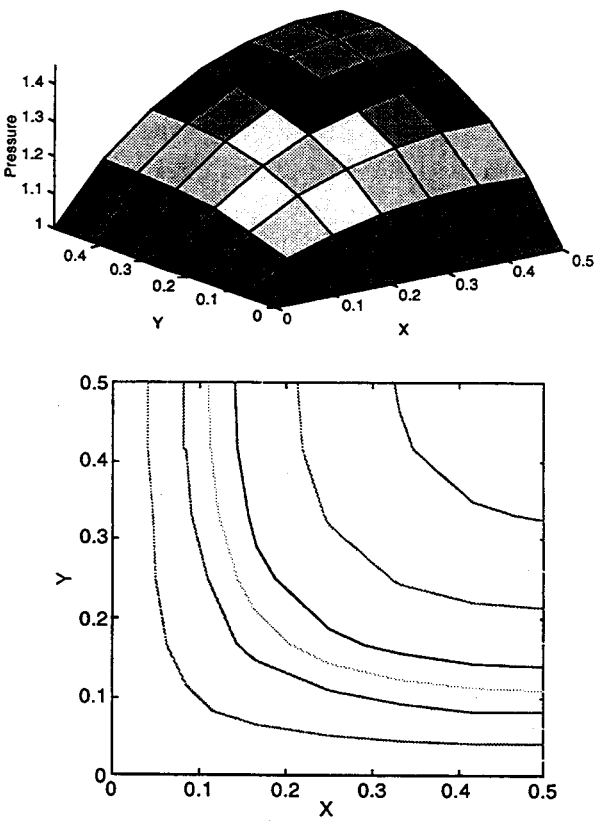


Fig. 8 3D-surface and contour plots of pressure distribution for one quarter of a square plate with large amplitude motion.

the plate undergoes a larger pressure resistant force when the plate moves down. Note that the velocity curve in Fig. 9 has been normalized. The figure also indicates that the transfer function between force and velocity is clearly nonlinear.

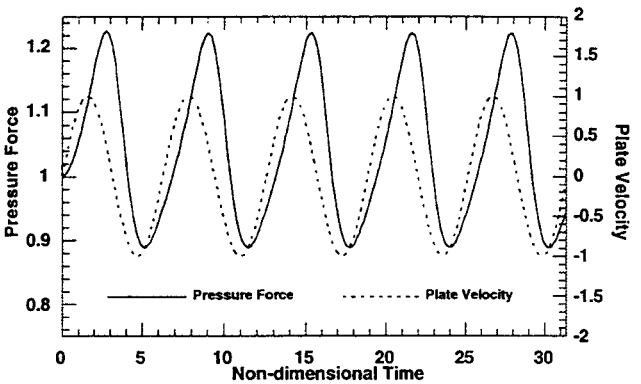


Fig. 9 *Nonlinear time response of the total pressure back force for a square plate with a large amplitude oscillation.*

Table 1 summarizes the maximum and minimum pressures on the plates for the four cases in Fig. 7 when the plates are sinusoidally oscillated in normal relative motion or torsional motion. Note that the pressure caused by torsional motion, which is also a potential second mode of motion for the accelerometer shown in Fig. 4, is much less than that caused by normal relative motion (case (A) in Fig. (7)) under ambient pressure. The design with the perforated plate also reduces the pressure significantly.

CASE	Max. d/h0	Max. P/Pa	Min. P/Pa
(A) Accel.	0.3	1.4484	0.7688
(B) Perforated Accel.	0.3	1.1007	0.9079
(C) Gyro	0.3	1.0902	0.9246
(D) DMD	0.3	1.0635	0.9446

Table 1 *Summary of maximum and minimum pressure on the plates for the four cases in Fig. 7. Note that* **d** *is the plate oscillation amplitude,* **h0** *is the initial gap thickness,* **P** *is the air film pressure on the plate, and* **Pa** *is the ambient pressure. The oscillation amplitude for case (C) and (D) is the maximum displacement of the plate tip.*

SUMMARY AND CONCLUSION

The numerical simulations of squeeze-film damping effects, based on the isothermal Reynold's equation, have been explored. The linearized model can be simulated by the finite element package FIDAP under small amplitude condition for plate geometries too complex for analytical solution. The viscous damping and spring effects of squeeze-films can be decoupled and modeled as frequency-dependent parameters for the small amplitude frequency response analysis of MEMS devices. An example, the dynamics of an accelerometer, is demonstrated. For the device operated in ambient pressure, the viscous damping effect dominates at low frequencies. The spring effect of the squeezed air film becomes significant at high frequency, and shifts the resonance by a large amount. The squeeze-film effect for large amplitude motions is simulated by finite difference codes. Four examples of different types of motions are demonstrated. The nonlinearity of the governing equation gives rise to a nonlinear pressure back force on the plate. The maximum pressure back forces for each case under specific motion amplitude are summarized.

ACKNOWLEDGMENTS

This work was supported in part by Semiconductor Research Corporation and by ARPA. The authors wish to thank Prof. K. Breuer for helpful discussions, and Charles Hsu for providing the accelerometer example.

REFERENCES

[1] W. E. Langlois, "Isothermal Squeeze Films," Quarterly 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, pp. 615-620.

[3] W. E. Newell, "The Miniaturization of Tuning Forks," Science, Vol. 161, 1968, pp. 1320-1326.

[4] H. Seidel, H. Riedel, R. Kolbeck, G. Mück, W. Kupke and M. Königer, "Capacitive Silicon Accelerometer with Highly Symmetrical Design," Sensors and Actuators, Vol. A21-23, 1990, pp. 312-315.

[5] M. Andrews, I. Harris and G. Turner, "A Comparison of Squeeze-film Theory with Measurements on A Microstructure," Sensors and Actuators A , Vol. A36, No. 2, March 1993, pp. 79-87.

[6] J. B. Starr, "Squeeze-film Damping in Solid-State Accelerometers," Tech. Digest, IEEE Solid State Sensor and Actuator Workshop, Hilton Head Island, SC, June 1990, pp. 44-47.

[7] T. Veijola, H. Kuisma, J. Lahdenperä and T. Ryhänen, "Equivalent-circuit model of the squeezed gas film in a silicon accelerometer," Sensors and Actuators A, Vol. A48, 1995, pp. 239-248.

[8] M. Novack, "Design and Fabrication of a Thin-film Micromachined Accelerometer," M.S. Thesis, Massachusetts Institute of Technology, 1992.

[9] FIDAP 7.0 manual, Fluid Dynamics International, Inc., 1993.

[10] C. Hsu and M. Schmidt, private communication.

[11] S. Dushman and J. M. Lafferty, Scientific Foundations of Vacuum Technique, Wiley, New York, 1962.

[12] J. Mizuno, K. Nottmeyer, C. Cabuz, T. Kobayashi, and M. Esashi, "Fabrication and Characterization of a Silicon Capacitive Structure for Simultaneous Detection of Acceleration and Angular Rate," Transducers '95 and Eurosensors IX, Stockholm, Sweden, June 1995, Vol. 2, pp. 679-682.

PROGRESS IN TUNNEL SENSORS

J. Grade, A. Barzilai, J.K. Reynolds, C.H. Liu, A. Partridge, and T.W. Kenny
Department of Mechanical Engineering, Terman 551
Stanford University, Stanford, CA 94305-4021

T.R. VanZandt, L.M. Miller, and J.A. Podosek
Jet Propulsion Laboratory, Pasadena, CA 91109

ABSTRACT

For several years, research has been underway on the use of electron tunneling as a displacement transducer in microsensors. Several prototype devices have been developed which illustrate the potential advantages of tunneling transducers. In this report, we focus on recent measurements of noise and drift which indicate that tunneling transducers can perform very well over long periods. In addition, we report on recent tunnel sensor development efforts, including a MCNC-based test device, an integrated controller for tunneling transducers, and a high-performance tunneling seismometer based on a modified geophone.

INTRODUCTION

Silicon micromachining has enabled the miniaturization of nearly every displacement transducer-based sensor. In many cases, the miniaturization of these devices has come with a reduction in sensitivity. The reduced sensitivity is due to two effects. The first effect is a reduction in the coupling of the external signal to the displacement of a sensor component. A miniaturized pressure sensor, for example, undergoes a smaller deflection because the area of the diaphragm has been reduced. The second effect is a reduction in the performance of the displacement transducer with the size of the device. A capacitive pressure sensor, for example, becomes less sensitive as the area of the capacitive electrodes is reduced while miniaturizing the device.

Electron tunneling has been studied as an important displacement transducer for microsensors because of its high sensitivity and scale independence. In tunneling, a nA current is measured, and 1% fluctuations in this current correspond to displacements of less than 0.1Å. This sensitivity is independent of the lateral dimensions of the electrodes, since tunneling only requires 1 metal atom on the surface of each side of the gap.

In the last 2 years alone, published work on tunneling sensors has included an infrared detector[1], several accelerometers[2-4], a pressure sensor[5], and a magnetometer[6]. All of this work serves to illustrate the breadth of applications for the tunneling transducer and the variety of possible fabrication techniques which may be used to build them. A complete description of the issues to be considered during design, fabrication, and operation of micromachined tunneling transducers has been published.[7]

Despite this growing research effort, there is still broad concern about the utility of tunneling sensors for *real* applications. Generally, concerns are expressed about the fabrication processes necessary to produce appropriate characteristics, the stability of tunneling transducers, and the packaging requirements necessary to preserve sensor performance. In this report, we will review recent work directed towards these issues, and describe several development efforts which serve to illustrate the simplicity, applicability and performance of tunneling transducers.

LOW-FREQUENCY NOISE MEASUREMENTS

The properties of tunneling transducers are known to rely on the location of atoms on a pair of electrodes. If the atoms re-arrange themselves, the current between the electrodes will fluctuate, leading to errors in the measurement of signals applied to the device. This dependence on the position of individual atoms is unique to tunneling; capacitive transducers are expected to be more stable because they are only sensitive to the average separation between extremely large numbers of atoms. Gold surfaces are favored for tunneling electrodes because gold resists chemical reaction with the constituents of most potential operating environments. Unfortunately, gold is a soft metal, and gold surface atoms are known to re-arrange at regular intervals at room temperature. Arrangements of atoms which feature extremely sharp tip-like structures can be expected to be particularly unstable because energy considerations favor flat (111) crystal surfaces.

In addition to gold atom rearrangements, errors can be expected due to the migration of physically-adsorbed molecules, such as H_2O, through the tunneling region. Other possible sources of error include variations in the mechanical characteristics of the structure supporting the electrodes, and variations in the electrical properties of the feedback circuit elements.

Since we are interested in the fundamental performance of tunneling transducers, we need to consider how these various sources of noise can contribute to significant errors in position. For example, the migration of adsorbates through the tunneling region should only be important at high frequencies because this mechanism cannot produce an accumulation of error. Changes in the structure due to thermal expansion should only be important at low frequency, because the temperature cannot change quickly. To understand the phenomena which limit the performance of tunneling sensors, we have focused on noise measurements.

An example of a micromachined device which uses tunneling currents to detect small forces is the membrane transducer, shown in Fig. 1. This device operates by electrostatically deflecting the membrane down towards the tip. When a tunneling current of order 1 nA appears between the tip and the membrane, a feedback voltage is added to the voltage on the deflection electrode so as to maintain the current at a constant value. If an external force is applied to the membrane, the feedback circuit responds by applying a balancing electrostatic force. By monitoring the feedback control voltage, it is possible to determine the external forces.

0-9640024-1-8 $3.00 © 1996 TRF

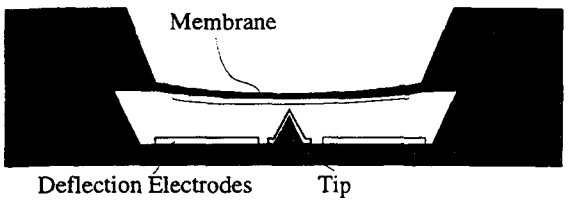


Fig. 1 This drawing shows the design of the membrane transducer. In this design, the gap between deflection electrode and membrane is about 40 μm, and the tip is about 40 μm tall. The membrane is 2 mm x 2 mm x 0.5 μm .

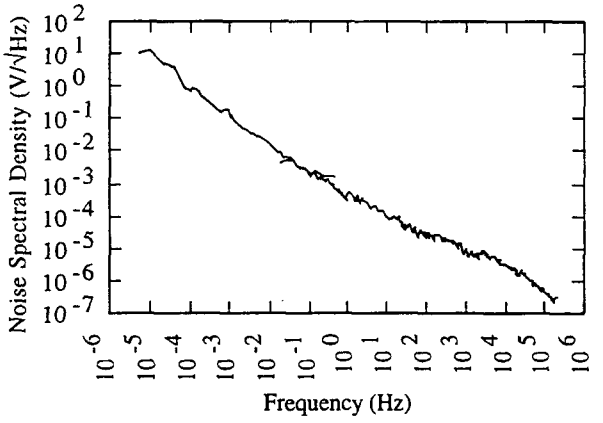


Fig. 2 This graph shows the measured noise spectral density for membrane tunneling transducers over a very broad range of frequencies.

A complete fabrication process and experimental characterization of this membrane transducer has been presented elsewhere. [7] In summary, this device features a control bandwidth in excess of 50 kHz, a force resolution of 10^{-11} N/√Hz, and a very simple fabrication and assembly technique. Devices of this type have been operated continuously for more than 3 years without failure. This device is the basis of the successful tunneling infrared sensor and magnetometer, and has been studied extensively at JPL and elsewhere.

All of the tunnel sensor development projects carried out in the JPL and Stanford groups have been constrained to applications which are within the known performance capabilities of tunneling transducers. One component of this constraint has been the known noise spectra of tunneling transducers. Figure 2 shows a measurement of the noise spectral density of the control voltage for membrane-style sensors over a very broad frequency range. This measurement has been repeated many times for many sensors in several locations, and is characterized by a 1/√F shape. The sources for this noise were not identified, but its consistency persuaded us to avoid applications which would require improvement in this noise performance. The displacement sensitivity of these devices is generally around 10 mV/Å, so the resolution at 1 Hz is 0.02Å/√Hz.

Several months ago, our group began logging data on the long-term drift in tunneling transducers. The purpose of those measurements was to look for an upper limit to the 1/√F noise characteristic, because there are practical limits to the range of

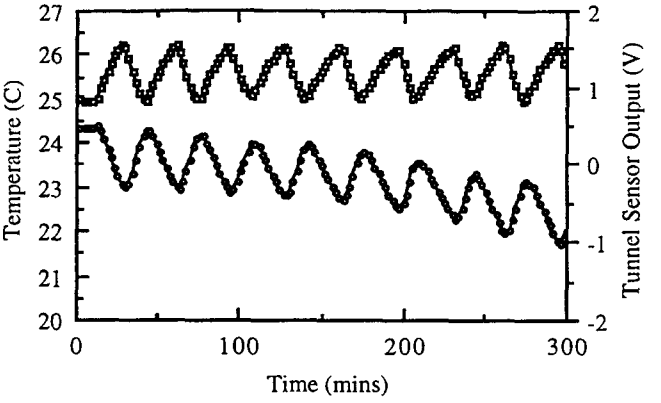


Fig. 3 This graph shows the measured output signal for a membrane transducer (lower curve), along with a measurement of the temperature near the sensor in the laboratory (upper curve)

error possible in a device of this sort. Figure 3 shows the result of an early measurement of the output signal from a membrane transducer operated at Stanford.

This initial low-frequency drift measurement is characterized by a very low-frequency oscillation. After some investigation, the source of this oscillation was identified : the laboratory is air-conditioned by a central air system which features a limit-cycle oscillation with a period of about 30 min, depending on the heat load applied to the building. This hypothesis was confirmed by a repeat experiment which included a thermometer whose output was also logged. The result of this experiment is shown in Figure 3 along with the sensor output oscillation. Clearly, there is a strong correlation between the temperature and the tunnel sensor output.

Figure 4 is a graph of the coherence between the tunnel sensor noise spectrum and the thermometer noise spectrum. It is clear that, for frequencies below 0.1 Hz, these signals are very well correlated (correlation > 95%). There is one very strong conclusion which may be drawn from this measurement : More than 95% of the noise in tunneling transducers is due to temperature coefficients of the mechanical properties of the sensor or the package. If these thermal sensitivities could be eliminated, the low-frequency noise of these sensors would improve by a significant factor.

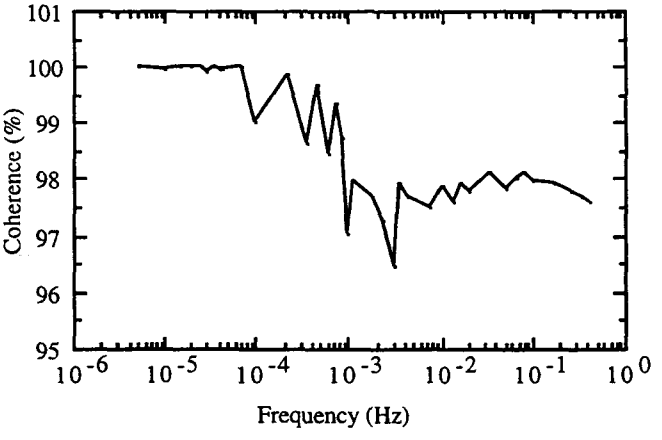


Fig. 4. This graph shows the measured coherence between the output of the membrane sensor and the thermometer.

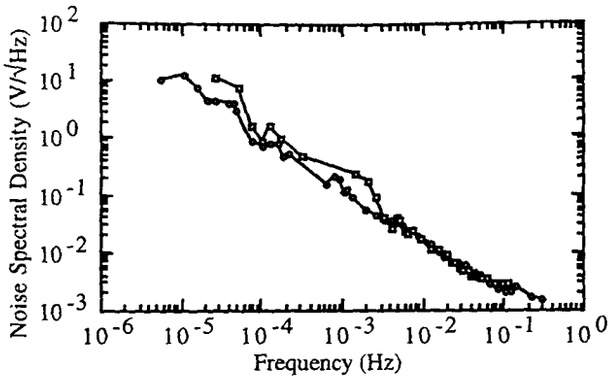


Fig. 5. This graph shows the measured Noise Spectral Density for the output signals of a sensor operated with metal spring clamps (circles) and with epoxy (squares) holding the structure together. It is clear that the replacement of clamps with epoxy has not had an important effect on the noise in this sensor. (The peaks due to air conditioner oscillation have been removed to increase clarity.)

We initially expected that this thermal sensitivity was related to the epoxy used to bond the halves of a tunneling transducer together. To test this hypothesis, we measured the noise spectrum of a tunnel sensor when its halves were clamped together with folded metal wire springs, and then when these springs were replaced with epoxy. The result is shown in Fig. 5, which indicates that the low-frequency noise in the tunneling device is not changed when the metal springs are replaced with quick-cure epoxy. This result can be explained in one of two ways : either the temperature coefficient of the stiffness of the springs and the epoxy are exactly the same, or the thermal sensitivity of the membrane sensor is dominated by some other element of the structure. We believe the latter hypothesis, and are planning a series of experiments to test this.

Figure 6 shows a plot of the noise in a membrane transducer, and a plot of the noise in the transducer which is uncorrelated with the output of our thermometer. We can clearly say that the lower curve represents the new upper bound to the fundamental noise in tunneling transducers.

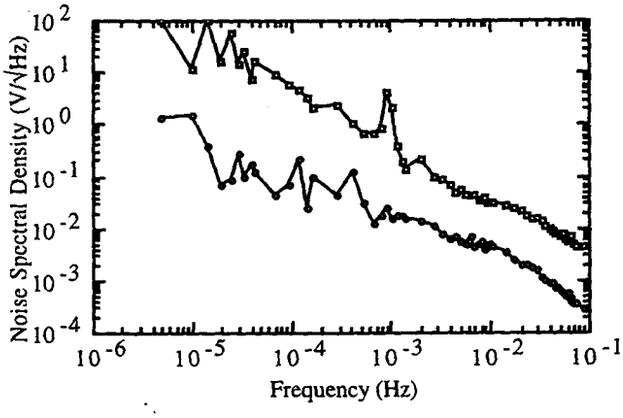


Fig. 6. This graph shows the measured noise of a typical membrane transducer(upper) and the noise which is uncorrelated with temperature measurements(lower). The lower curve represents a new upper limit to the fundamental noise in tunneling transducers of this type.

This series of measurements indicates that temperature fluctuations in a normal laboratory environment can dominate the low-frequency noise in simple tunneling structures. For our laboratory, more than 95% of the noise at all frequencies below 0.1 Hz is due to environmental temperature changes. Therefore, a capacitive transducer built with this mechanical structure could perform no better. It is important to note that these measurements may be limited by the accuracy of the diode thermometer that was used, and that an even greater fraction of the noise may be due to this source. The way to improve this situation is clearly by changing the mechanical design of this sensor structure. This work is currently underway.

Also, this data shows that, despite the temperature sensitivity of these membrane devices, the errors are remarkably small. These measurements allow us to be quantitative about this result. The drift in the output of a tunneling sensor may be calculated as follows :

$$\Delta V_{RMS} = \sqrt{\int_{F1}^{F2} (NSD(F))^2 dF} \quad , \qquad (1)$$

where NSD(F) is a fit to the noise spectral density (Fig. 2) as a function of frequency between the two frequencies which represent the measurement bandwidth. Across the frequency range shown in Fig. 2, this function is reasonably well-approximated by

$$NSD(F) = \frac{0.5mV/\sqrt{Hz}}{\sqrt{F/1Hz}} \quad . \qquad (2)$$

Suppose we were to average the output of a membrane sensor for 1 second, and repeat the measurement an hour later. The RMS average difference between the results would be approximated by taking the above integral from F1 = 0.0003 Hz to F2 = 1 Hz. The result of this calculation is a change in output voltage of 1.5 mV, which corresponds to an average error in position of only 0.15 Å. This is considerably smaller than the lattice spacing between atoms in Gold, and indicates that atomic motion of the atoms involved in the tunneling process is extremely rare. Clearly, this level of stability could not be achieved if the control system was unable to prevent "crashes" between the electrodes. Remember that these membrane transducers operate with extremely wide control bandwidth (>50 kHz), thereby preventing "crashes". It might be unreasonable to expect this level of stability from any tunneling system with a smaller control bandwidth.

These measurements also indicate the importance and value in measurements of noise in tunneling sensors. Researchers who have successfully implemented any tunneling device and are seeking to publish the results are strongly encouraged to report not just the signal measurements, but also the noise measurements. Both measurements are necessary to determine the applicability of these devices, and both measurements should be considered essential in any conclusive device characterization.

A PROTOTYPE TUNNELING SEISMOMETER

Seismology and resource exploration is currently carried out with inexpensive vibration sensors called geophones.[8] Geophones operate by detecting the velocity of a spring-supported coil through voltages generated as it moves through a magnetic field. Typical geophones have a suspended mass of 20 gm, a resonant frequency near 5 Hz, total dimensions of order 3 cm

diameter x 5 cm tall, a mass of 100-200 gm, and a cost of under $50. The suspension within the geophone uses circular BeCu leaf springs and is significantly stiffer in the cross-axis directions than along the sensitive axis. Near their resonant frequency, geophones can detect acceleration as small as 10^{-9}g when operated with a 10 Hz bandwidth. As such, the geophone represents a very mature, high performance technology. It would be very difficult to micromachine such a device, because the low resonant frequency would require a very compliant flexural support and a large suspended mass. Significant reduction of the mass would degrade the performance of the sensor because thermal noise increases as $1/\sqrt{m}$, and would be greater than 10^{-9} g/$\sqrt{Hz}$ for m < 2 gm.[9]

Because the transducer measures coil velocity, at frequencies away from the resonance, the sensitivity degrades linearly with frequency. This poses a problem for seismology because many interesting signals occur at frequencies well below 1 Hz, where geophone sensitivity is poor. One method for improving the low-frequency performance is to introduce a position transducer, and use the magnet/coil system to generate a rebalance force to hold the coil position constant. To resolve acceleration signals near 10^{-9}g at frequencies below 5 Hz, the displacement transducer needs to resolve deflections of :

$$\Delta Z = \frac{A}{\omega_o^2} = \frac{10^{-8}\ m/s^2}{(2\pi\ 5Hz)^2} = 0.1Å. \qquad (3)$$

This is within the performance capabilities of the tunneling transducer, according to the data presented in Fig. 2 above. The tunneling transducer includes the advantage that the amplification and feedback circuitry can be implemented using discrete circuit elements, and does not need to be highly integrated in order to operate with the necessary performance.

Figure 7 shows a cross-section of a geophone which has been modified to include a tunneling position transducer. One electrode of the tunneling transducer is supported by the geophone case, and the other is attached to the moving coil. The position of the moving coil is controlled by passing a servo current through the coil. When there is an external acceleration, the servo force will balance the inertial force, thereby keeping the relative position of the coil constant.

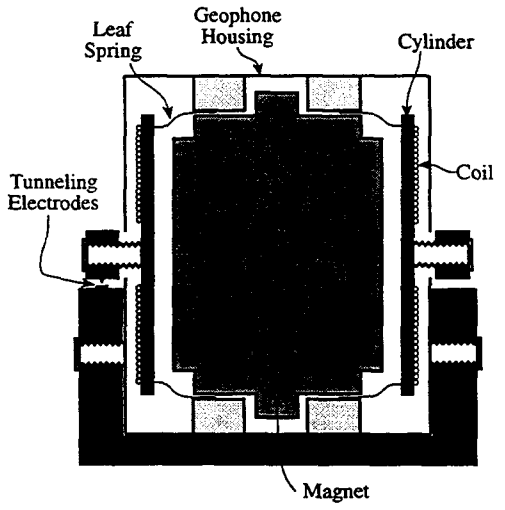


Figure 7. This drawing shows the internal structure of a geophone, and the modifications made to attach tunneling electrodes to the device.

As described elsewhere [7] there is no need for sharpened tips in a tunneling transducer. In this device, we have experimented with several geometries, including tunneling between the edges of a crossed pair of wires, between the end of a wire and a flat electrode, or between the side of a bent wire and a flat metal electrode. We have focused on these rather crude architectures because the low resonant frequency of the geophone will prevent the implementation of wide-bandwidth control.

During operation, motion of the tunnel tip leads to a change in tunnel current, which is converted to a voltage through a 10 MΩ resistor. An electronic control circuit (to be designed in the near future) generates a control voltage which is applied to the suspended coil in the geophone. If the control circuit is working properly, the inertial force on the coil (F = mA = 0.2 N/g for a 20 gm mass) will be balanced by the electromagnetic force on the coil (F = 0.02 N/V), so the sensitivity will be given by :

$$S = \frac{\partial V}{\partial F}\frac{\partial F}{\partial A} = (50\ V/N)(0.2N/g) = 10\ V/g. \qquad (4)$$

We have begun testing the prototype tunneling seismometer. We have encountered some difficulty implementing a controller which actually maintains a steady tunneling current because of the low resonant frequency of the proof mass One version of the control circuit operates with a large oscillation above 300 Hz, in which the tunneling electrodes are alternately too far apart or in contact. Low frequency acceleration signals cause the controller to adjust the duty-cycle of this oscillation. By filtering the output, it is possible to detect low-frequency oscillations with this system. In a preliminary experiment, we confirmed the sensitivity of 10V/g, and recorded the noise spectrum shown in Fig. 8. For frequencies below 15 Hz, the resolution of this "almost"-tunneling geophone is already near 1 μg/$\sqrt{Hz}$, which is an excellent early result, and is competitive with good lab-grade accelerometers. With stable feedback control of the tunneling current, the fraction of the time that the sensor is tunneling can be increased dramatically, and we expect to achieve resolution approaching 10^{-9} g/$\sqrt{Hz}$. If successful, we plan to bring an operational tunneling geophone to our poster at the conference.

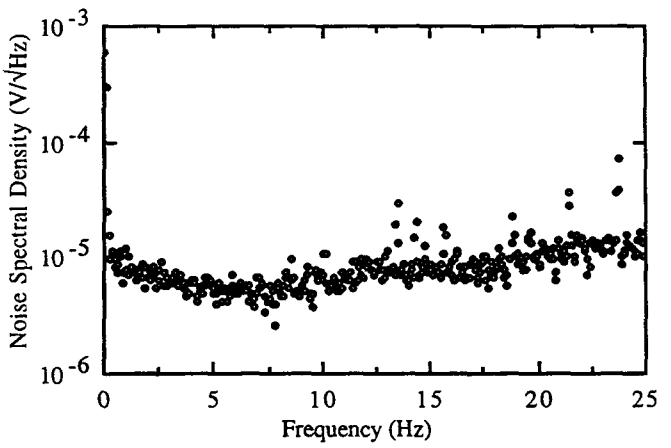


Figure 8 This graph shows the noise spectral density recorded at the output of the tunneling geophone. Since this device has a sensitivity of 10V/g, the resolution is near 1 μg/√Hz across much of this low frequency range.

SURFACE MICROMACHINED TUNNELING TRANSDUCER

To understand the sources of noise in tunneling devices, we have been interested in the development of test devices with structures other than the membrane device. One interesting approach is based on the work of Kobayashi and Fujita,[10] in which surface micromachining has been used to build a "Lateral Tunneling Unit". In their device, an electrostatic comb-drive is used to deflect a metal-coated polysilicon "finger" towards a fixed, metal-coated electrode.

We have used the MCNC MUMPS process for surface micromachined polysilicon structures, and have evaporated Ti/Pt/Au metallization onto the structures after the HF-release step. This metal layer recipe provides good adhesion to the surface, an excellent diffusion barrier to prevent silicon diffusion through the gold, and a clean gold surface for the tunneling contacts. By evaporating after release, we rely on the overhanging polysilicon of each structural layer to provide shadow-mask isolation between the electrodes. Using this process, we are able to achieve the necessary clean metal electrodes, electrical isolation, and mechanical control that is necessary for the operation of a tunneling device. Figure 9 shows a drawing of the device design.

Difficulties encountered during early testing of this device have included stress in the evaporated metal layer which causes out-of-plane warp in the released structures. We recommend use of both structural polysilicon layers in the MUMPS process to stiffen the released structures, and careful minimization of the metal layer thickness. In our metal process, good metal electrodes can be achieved with as little as 80Å/80Å/800Å of the Ti/Pt/Au, respectively.

Preliminary testing of these devices is underway. Limitation of the out-of plane warp through the use of the 2-layer poly and minimization of the metal layer thickness has produced a device which acts as a reliable relay. Implementation of feedback control is underway, and we will present recommendations for feedback control in a subsequent publication. Once successful, a complete description of this device and its performance characteristics will be published as well. A significant advantage of fabrication of a test tunnel sensor structure through the MCNC process is that it will enable many new researchers to begin experimentation with tunneling devices.

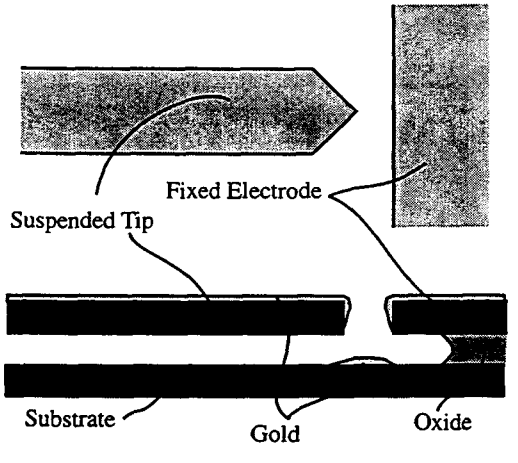


Figure 9 This drawing shows the method for obtaining tunneling electrodes from a comb drive-supported tip and a fixed electrode. The overhang which results from the release etch forms a natural shadow mask. The metals are evaporated with the sample mounted on a tilted planetary mount, so that the facing edges of the tunneling contact are partially coated. (After Kobayashi, et. al. [10]).

INTEGRATED TUNNEL CONTROLLER IN STANDARD CMOS

In a previous publication, we have described a simple, discrete-component circuit which can be used to provide wide-bandwidth control for the membrane tunneling transducer. In this work, we report on work towards a single-chip control circuit for this device which can be made through MOSIS or the Stanford in-house CMOS process.

While tunneling servo controllers can be quite simple to build from discrete components, these embodiments cannot generally be directly integrated. The tunneling current must be sensed and converted to a voltage. In discrete circuits this is accomplished with a large-value resistor (10MΩ), which would use significant area in an integrated form. In addition, the common method to develop large resistances, well-implant snakes, would have prohibitively large leakage to substrate. Therefore we are developing a replacement for this component. We have constrained our search to implementations in common industrial CMOS, and assume that special purpose resistors and high quality bipolar devices are not available.

Many options are available for this application, including switched capacitor techniques, scaled current mirrors, and others. We have elected to convert the tunneling current to a voltage by dropping the current across a forward biased diode. Because of the logarithmic nature of the diode, this develops an easily measured voltage. Advantages of this approach include the small silicon area required by the diode and the high effective impedance of the diode at the small tunneling currents. At a current of a few nA, a diode will exhibit a bias of a few hundred millivolts, an equivalent (linearized) resistance on the order of a hundred megohms. This large equivalent resistance reduces the gain requirements of the amplifier stage. Also, the exponential dependence of tunneling current with tunneling gap displacement is linearized by the logarithmic diode, so displacement in the tunneling gap is roughly proportional to the developed voltage.

The forward current vs. voltage for a diode varies strongly from device to device and with temperature. Therefore, the control voltage must be generated and temperature compensated. The direct approach would be to drop a reference current across a reference diode identical to the sense diode and servo to the voltage across the sense diode to the voltage developed across the reference. While conceptually simple, developing the small well regulated reference current can be difficult in CMOS. Instead, we employ a method of scaling a voltage developed across matched diodes at larger currents.

The servo controller employs simple integral feedback with a bandwidth-limited op-amp topology. Since the electrostatic feedback of present tunneling transduces requires large drive voltages, we include a high voltage class A output stage built with a high-voltage-FET transistor [11]. This transistor employs a lightly doped drain structure built from standard well implant to provide a high breakdown voltage. It can be fabricated in standard

CMOS and provide over 100 V drive in our MOSIS parts and 60V drive in our in-house CMOS.

The components of the circuit have been developed and tested and the circuit has been fabricated and is now under test.

CONCLUSIONS

In the past 2 years the pace of tunnel sensor development has quickened. Our research activities have been focused on the determination of the fundamental characteristics of tunneling transducers, and on the development of techniques to support commercialization. Recent low-frequency noise measurements have proven that the drift in tunnel sensors can be low enough to be dominated by thermal-mechanical effects in the supporting structure, and that the fundamental stability of these transducers is remarkably good. This work illustrates the importance of noise measurements for device performance characterization and points the way to performance improvements. We plan to continue and extend these measurements to include studies over wide temperature ranges and in the presence of vacuum and other controlled environments.

In parallel with the noise studies, we have begun development of a tunneling seismometer, and have presented some early, very encouraging results. In addition, we are developing a tunnel sensor fabrication process which can be carried out at MCNC. Finally, we are developing a single-chip, CMOS-compatible tunnel sensor control circuit which can be obtained through MOSIS. Availability of these devices through foundries should further quicken the pace of tunnel sensor-based device and product development.

ACKNOWLEDGMENTS

We would like to thank W. Kaiser, T. George, S. Manion, C. Storment, and D. Marcus for discussions and support. N. Maluf and G. Kovacs have contributed directly to the integrated controller; the high-voltage circuitry was developed by Maluf. This work was supported by the Center for Space Microelectronics Technology, Jet Propulsion Laboratory, California Institute of Technology, and is sponsored by the National Aeronautics and Space Administration, Office of Space Access and Technology. We also acknowledge the National Science Foundation CAREER Award (ECS-9502046), the Charles Lee Powell Foundation, and the Terman Fellowship. We also thank The National Science Foundation for support of travel to this conference.

REFERENCES

[1] T.W. Kenny, J.K. Reynolds, J.A. Podosek, E.C. Vote, L.M. Miller, H.K, Rockstad, and W.J. Kaiser, "Micromachined Infrared Sensors using Tunneling Displacement Transducers,"*Rev. Sci Instrum, 67* (1996), p. 112.

[2] H.K. Rockstad, J.K. Reynolds, T.K. Tang, T.W. Kenny, W.J. Kaiser, and T.B. Gabrielson, "A Miniature, High-Sensitivity, Electron Tunneling Accelerometer," *Proceedings, Transducers '95, Vol. 2*, (1995) p. 675 .

[3] F.T. Hartley, B. Dolgin, and P.M. Zavracky, "Progress Towards a Bulk Micromachined Tunneling Tip Microaccelerometer," *Proceedings, Transducers '95, Late News Volume*, (1995) p. 75.

[4] C. Yeh and K. Najafi, "A Low-Voltage Bulk-Silicon Tunneling-Based Accelerometer," *to be published*, (1995).

[5] C. Yeh and K. Najafi, "A Bulk-Silicon TUnneling-Based Pressure Sensor", *Proceedings 1994 IEEE Solid State Sensor and Actuator Workshop*, (1994) p. 123.

[6] L.M. Miller, J.A. Podosek, E. Kruglick, T.W. Kenny, J.A. Kovacich, and W.J. Kaiser, "A μ-Magnetometer Based on Electron Tunneling," *Proceedings 1996 IEEE-MEMS Workshop*, (1996) p. 467 .

[7] T.W. Kenny, W.J. Kaiser, H.K. Rockstad, J.K. Reynolds, J.A. Podosek, and E.C. Vote, "Wide-Bandwidth Mechanical Elements for Tunneling Transducers," *J. MicroElectroMechanical Systems 3*, (1994) p. 97 .

[8] Mark Products L-28, Mark Products, 10502 Fallstone Road, Houston, Texas, 77099

[9] T.B. Gabrielson, "Mechanical-Thermal Noise in Micromachined Acoustic and Vibration Sensors," *IEEE Electron Devices, 40*, (1993) p. 903.

[10] D. Kobayashi, T. Hirano, T, Furuhata, and H. Fujita, "An Integrated Lateral Tunneling Unit," *Proceedings 1992 IEEE MEMS Workshop*, (1992) p. 214 .

[11] N.I. Maluf, R.J. Reay, and G.T.A. Kovacs, "High-Voltage Devices and Circuits Fabricated using Foundry CMOS for use with Electrostatic MEM Actuators," *Proceedings, Transducers '95, Vol. 1*, (1995) p. 158 .

A MICROMACHINED VARIABLE CAPACITOR FOR MONOLITHIC LOW-NOISE VCOS

Darrin J. Young, Bernhard E. Boser
Department of EECS, University of California
Berkeley, California 94720

ABSTRACT

An aluminum micromachined variable capacitor is proposed as the tuning element in voltage-controlled oscillators (VCOs) in applications such as cellular phones. Compared to the traditional solutions relying on varactor diodes, this approach is amenable to monolithic integration in a standard electronic circuit process without sacrificing performance. An experimental device consisting of four parallel microstructures is tunable between 2.11pF and 2.46pF with a 5.5V DC tuning voltage. At 1GHz, the quality factor (Q) is 62, limited by the resistive loss in the interconnect traces.

INTRODUCTION

Increased demand for wireless communication motivates a growing interest in monolithic all-CMOS personal communication transceivers [1]. Sub-micron technologies meet RF circuit speed requirements, but still depend on off-chip components to implement some key building blocks. Off-chip devices result in increased fabrication cost and power dissipation, and reduced reliability.

The various cellular phone standards, for example, require VCOs with frequencies in the low Gigahertz range and a tuning range of 50 to 100MHz. Channel spacing of only 30KHz calls for extremely low phase-noise of the oscillator. Monolithic solutions based on ring oscillators [2] cannot meet this requirement. Since the phase-noise exhibited by this type of oscillator is a function of power dissipation rather than electronic circuit speed, little improvement can be expected from future technology scaling.

Current VCO designs in personal communication devices rely on an off-chip LC tank circuit with a Q of at least 10 to meet the low phase-noise requirements. Typical element values are on the order of 2pF for the capacitor, and 6nH for the inductor. Frequency tuning is achieved by modulating the depletion width of a varactor diode. Unfortunately, this solution has been proven difficult to realize in monolithic form. While in recent years substantial progress has been made to improve the quality of on-chip spiral inductors, high-quality varactor diodes require special materials that are not available in standard IC processes and hence are not suited for monolithic integration. Silicon diodes suffer both from an excessive series loss and hence low quality factor, and from a limited tuning range [3].

In this paper, a new solution based on a micromachined parallel-plate variable capacitor is proposed. The device consists of a thin sheet of aluminum suspended in air approximately 1.5μm above the substrate. A DC voltage results in an electrostatic force pulling the movable plate closer to the substrate, hence increasing the capacitance. Theoretically, a tuning range of up to 50% can be achieved. The required tuning voltage can be chosen by appropriately dimensioning the suspension. This represents an important advantage over varactor diodes which have a tuning range that is a strong function of the supply voltage. For a typical 3V maximum tuning voltage, the adjustment range of varactor diodes usually cannot cover both the required oscillator tuning range plus typical manufacturing tolerances; hence, it requires (manual) tuning. Compared to varactor-based oscillators, a larger oscillation amplitude is possible with micromachined capacitors for two reasons. First, unlike the varactor, the proposed capacitor does not have the potential problem of becoming forward-biased. Second, it does not suffer from high-frequency distortion because of its mechanical damping.

Despite the better mechanical properties and the vast experience with polysilicon microstructure fabrication [4], aluminum [5, 6] is used as the structural material in this work for several reasons. First, its lower sheet resistance is critical to minimize ohmic losses and guarantee an adequate quality factor. The second key advantage of aluminum is the low processing temperature of only 150^0C for the proposed fabrication procedure. Because of this, the variable capacitors can be fabricated on top of wafers with completed electronic circuits without affecting device characteristics. This is particularly crucial in applications such as RF where the availability of the most recent IC technology is a critically competitive advantage. The inferior mechanical properties of aluminum compared to polysilicon are not critical for this application.

MICROMACHINED VARIABLE CAPACITOR

Figure 1 shows top and cross-sectional views of the variable capacitor. It consists of a thin aluminum plate suspended in air nominally 1.5μm above a bottom aluminum layer and anchored with four mechanical folded-beam suspensions acting as springs. The folding relaxes the stress in the film and thus reduces the warpage [7]. The symmetry also avoids systematic tilt of the structure. The warping is mainly caused by the thermal expansion mismatch between the sacrificial material (photoresist) and the top aluminum layer. The elevated temperature during the deposition and subsequent cooling to the room temperature further contributes to a built-in strain gradient in the aluminum film and causes the structure to warp after removal of the sacrificial layer.

The chosen plate size of 200 by 200 μm^2 and 1.5μm nominal air-gap result in a nominal capacitance value of approximately 200fF. These dimensions represent a compromise between the achievable capacitance value and the maximum structure size that can be fabricated reliably without excessive warping. Difficulties in removing the sacrificial material preclude a smaller air-gap despite the potential for higher capacitance density and hence smaller overall size. The achieved dimensions present a substantial advance for micromachined aluminum structures. Previously reported results describe much smaller structures with larger gap size [5, 6].

A DC bias voltage applied across the capacitor results in an electrostatic pull-down force and consequent reduction of the air-gap. The suspended plate can be pulled down at most by 1/3 of the original gap size before the pull-down force exceeds the mechanical restoring force causing the plate to be pulled all the way to the sub-

0-9640024-1-8 $3.00 © 1996 TRF

Solid-State Sensor and Actuator Workshop
Hilton Head, South Carolina, June 2-6, 1996

strate. This corresponds to an up to 50% increase in capacitance value.

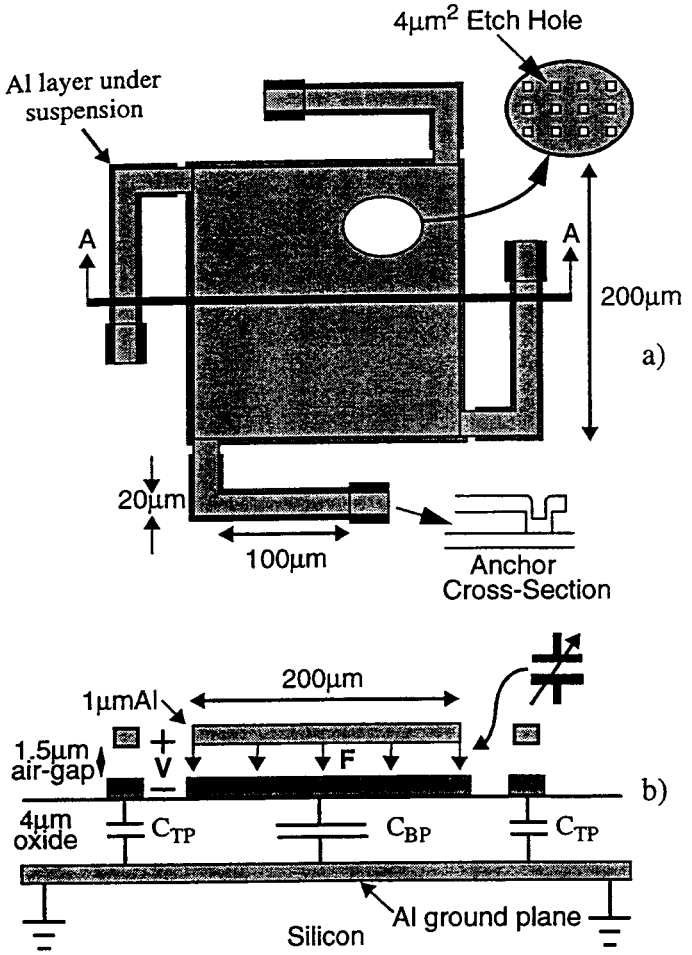


Figure 1. a) Top view, b) cross-section of a micromachined variable capacitor

In RF transceivers, the DC tuning voltage is typically limited to 3V or less by the supply voltage. Assuming a 1.5μm nominal gap size and 80ng mass of the 200 by 200 μm² plate, this calls for a suspension spring constant of 3.8N/m and corresponds to a resonant frequency of 35KHz. The suspension consists of four beams of 100μm length and 20μm width.

To demonstrate the concept, four variable capacitors are wired in parallel to obtain a 400fF tuning range and a 800fF nominal capacitance (Figure 5). The actual overall capacitance is larger because of fixed parasitics. A Q of at least 40 at 1GHz is needed to achieve a similar performance as a discrete varactor diode. This corresponds to a total series resistive loss of less than 2Ω.

Figure 1b shows the parasitics capacitances, C_{TP} and C_{BP}, between the top and bottom plates of the variable capacitor and the substrate. In a VCO application, the bottom plate is grounded and C_{BP} is therefore shorted. However, the top-plate parasitics, C_{TP}, appear in parallel with the variable capacitor. This not only reduces the tuning range, but can also lower the quality factor of the overall capacitor. These problems are alleviated with a separated aluminum layer directly on the silicon substrate isolated from the capacitor bottom-plate with a 4μm thick oxide. The thick oxide minimizes the value of the parasitic capacitance. The aluminum layer shields the parasitics from the lossy substrate, thus ensuring a high Q. Possibly the aluminum ground plane could be replaced with an interconnect aluminum layer to eliminate some fabrication steps.

In this design the bottom-plate aluminum extends from the anchor under the suspension up to the edge of the movable plate. This conservative design has been chosen to prevent the capacitors from shorting out when the suspension touches the bottom plate. However, it results in a relatively large parasitic capacitance of 220fF for all four suspensions and substantially reduces the tuning range. The capacitor shorting has been found not to be a problem in experimental devices. In the next version, the layout shown in Figure 2 will be used, resulting in an increase of the variable capacitance by about 100fF and a decrease of C_{TP} to 70fF. The interconnect traces of the four individual capacitors shown in Figure 5 add approximately 280fF parasitic capacitance and 0.7Ω and 0.3Ω series resistance to the wiring of the top and bottom plates respectively.

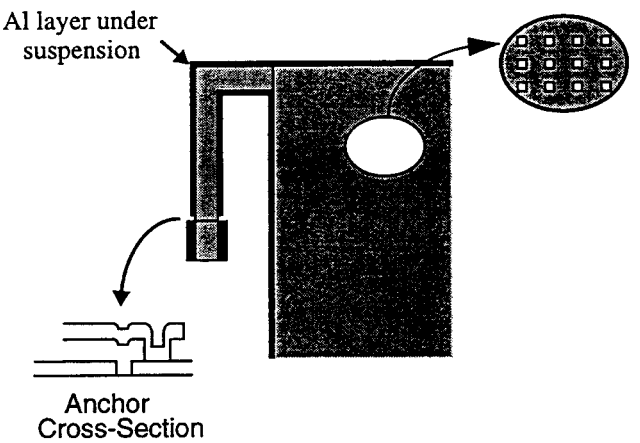


Figure 2. Suggested new layout

FABRICATION PROCESS

To experiment with the processing, the variable capacitor has first been fabricated on a bare silicon wafer. Fabrication on top of completed electronic circuits is planned as a next step.

Figure 3 illustrates the process flow. First, a 1μm thick aluminum film is sputtered onto the bare silicon wafer to form a low-resistance ground plane. Next, a 4μm low-temperature oxide is deposited (Fig. 3a). The bottom plate of the capacitor and interconnect trace consist of another 1μm thick sputtered aluminum film (Fig. 3b). Wet etching is chosen to pattern this layer since the lateral dimensions of the structures are not critical. A 1.5μm thick photoresist serves as a sacrificial layer. This material offers excellent control of thickness and uniformity, is photo-definable and easy to process. Contact windows to the bottom aluminum are opened in the resist (Fig. 3c). Then it is baked at 120°C for 45 minutes to prevent out-gassing during the subsequent deposition of the top aluminum layer. Longer baking time would result in hardening of the resist and consequent difficulties in its removal. The final 1μm aluminum layer is sputtered on at a reduced power and also wet etched since a plasma dry etch would cause the sacrificial layer to outgas. This layer forms anchors to the bottom aluminum inside the contact windows through the resist (Fig. 3d). In addition, it is deposited on all interconnects to lower the resistance (not shown in the figure). Etch holes of 2 by 2μm² size spaced 10μm apart ensure a complete removal of the sacrificial photoresist.

At this point, the wafers are coated with photoresist to protect against particulates during dicing. After dicing and removal of particulates, the protective resist and sacrificial layer are removed with an oxygen-based dry etch, thus releasing the capacitor top-plate (Fig. 3e).

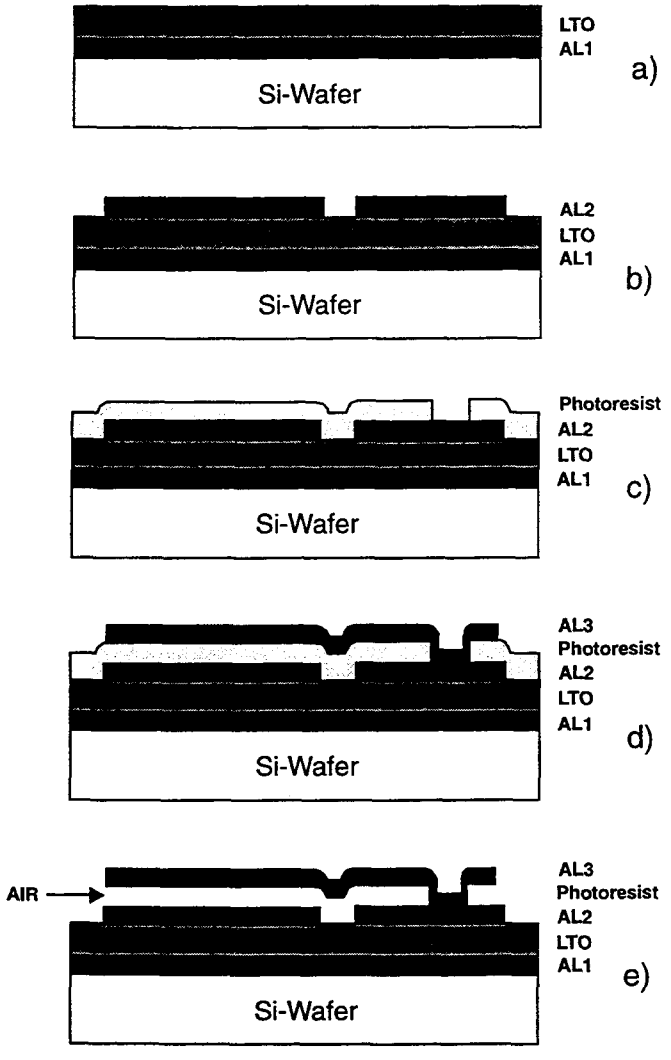


Figure 3. Process flow of the post-CMOS Al micromachining technology:
a) ground plane (1μm aluminum) and LTO deposition,
b) first aluminum layer deposition (1μm) and patterning,
c) photoresist sacrificial layer (1.5μm),
d) second aluminum layer (1μm) with anchor,
e) released structure after etching photoresist.

The oxygen gas pressure and plasma power must be set carefully to ensure a complete removal of the resist and prevent the aluminum film from warping significantly. The structure heats up to 150^{0}C during the 180-minute long etch. An optimal pressure of 500mT and 150W power have been determined experimentally. The dry etching process virtually eliminates the problem of the structure sticking to the substrate. This is an important advantage over the wet release procedure typically used with polysilicon microstructures.

EXPERIMENT RESULTS

Figure 4 shows SEM micrographs of a single tunable capacitor. Optical inspection reveals doming of the capacitor top-plate. At the center of the structure, the gap size is 0.5μm larger than the nominal value of 1.5μm, and the corners are 0.3μm below the nominal height.

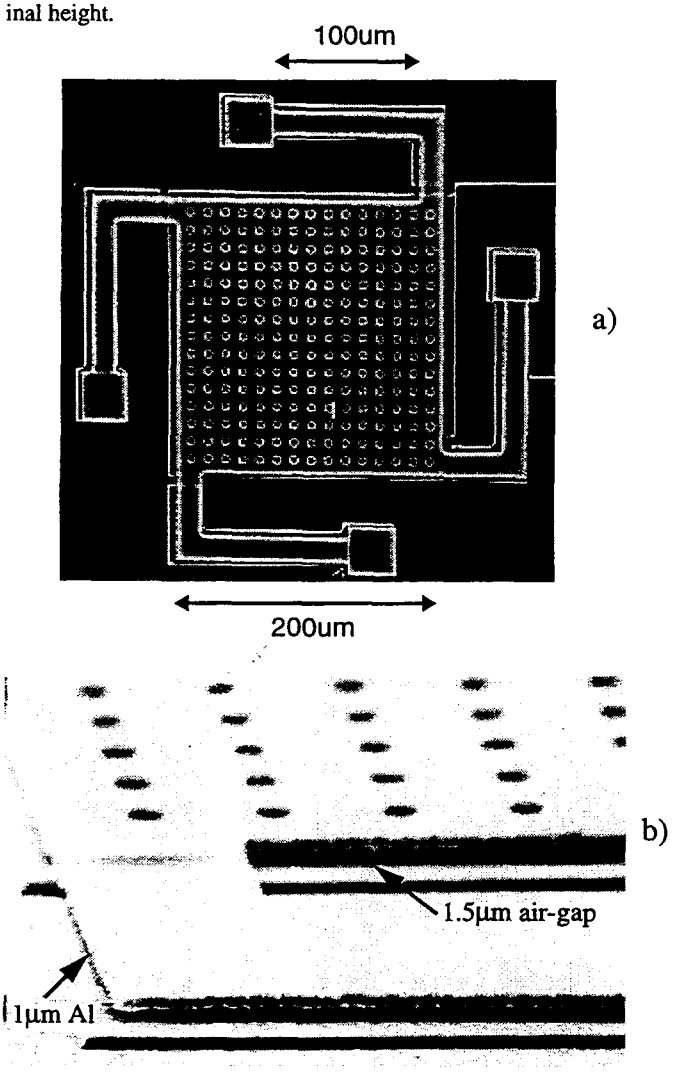


Figure 4. SEM of 200 x 200 μm^2 micromachined aluminum tunable capacitor
a) Top view, b) Suspension detail

Figure 5 presents the parallel connection of four variable capacitor with a nominal capacitance value of 2pF designed for a

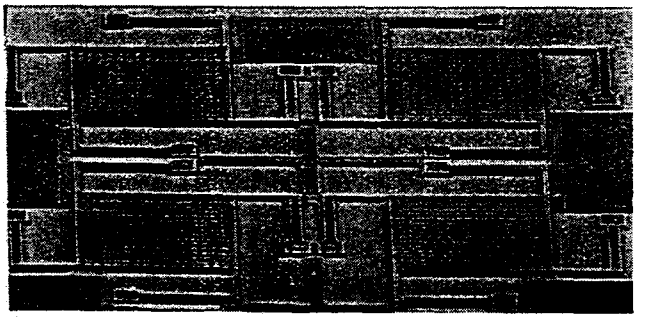


Figure 5. Four parallel microstructures

experimental VCO. With a 5.5V tuning voltage, the capacitance value can be varied between 2.11pF and 2.46pF. This corresponds to a tuning range of 16% shown in Figure 6. The tuning voltage is larger than the designed 3V but can be lowered by decreasing the width of the suspension. The tuning range can be increased by appropriate layout changes as illustrated in Figure 2 and described above.

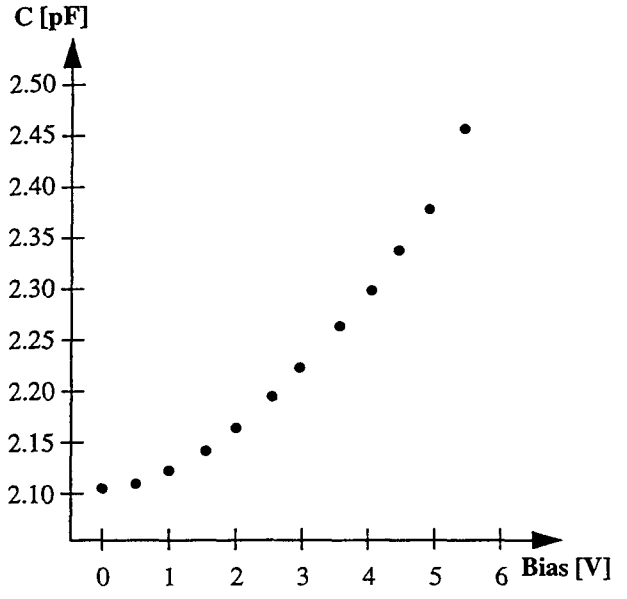
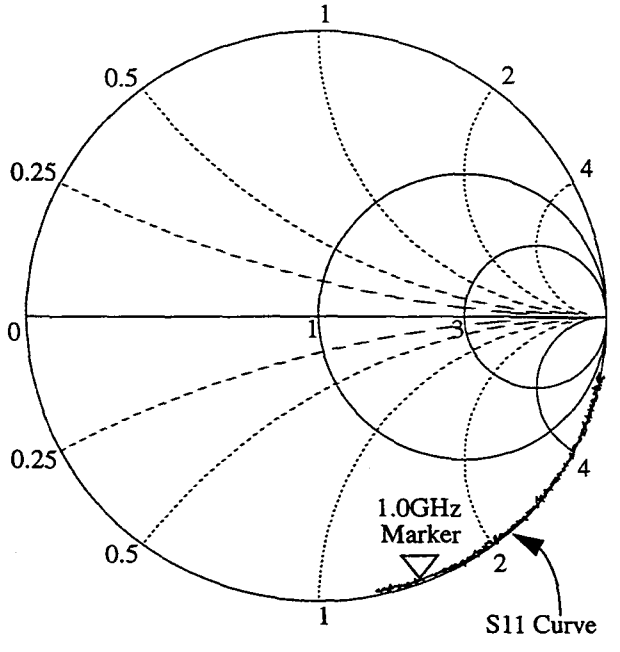


Figure 6. Measured capacitance as a function of the applied DC Bias

Aside from the capacitance, the quality factor is the most important electrical parameter of the structure. It is most accurately determined with an S-parameter measurement. Figure 7 shows the Smith Chart for the four parallel structures measured with an HP8719 network analyzer for frequency up to 1.2GHz. The S11



START: 0.1600GHz, STOP:1.200GHz

Figure 7. Measured S11 in a Smith Chart

curve follows closely the lower-half of the unit circle, as is expected for a capacitor. At 1GHz, the series resistance is 1.2Ω,. This value only slightly exceeds the projected wiring resistance and corresponding to a Q of 62. It matches or exceeds the quality factor of varactor diodes and is at least an order of magnitude larger than that of a typical junction capacitor implemented in a standard IC process.

CONCLUSION

Aluminum micromachined variable capacitors constitute an attractive alternative to conventional varactor diodes as the tuning element in voltage-controlled oscillators for personal communication devices. The proposed device meets or exceeds the performance of typical varactors and can be added to standard IC processes with a few simple back-end processing steps. Further advantages over a varactor-based solution include the possibility to operate at lower supply voltages and a larger oscillation amplitude.

Unlike polysilicon micromachining processes, this technology requires no changes of the underlying IC process, a feature that may be attractive also for other applications, for example inertial sensors. The performance of such sensors has not yet been determined.

ACKNOWLEDGMENTS

The authors thank James Bustillo from the Berkeley Sensor and Actuator Center for his help with device fabrication. All experiments have been carried out in the Berkeley Microfabrication Laboratory. Financial support from NSF is gratefully acknowledged.

REFERENCES

[1] P. R. Gray, R. G. Meyer, "Future Direction in Silicon ICs for RF Personal Communications", *Technical Digest*, IEEE Custom Integrated Circuits Conference, Santa Clara, CA, May 1995, pp. 83-90.

[2] M. Thamsirianunt, T. A. Kwasniewski, "A 1.2μm CMOS Implementation of a Low-Power 900-MHz Mobile Radio Frequency Synthesizer", *Technical Digest*, IEEE Custom Integrated Circuits Conference, San Diego, CA, May 1994, pp. 383-386.

[3] M. Soyuer, K. A. Jenkins, J. N. Burghartz, M. D. Hulvey, "A 3V 4GHz nMOS Voltage-Controlled Oscillator with Integrated Resonator", *Technical Digest*, IEEE International Solid-State Circuits Conference, San Francisco, CA, February 1996, pp.394-395.

[4] R. T. Howe, "Recent Advances in Surface Micromachining", *Technical Digest*, The 13th Sensor Symposium, 1995, pp. 1-8.

[5] L. J. Hornbeck, "Deformable-Mirror Spatial Light Modulators", in Spatial Light Modulators and Application III, *Proceeding of The SPIE - The International Society For Optical Engineering* (1990), vol. 1150, San Diego, CA, August 1989, pp. 86-102.

[6] C. W. Storment, D. A. Borkholder, V. Westerlind, J. W. Suh, N. I. Maluf, and G. T. A. Kovacs, "Flexible, Dry-Released Process for Aluminum Electrostatic Actuators", in *Journal of Microelectromechanical System*, vol. 3, No. 3, September, 1994, pp.90-96.

[7] A. P. Pisano, Y. H. Cho, "Mechanical Design Issues in Laterally-Driven Microstructures", *Sensor and Actuators*, A21-A23, 1990, pp. 1060-1064.

A Fully Differential Lateral ΣΔ Accelerometer with Drift Cancellation Circuitry

Mark A. Lemkin
Berkeley Sensor and Actuator Center
497 Cory Hall

Bernhard E. Boser
Electronics Research Laboratory
572 Cory Hall

David M. Auslander
Department of Mechanical Engineering
5120 Etcheverry Hall

University of California, Berkeley, CA 94720

ABSTRACT

A monolithic, surface micromachined, force-balanced capacitive accelerometer is presented. The mechanical sense element provides a true differential capacitive output allowing differential circuitry to be used throughout the system. Even in this differential topology low frequency noise and offset in position sense electronics appear in the output, harming bias stability and dynamic range. An extended feedback topology is presented to eliminate these problems. The effectiveness of this method is confirmed with experimental measurements and simulation results.

INTRODUCTION

Sigma-Delta micromachined accelerometers provide digital output, high linearity, large dynamic range, and are significantly less sensitive to processing fluctuations than their open loop counterparts as reported in [1, 2]. The digital output provided by ΣΔ accelerometers lends itself to embedded applications such as head mounted displays for computer I/O devices, or digital control in the automobile market where digital signal processing of acceleration measurements are desired. CMOS is desirable for building digital circuits due to its scalibility and low power consumption. Analog circuit design in CMOS, however, requires particular attention to low frequency flicker noise and offset. This paper addresses these analog circuit design issues and provides a system topology which solves the problem of drift and offset in ΣΔ accelerometers, extending the useful bandwidth of these sensors down to DC.

A block diagram of a surface micromachined ΣΔ accelerometer is shown in Figure 1. Above its resonant frequency the sense element acts as a mechanical double integrator, integrating acceleration twice in order to obtain position. Position is measured through electrical integration of charge resulting from a voltage pulse applied to a differential sense capacitor array [3]. The output of the position sense circuitry is amplified before entering a comparator which determines whether the proof-mass is above or below the zero position. Electrostatic feedback, controlled by the 1-bit position output, force balances the proof-mass. High resolution is achieved by averaging, or filtering, the 1-bit output. The filtered output represents the amount of force needed to balance the proof-mass at the zero position. This ability to trade bandwidth for accuracy is the basis by which ΣΔ data converters operate [4].

In this paper, a fully differential ΣΔ accelerometer is first described. Next, sources of low frequency noise and offset are discussed and their effect on sensor output analyzed. A method for cancelling low frequency noise and offset in the position sense circuitry is then presented. The effectiveness of this method is confirmed with measurement results from a fabricated device and computer simulations.

A FULLY DIFFERENTIAL ΣΔ ACCELEROMETER

A block diagram of a fully differential, surface micromachined ΣΔ accelerometer is shown in Figure 1. A schematic diagram of the differential sense capacitor array used for sensing position is shown in Figure 2. Sense capacitors are formed between stationary comb fingers and the proof-mass comb fingers, which act as a common node to the sense capacitors. Common centroid layout of the sense capacitors provides first order cancellation of off-axis translational and angular accelerations [3]. The ΣΔ loop is implemented as a sampled data system, allowing time multiplexing of different phases for sensing and feedback. Position is measured by applying a voltage pulse to the proof-mass and integrating the resulting charge difference [3, 5]. A fully differential charge integrator is used for position sensing, providing a differential output voltage directly proportional to proof-mass displacement. The output of the charge integrator is amplified by a preamp before entering a comparator. The comparator provides additional amplification of the small voltage difference from the preamp output and incorporates a regenerative latch to determine the sign of small voltage differences. A regenerative latch utilizes positive feedback to convert small voltage differences to digital output values providing the one-bit A/D conversion necessary for a ΣΔ converter.

The closed loop system is unstable without compensation, as described in [6]. Compensation of this system is accomplished in the feedback path via a 2-tap FIR lead filter, which provides 27 degrees of phase lead at one quarter the clock frequency ($f_s/4$). The 2-bit output of the compensator is converted to an analog voltage by an on chip D/A converter. This analog voltage provides electrostatic force feedback during the feedback phase of the cycle through a separate set of comb fingers.

All circuit components are integrated with the polysilicon proof-mass on the same die, except the external clock generator. Off chip clock generation allows comparison of system performance with and without the drift cancellation technique discussed in the section on noise and offset cancellation.

NOISE AND OFFSET

Noise and offset enter at the mechanical sense element as well as in the position sense interface. Through linear system analysis the transfer function between noise injected at the sensor (N_s^2 in Figure 1) and the output may be found to be unity, while noise injected in the position sense interface (refered to N_i^2 in Figure 1) is shaped by the inverse of the proof-mass transfer function. For a pure double integrator, corresponding to an unconstrained mass in a ΣΔ accelerometer, the transfer function between electrical noise at a frequency f injected in the position sense interface and the output is $(2\pi f)^2$. Noise power is attenuated by $(2\pi f)^4$, the square of the noise

0-9640024-1-8 $3.00 © 1996 TRF

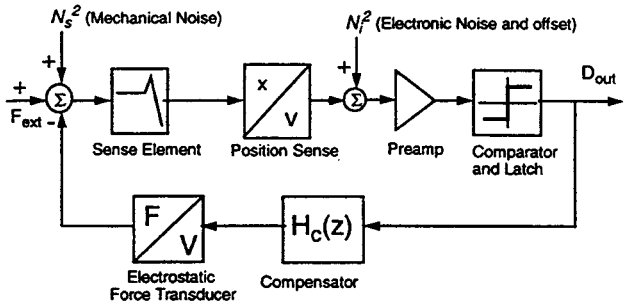


Figure 1 ΣΔ force feedback loop. Mechanical (N_s^2) and Electrical (N_i^2) noise sources shown. Electrical noise sources referred to N_i^2.

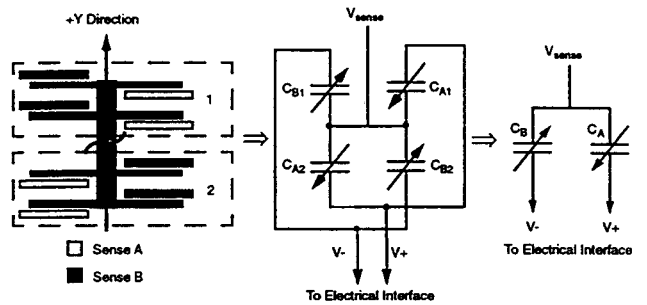


Figure 2 Schematic of fully differential, common centroid, sense capacitor array. Drive capacitors not shown for clarity.

transfer function. This relation shows noise and offset are greatly attenuated at low frequencies. This is also approximately the case in ΣΔ voltage A/D converters, where the integrator pole is set by opamp gain and can have a very low frequency. However, for surface micromachined sensors practical considerations such as warpage and shock survival set a lower bound of several kHz on the natural frequency of the sense element. This results in a noise shaping characteristic in which attenuation of sensor interface noise stops for frequencies below the proof-mass resonant frequency.

Brownian motion, due to the small mass of the sensor, results in a white acceleration noise in the sense element. At atmospheric pressure the Brownian noise for the sensor described in this paper is approximately 200 μg/√Hz. This may be attenuated by running the accelerometer in a vacuum. Mismatch between sense capacitors provides a DC offset indistinguishable from an acceleration input. Typically this mismatch is a relatively small percentage of the sense capacitance due to good matching commonly achievable with modern photolithography.

Noise and offset in the sense electronics can be quite large however, especially in a CMOS technology. Flicker noise in the input transistors of CMOS amplifiers usually starts to dominate the electrical noise below several hundred kilohertz, with noise power increasing proportional to the inverse of frequency. For example, flicker noise from the input transistors of the preamp, when referred to the output, results in a measured output noise of approximately 2 mg/√Hz at 1 Hz for this sensor. In addition input voltage offsets due to threshold voltage variations and transistor size mismatch, typically on the order of 15mV, generate an output signal of order 2 g, substantially reducing the dynamic range of the system. Changes in temperature affect this amplifier offset resulting in a temperature dependent offset at the sensor output.

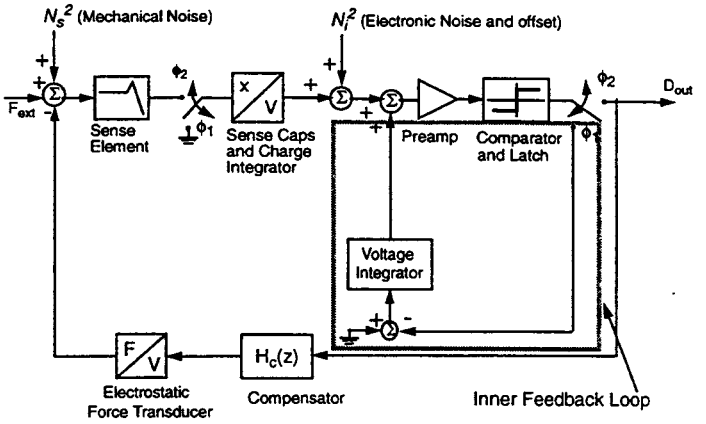


Figure 3 ΣΔ force feedback loop with drift cancellation circuitry. Noise and offset from electronics are referred to the node labeled N_i^2.

NOISE AND OFFSET CANCELLATION

The fundamental reason for poor rejection of low frequency noise and offset in the topology shown in Figure 1 is the high resonant frequency of the proof-mass. At frequencies below the sense element resonant frequency electrical interface noise appears in the output attenuated by a constant factor equal to:

$$\frac{N_{Output}^2}{N_i^2} = \left(\frac{C_i \omega_n^2}{\left(\frac{d}{dx} Cs\right) V_s} \right)^2, \omega < \omega_n \qquad (1)$$

where ω_n is the natural frequency of the sensor, C_i is the capacitance of the charge integrator, V_s is the sense pulse magnitude, and dC_s/dx is the differential change in sense capacitance for a proof-mass displacement. The sensitivity, dC_s/dx, as well as minimum achievable resonant frequency is set by the micromachining technology. Settling constraints in position sense electronics set the value of C_i for a given CMOS technology and sense capacitance. Increasing the sense pulse voltage improves performance only to a point dependent on technology and sampling rate. Further increases harm performance due to distortion imparted on the proof-mass during the sense part of the cycle [6].

Addition of a separate feedback loop to the system can drastically improve the noise transfer function between the sense electronics and the output. Figure 3 shows a diagram of the drift cancellation loop incorporated in the ΣΔ accelerometer. This inner loop is a purely electrical realization of a first order sigma-delta converter. As such the integrator pole of the noise shaping element may be made arbitrarily low. Thus the transfer function between sense position electronic noise and the output will decrease with decreasing frequency down to arbitrarily low frequencies for the inner feedback loop.

By adding one additional clock phase the integrator pole of the inner loop may be used to attenuate offset in the position sense interface. During phase 1 the sense capacitors are zeroed, the position sense electronics are allowed to settle and the comparator is clocked. At this point the comparator is unaffected by the position of the proof-mass, since no sense pulse has been applied. Any offset present in the sense electronics will cause the comparator output to change sign in the direction of the offset. The

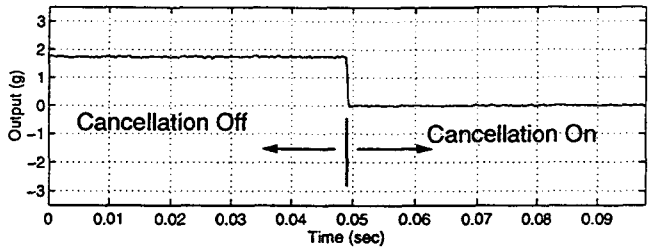


Figure 4 Simulation of a 15mV preamp offset in a 1.9 kHz bandwidth with and without cancellation. Without cancellation the DC output is at 1.8g reducing the useful dynamic range of the sensor.

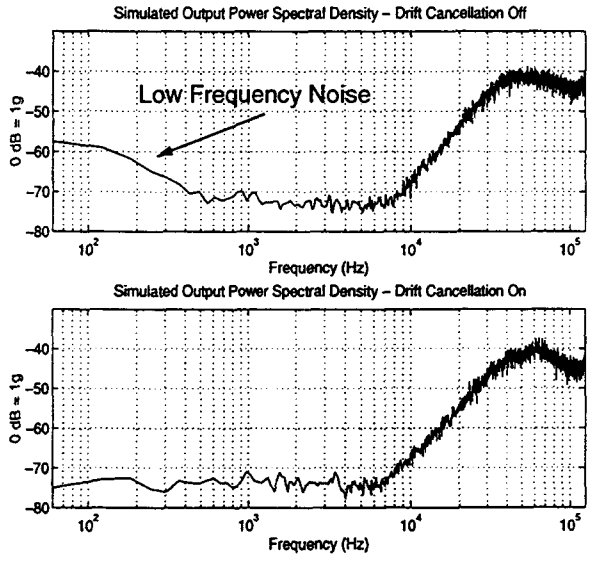


Figure 5 Simulation of ΣΔ accelerometer with low frequency colored noise. The top plot (a), without drift cancellation shows coupling of low frequency noise directly to the output. The inner feedback loop attenuates this noise as shown in the bottom plot (b).

comparator output is negatively fed back to the voltage integrator driving the output of the voltage integrator in the opposite direction of the offset. After many cycles the value needed to just cancel the sense interface offset is reached at the voltage integrator output. At this point the comparator output alternates sign every cycle. The voltage integrator gain is chosen so perturbations from this oscillation are below the quantization noise floor of the outer ΣΔ loop. Low frequency noise, which may be thought of as a slowly varying offset, is canceled in exactly the same manner as a stationary offset.

During the second, or sense phase, a voltage pulse is applied to the sense capacitors, the electronics are allowed to settle, and the comparator is clocked again. Note the voltage integrator, from the inner feedback loop, is still connected to the summing block at the preamp input. Its value is added to the input of the preamp during this phase, cancelling offset and low frequency noise from the sense circuitry. Since the sense pulse is not applied during phase 1, the voltage integrator in the inner loop does not respond to changes in sense element position. Thus the desired signal, from capacitance differences in the differential sense capacitor array, remains unattenuated. The corrected digital output is used for feedback during phase 3, which completes the cycle.

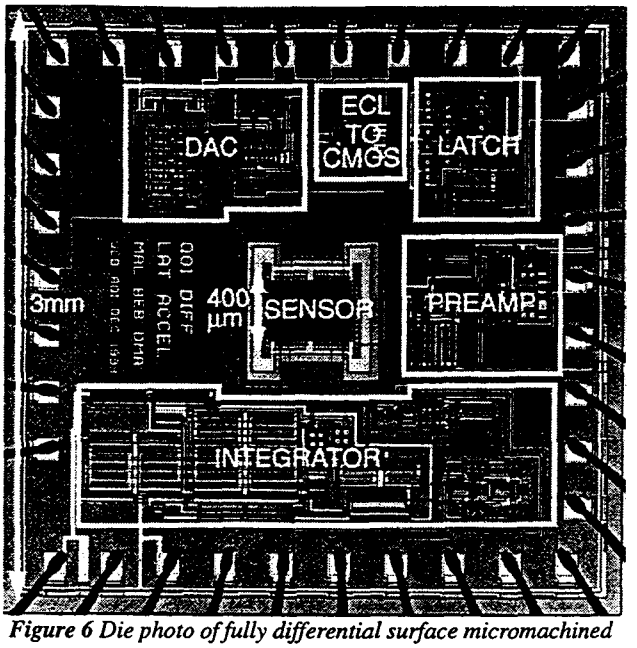


Figure 6 Die photo of fully differential surface micromachined accelerometer. The proof-mass is 400µm x 400µm and weighs 0.2 µgrams.

RESULTS OF SIMULATIONS

The system has been modeled and simulated in MATLAB. Figure 4 shows simulated rejection of a 15 mV preamp offset with zero input acceleration. Without cancellation the DC output is 1.8 g out of a full scale range of +/- 3.5 g, reducing the useful dynamic range of the sensor. When cancellation is turned on the output of the sensor is brought to zero.

Figures 5a and 5b are power spectral density (PSD) plots of the complete sensor output spectrum with colored noise added at low frequency to simulate the effect of flicker noise in the position measurement interface. Because ΣΔ converters are oversampled, the sensor output will be filtered and downsampled to a smaller useful bandwidth. For this sensor the downsampling ratio is on the order of one-thousand times. The top plot, without drift cancellation, shows coupling of low frequency noise directly to the output in typical sensor bandwidths of up to 1 kHz. The inner feedback loop attenuates this noise as shown in the bottom plot.

MEASUREMENT RESULTS

A die photo of the differential test accelerometer is shown in Figure 6. Table 1 lists important accelerometer specifications. For evaluation the accelerometer is mounted on a vibration table and excited with a sinusoidal input signal. A reference accelerometer mounted beside the test accelerometer is used for calibration of the test acclerometer. The one-bit output is decimated with a digital signal processing card connected to a workstation. Data analysis to determine signal to noise ratio and generate power spectral density plots is done off line with MATLAB.

This accelerometer has been tested with and without drift cancellation circuitry enabled. Figure 7a shows measured power spectral density of the sensor output with drift reduction circuitry disabled. A 110 mg RMS input signal at 16 Hz is used as a reference. Flicker noise from the preamp input dominates the noise floor below 3.2 Hz. At 0.1 Hz the noise floor due to 1/f noise in the

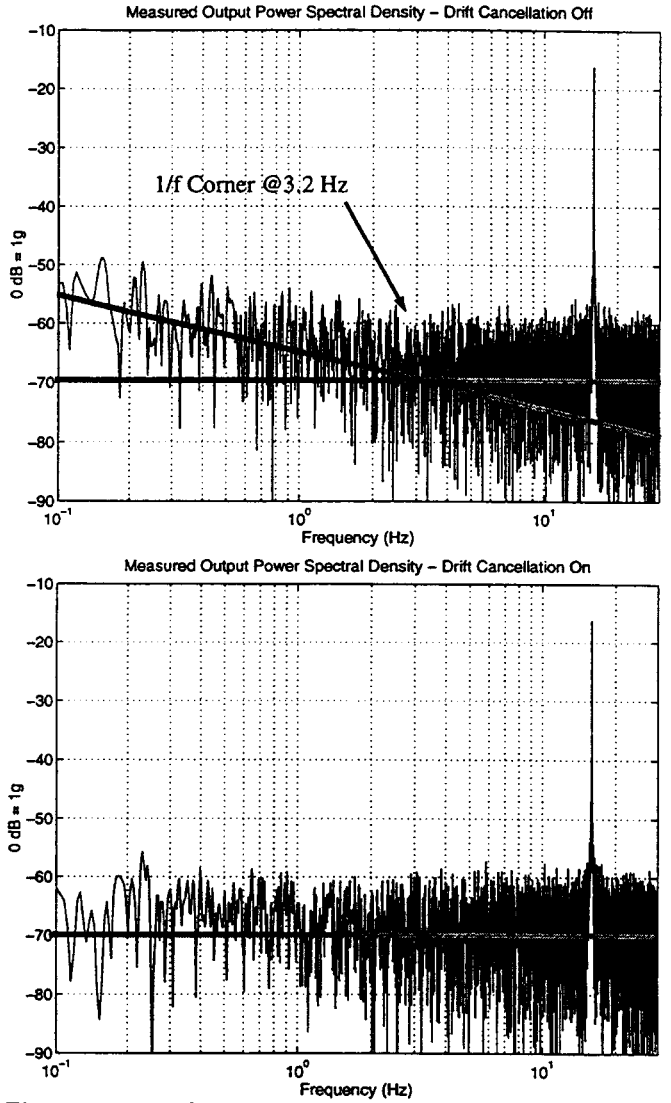


Figure 7 Measured output spectrum. The top plot (a), without drift cancellation, shows coupling of preamp noise directly to the output. The inner feedback loop cancels this noise as shown in the bottom plot (b), providing 13 dB improvement in the noise floor at 0.1 Hz. The two peaks are 110 mg RMS reference inputs at 16 Hz used for calibration. PSD constructed from 415 second data train.

preamp has risen 13 dB to 5 mg/√Hz. To obtain this plot an offset voltage added to the auxilliary stage of the preamp had to be manually adjusted to bring the sensor output to mid-scale, since mismatch in the sense interface caused the sensor output to saturate.

Figure 7b shows the measured sensor PSD with drift reduction enabled. The elimination of 1/f noise in the output is clearly demonstrated by the flat power spectrum. The offset necessary to bring the output to the center of its range was automatically provided by the inner feedback loop.

CONCLUSIONS

The relatively high resonant frequency typical of surface micromachined accelerometers reduces attenuation of low frequency noise and DC offset from the electrical position sense interface. This noise and offset appear in the output and reduce bias stability as well as dynamic range. A time multiplexed, dual feedback loop ΣΔ topology has been presented using an electrical voltage integrator to provide cancellation of noise, offset and drift in position sense electronics. Proof of efficacy has been demostrated both through simulation and measurements from a fabricated device, with an observed 13dB improvement in output noise at 0.1 Hz.

Table 1: Accelerometer Specifications

Parameter	Value
Process	3μm BiCMOS
Power Supply	3.7 mA @ 5V
Die Size	3mm x 3mm
Full Scale	+/- 3.5 g
Noise Floor (fs = 500 kHz) Sense Pulse = 2.5V	500 μg/√Hz
Noise Floor (fs = 250 kHz) Sense Pulse = 1.25V	1.6 mg/√Hz
Resonant Frequency	8.1 kHz
Proof-Mass	0.2 μgram
Total Sense Capacitance	84 fF
Sensitivity	64 fF/μm

ACKNOWLEDGEMENTS

The authors would like to thank Analog Devices for device fabrication and the California PATH program and ARPA for supporting this project. Discussions and contributions by Professors Howe and Pisano are gratefully acknowledged.

REFERENCES

[1] Henrion, W., *et. al.*, "Wide dynamic range direct digital accelerometer," *IEEE Solid-State Sensor and Actuator Workshop*, Hilton Head Island, SC, June 4-7, 1990 pp. 153-157.

[2] Yun, W., Howe, R.T., Gray, P.R., "Surface micromachined, digitally force-balanced accelerometer, with integrated CMOS detection circuitry," *IEEE Solid-State Sensor and Actuator Workshop*, Hilton Head Island, SC, June 22-25, 1992 pp. 126-131.

[3] Lemkin, M. and Boser, B.E., "A micromachined fully differential lateral accelerometer," *CICC Dig. Tech. Papers.* IEEE Custom Integrated Circuits Conference, 1996,

[4] J. C. Candy and G. C. Temes. Oversampling Methods for A/D and D/A conversion. In J. C. Candy and G. C. Temes, editors, *Oversampling Delta-Sigma Data converters*, pages 1-29. IEEE Press, 1992.

[5] Lu, C., Lemkin, M., Boser, B.E., "A monolithic surface micromachined accelerometer with digital output," *IEEE Journal of Solid State Circuits*, pp. 1367-1373, December 1995.

[6] Boser, B. E., Howe, R. T, "Surface Micromachined Accelerometers," *CICC Dig. Tech. Papers.* IEEE Custom Integrated Circuits Conference, 1995, pp. 337-344.

A BI-DIRECTIONAL MAGNETIC MICROPUMP ON A SILICON WAFER

Wenjin Zhang and **Chong H. Ahn**

University of Cincinnati
Center for Microelectronic Sensors and MEMS
Department of Electrical and Computer Engineering and Computer Science
P.O. Box 210030, Cincinnati, Ohio 45221-0030

ABSTRACT

A new electromagnetically driven peristaltic micropump on a silicon wafer with a low driving voltage less than 3 volts has been designed, fabricated, and tested. To construct the micropump, electromagnets on a Pyrex glass wafer and micropump parts on a silicon wafer are separately fabricated, and then assembled together using low temperature polymer bonding techniques. This new fabrication technique of magnetic microactuators using low temperature post-bonding techniques provides a flexible and unique approach in realizing various magnetic MEMS devices. The realized magnetic micropumps have been functionally operated in peristaltic motions with a DC 300 mA and 3 volts allowing bi-direction pumping flows. The maximum attainable flow rate was approximately 20 $\mu l/min$ at 5 Hz. In this work, a new electromagnetically driven bi-directional micropump has been successfully realized on silicon wafers.

INTRODUCTION

There has been a large demand for a micropump which has a high flow rate with a low driving voltage. Most micropumps which have been realized to date are suffering from high driving voltages to attain high flow rates depending on their application targets. In specific, if a micro fluid regulation system is sharing a power supply with signal control or processing IC chips which require less than 3-5 volts, low driving voltage of micropumps is critically essential in practical applications. If the driving voltages of the micropump require those high values that would not be attainable from the power supply with the IC chips, the micropump system must require an extra power supply or voltage regulator. Furthermore, micro fluid regulation systems for biomedical or biological applications such as portable or implantable drug delivery systems require low driving voltages, since electrical breakdown hazard from the high voltages is seriously concerned due to its safety issues.

Micropumps which have a function of bi-directional fluid driving mechanism are usually required in biomedical or chemical analysis systems [1], where the bi-directional pumping actions are using to produce either a washing flow or a sampling flow. For a preliminary bi-directional magnetic micropump, a peristaltic micropump [2-4] using magnetic variable-reluctance principles has been designed and fabricated in this work. Thus, the micropump fabricated in this work consists of an electromagnetically actuated pump chamber and two peristaltic microvalves as shown in Figure 1. By controlling the exciting sequence of the two valves and the pump chamber, its pumping operation can be achieved as well as its flow directions can also be changed. Magnetic forces have been used for the microactuators which require low driving voltages as well as large driving forces [5-7].

In this work, a new proto-type magnetic microactuator as shown in Figure 2 has been proposed and realized using both bulk micromachining and wafer bonding techniques, where magnetic forces can be produced between the electromagnets and the permalloy films plated on silicon diaphragms (i.e., permalloy/silicon diaphragm or magnetic diaphragm). When a current is applied to the inductor mounted on a glass wafer, the permalloy/silicon diaphragm is attracted toward the upper electromagnet, lifting the valve bosses and thus opening the valve as described in [8]. In realizing this structure, the planar inductors and the magnetic diaphragms are separately fabricated and then assembled using low temperature polymer bonding techniques.

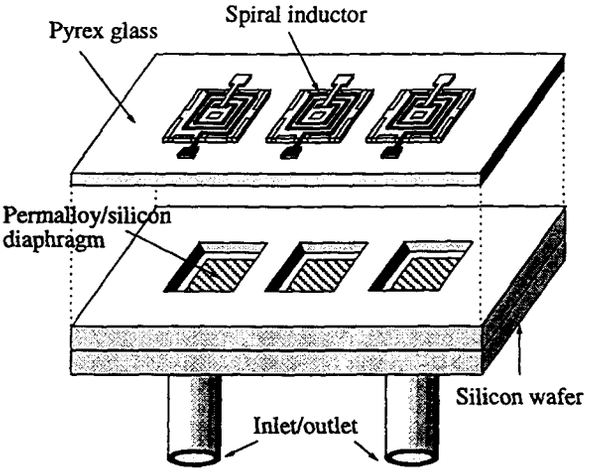


Figure 1. Schematic diagram of a magnetically driven bi-directional micropump.

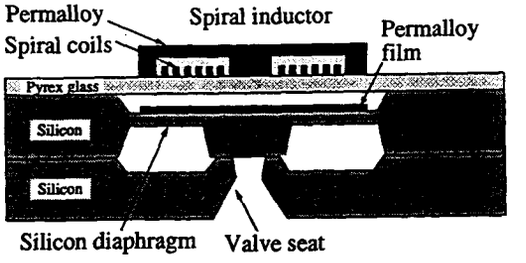


Figure 2. Schematic diagram of a prototype magnetic microactuator providing vertical actuation.

0-9640024-1-8 $3.00 © 1996 TRF

Pyrex glass wafers are serving as an inductor substrate as well as an insulation substrate which passes magnetic flux through it. The magnetic diaphragms and valve chambers are fabricated using conventional bulk micromachining and low temperature eutectic bonding techniques. Thus, a new fabrication technique in realizing magnetic microactuators based on the combination of bulk micromachining and wafer bonding techniques has been introduced, which provides a flexible and unique approach in the fabrication of magnetic microactuators. In this work, a new magnetically driven bi-directional micropump on a silicon wafer with a low driving voltage of 3 volts has been designed, fabricated, and tested.

INDUCTOR AND PUMP DESIGN

Several micromachined planar inductors for the applications of microactuators have been reported [9-11]. Among them, the planar spiral-type inductor [9] has a very favorable structural affinity to the structure of micropumps, microvalves, and several planar microactuators which usually require a vertical actuation of diaphragm. The major concerns in designing this type of inductor are to achieve a high inductance, a low coil resistance, and a low magnetic field interference. In the spiral inductor structure, if magnetic cores with a high permeability encapsulate spiral coils constructing a closed magnetic circuit, the achievable inductance value increases due to the low reluctance of the inter-linked magnetic circuit between the coils and the cores. A micromachined planar spiral inductor which has fully encapsulated magnetic cores was reported in [9]. Accordingly, the magnetic field interference can be minimized, confining the magnetic fluxes in the closed magnetic core. However, the bottom magnetic core from the fully encapsulated inductor is simply removed from the inductor proposed in this work for micropump applications. As described in Figure 2, magnetic path will be completed through the magnetic core electroplated on a silicon diaphragm (i.e., permalloy/silicon diaphragm). The magnetic forces generated in this actuator depend on the variation of magnetic reluctance through this magnetic path including air gaps. A designed spiral inductor to be fabricated on a Pyrex glass wafer is described in Figure 3, where spiral conductor lines are encapsulated from the top with an opening path at the bottom.

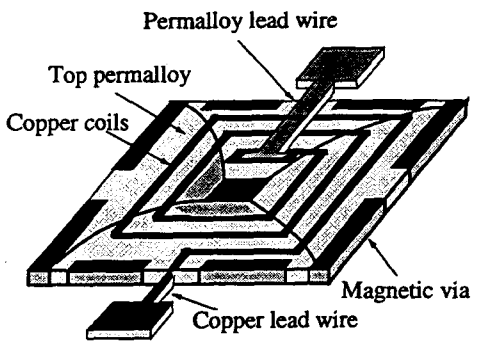


Figure 3. Schematic diagram of the designed spiral inductor.

In order to allow bi-directional fluid flow, this micropump is designed to have a pump actuator and two peristaltic microvalve actuators as shown in Figure 4. The top wafer is used for diaphragms and central bosses, and the bottom wafer is used for valve seats, flow channels, and inlet and outlet holes. The valves consist of a diaphragm with a central boss which mates to an etched silicon valve seat, composing normally-closed diaphragm valves [8]. Silicon diaphragms have a size of 8 mm x 8 mm. Micro flow channels which have a size of 1000 μm x 200 μm are designed to connect between the chambers, and the inlet and outlet holes are designed to have a dimension of 500 μm x 500 μm.

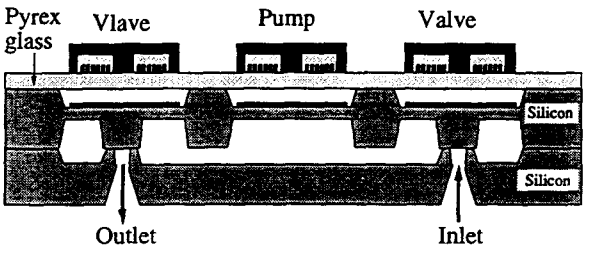


Figure 4. Cut-view of the designed micropump.

FABRICATION

Planar spiral inductors were fabricated using AZ-4000 series thick-photoresist with three mask process. The brief fabrication steps are described in Figure 5.

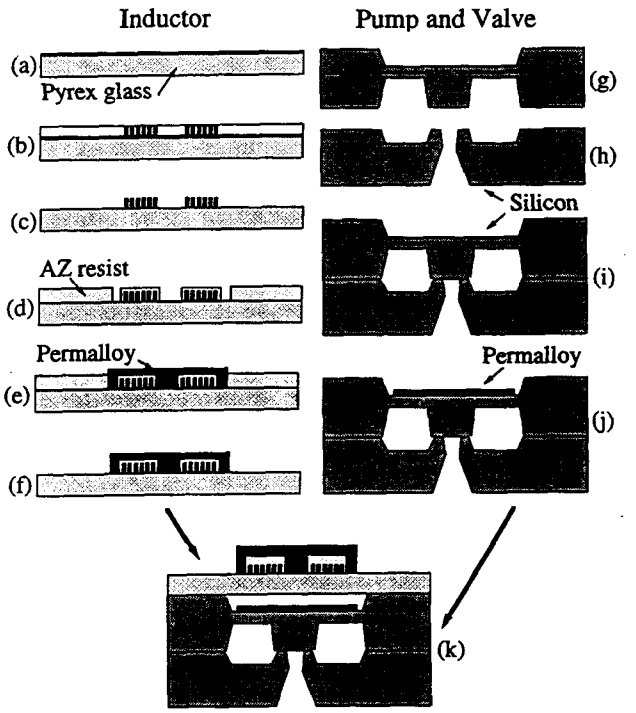


Figure 5. Fabrication steps of inductors and pump parts: (a) deposit seed layer; (b) plate bottom conductor lines; (c) strip PR and seed layer; (d) insulation layer; (e) plate permalloy; (f) strip PR; (g) etch pump chamber; (h) etch valve seat; (i) eutectic bonding; (j) plate permalloy film; (k) assemble two components.

The fabrication of the inductor started with 2-inch diameter and 10 *mil* thick Pyrex 7740 glass wafers. Onto the glass wafers, chromium (300 Å) / copper (3000 Å) / chromium (200 Å) was deposited to form a seed layer for electroplating. AZ-4000 photoresist was used to build the plating molds of spiral conductors with a thickness of 40 um, and then copper was electroplated through the molds. Upon completion of the plating, the photoresist was stripped with acetone, and the seed layer was then removed to isolate the spiral coils from each other. AZ-4000 photoresist was spun on the spiral coils, patterned for conductor via holes, and then cured at 150 °C for 1 hour to make a planarized insulation layer. Onto the cured photoresist, chromium (300 Å) / copper (3000 Å) / chromium (200 Å) was deposited again for the seed layer of the top magnetic core plating. To electroplate simultaneously both the top magnetic core and the permalloy lead wire from vias to bonding pads, the similar processes used in fabricating the spiral conductors were repeated. Figure 6 shows the photomicrograph of the fabricated spiral inductor.

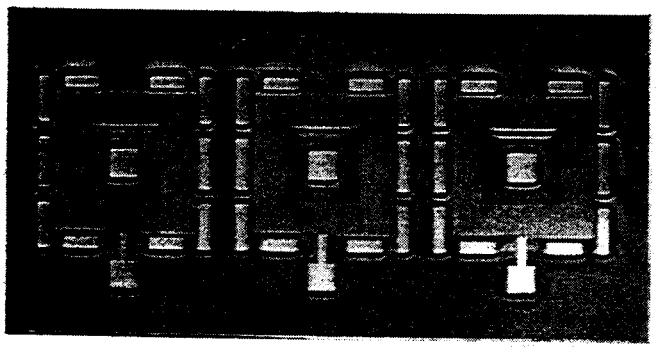


Figure 6. Microphotograph of the fabricated spiral inductor which is taken from rear side through the glass wafer.

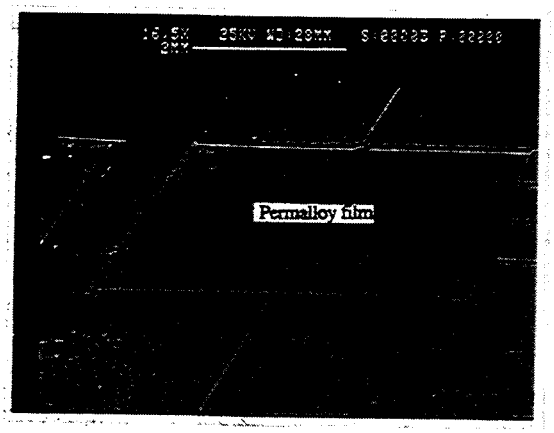


Figure 7. SEM of the fabricated pump part, where the permalloy/silicon diaphragm is shown.

On the other hand, the fabrication of the pump parts started with 2-inch {100} silicon wafers with a silicon oxide of 1.0 μm following the steps described in Figure 5. Diaphragms and central bosses were fabricated on the top wafer using bulk micromachining techniques in *KOH* solution. To provide an actuation gap between the glass wafer and the silicon diaphragm, a silicon of 30 μm was etched from the top of the diaphragm, which allows an ultimate diaphragm thickness of 10 μm. Valve seats, flow channels, and inlet and outlet holes were also anisotropically etched on the bottom wafers.

To bond two wafers using gold eutectic bonding techniques, both wafers were deposited with titanium (300 Å) / gold (3000 Å) layer and then bonded together at 380 °C for 10 minutes. Upon the completion of the bonding, chromium (300 Å) / copper (3000 Å) / chromium (200 Å) seed layer was deposited on the top of the diaphragm for the plating of the magnetic film. A nickel-iron permalloy of 7 μm thick was plated as a magnetic film on the silicon diaphragm [6, 12] to construct a permalloy/silicon diaphragm. Figure 7 shows scanning electron micrograph of the bonded diaphragm and valve seats, where the diaphragm has a size of 8 mm x 8 mm x 17 μm in square shape.

Finally, the separately fabricated spiral inductors and pump parts were assembled using low temperature polymer bonding techniques [13]. The micropump were diced into chips and glass pipes were attached on both the inlet and the outlet. Figure 8 shows the photograph of the micropump mounted on a packaging for test.

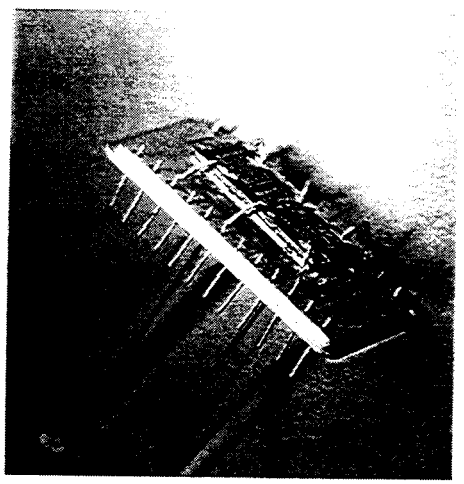


Figure 8. Microphotograph of the assembled micropump which is mounted on the package for test.

EXPERIMENTAL RESULTS

To evaluate the fabricated inductors, the planar inductors fabricated on the glass wafers were diced into chips. The chip was mounted on a ceramic package, and then the inductance of the device was measured using an HP-4192A impedance analyzer. For a typical 20-turn device of 8 mm x 8 mm in area, an inductance of approximately 1.7 μH was measured at 10 KHz from the fabricated spiral inductor. The measured inductance versus frequencies is plotted in Figure 9. The measured conductor resistance was approximately 3 ohms which allows a high exciting current with even low driving voltage less than 3 volts.

A preliminary pumping test was performed for the fabricated micropump using DI water. The DI water was supplied through the inlet and outlet glass pipes with an inside diameter of 1.5 mm. Using a power MOSFET micropump driving circuit, three phase square waves were applied to the pump in keeping the exciting sequence to achieve a sequential peristaltic micropump action. The pumping flow rate of the pump was evaluated by measuring the total amount of pumped fluid for a given duration at a fixed driving frequency. The measurements

were repeated over a frequency range of magnetic diaphragm actuation from 1 Hz to 20 Hz.

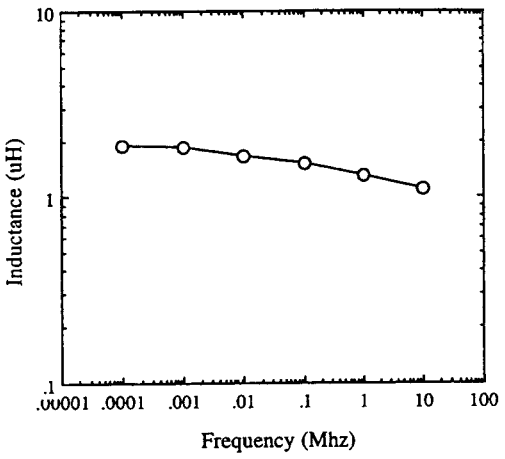


Figure 9. Inductance of the fabricated inductor as function of frequencies.

A maximum flow rate of 20 $\mu l/min$ was attained at 5 Hz. By changing the phase firing order in reverse, the direction of pumping flow was also changed in reverse achieving the bi-directional micropump function. Flow rate, developed pressure, and power consumption as a function of the driving frequencies and voltages are currently being investigated.

CONCLUSION

A new magnetic driven bi-directional peristaltic micropump on a silicon wafer with a low driving voltage less than 3 volts has been designed, fabricated, and tested. Using low temperature post-bonding techniques, the fabrication techniques demonstrated in this work provide a flexible and unique approach in realizing various magnetic MEMS components. The realized magnetic micropumps have been functionally operated with a DC 300 mA and 3 Volts showing bi-direction pumping flows. The maximum attainable flow rate was approximately 20 $\mu l/min$ at 5 Hz. In this work, a new electromagnetically driven bi-directional micropump has been successfully realized on Pyrex glass and silicon wafers.

ACKNOWLEDGEMENT

The authors wish to thank Profs. H. Thurman Henderson and Joseph H. Nevin and all members of Center for Microelectronic Sensors and MEMS at the University of Cincinnati for their help and technical discussion. The authors would also like to gratefully acknowledge Hoechst Celanese Co., DuPont, OCG Microelectronics Materials, and DOW Plastics for their donations of materials.

REFERENCES

[1] V. Gass, B. H. van der Schoot, S. Jeanneret, and N. F. de Rooij, "Integrated Flow-Regulated Silicon Micropump", Sensors and Actuators A, Vol. 43, pp. 335-338, 1994.

[2] J. A. Folta, N. F. Raley, and W. E. Hee, "Design, Fabrication and Testing of a Miniaturized Peristaltic Membrane Pump", Proc. IEEE Solid-State Sensors and Actuators Workshop, pp. 186-189, 1992.

[3] R. Zengerle, S. Kluge, M. Richter, and A. Richter, "A Bi-directional Silicon Micropump", Proc. IEEE Microelectromechanical Systems Workshop, pp. 19-24, 1995.

[4] J. G.Smiths, "Piezoelectric Micropump with Three Valves Working Peristaltically", Sensors and Actuators, A21-A-23, pp. 203-206, 1990.

[5] K. Yanagisawa, H. Kuwano, and A. Tago, "An Electromagnetically Driven Microvalve", Proc. Transducers'93, pp. 102-105, 1993.

[6] C. H. Ahn, Yong. J. Kim, and Mark. G. Allen, "A Planar Variable Reluctance Magnetic Micromotor with Fully Integrated Stators", IEEE/ASME Journal of Microelectromechanical Systems, Vol. 2, No. 4, pp. 165-173, 1993.

[7] C. H. Ahn and M. G. Allen, "Fluid Micropumps Based on Rotary Magnetic Actuators", Proc. IEEE Microelectromechanical Systems Workshop, pp. 408-412, 1995.

[8] H. Jerman, "Electrically-Activated, Normally-Closed Diaphragm Valves", Journal of Micromechanics and Microengineering, Vol. 4, No. 4, pp. 210-216, 1994.

[9] C. H. Ahn and M. G. Allen, "A Planar Micromachined Spiral Inductor for Integrated Magnetic Microactuator Applications", Journal of Micromechanics and Microengineering, Vol. 3, pp. 1-9, 1993.

[10] C. H. Ahn and M. G. Allen, " A New Toroidal-Meander Type Integrated Inductor with Multilevel Meander Magnetic Core", IEEE Transactions on Magnetics, Vol. 30, No. 1, pp. 73-79, 1994.

[11] C. H. Ahn, Y. J. Kim, and M. G. Allen, "A Fully Integrated Planar Toroidal Inductor With A Micromachined Nickel-Iron Magnetic Bar", IEEE Transactions on Components, Hybrids, and Manufacturing Technology, Vol. 25, No 3, pp. 356-403, 1994.

[12] R. L. Anderson, E. E. Castellani, P. M. McCaffrey, and L. T. Romankiw, "Method for Treating Magnetic Alloy to Increase The Magnetic Permeability", United States Patent #4,003,768

[13] J. C. Bolger and C. T. Mooney, "Die Attach in Hi-Rel P-Dips: Polyimides or Low Choride Epoxies", Proc. IEEE Electronics Components, pp. 63-67, 1984.

TWO-PHASE ACTUATORS: STABLE ZIPPING DEVICES WITHOUT FABRICATION OF CURVED STRUCTURES

J.R. Gilbert and S.D. Senturia

Microsystems Technology Laboratories, MIT, Cambridge MA 02139

ABSTRACT

The design of electrostatic actuators optimized for both high force generation as well as long and stable displacement is difficult. Several groups have proposed and/or implemented devices based on "zipping" to achieve improvements in both areas simultaneously. In these designs the implementation of a zipping actuator requires the fabrication of a curved surface on which one electrode touches another. The fabrication of controlled curved surfaces often presents difficulties, especially those perpendicular to the wafer surface. In this paper we propose and model a class of devices, two-phase actuators, which can obtain similar effects without need for fabrication of curved structures.

INTRODUCTION

Practical use of electrostatic actuation often encounters two difficulties: (1) the electrostatic force scales as 1/x, giving very short displacement if large forces are needed, and (2) electrostatic spring-mass systems exhibit pull-in instabilities and so have stable controllability only over a fraction of their travel. Much of the design of electrostatic actuators is devoted to working around or utilizing these two effects. Several groups have proposed and/or implemented devices based on "zipping" to overcome both of these effects [1,2,3,4]. In these designs the implementation of a "zipping" actuator with stable motion requires the fabrication of a curved surface on which one electrode touches another. We propose a class of devices, two-phase actuators, which can obtain similar effects without the need for fabrication of curved structures. This design concept emerged in our work on MEMCAD and the CoSolve-EM package for 3D simulation of coupled electromechanics[5] and contact electromechanics[6]. In this paper we use those tools to simulate a voltage controlled torsion mirror (VCM) as a representative example of this class of devices.

ZIPPING ACTUATORS

The simplest conceptual design of a zipping actuator is given by Branedbjerg. An S-shaped flexible electrode is positioned next to a flat ground plane and separated from it by a thin dielectric layer. Such a device is sketched in Figure 1. The flexible electrode has one end clamped high and the other end clamped low, near the ground layer. The device has an S-shaped transition region from its high end to its low end. A voltage is applied between the electrode and ground, and the transition region moves from right to left. This motion of the transition region is one measure of the "displacement" of the device. Because the electrostatic loads on the structure vary as 1/x, where x is the distance between the electrode and ground, the circled region in Figure 1B gives the dominant contribution to the driving force. The shape of this driving region is essentially maintained as the transition region moves from right to left, and so the driving force is nearly constant as the transition region moves.

Most zipping actuators share these two qualitative features: a curved transition region that is displaced under operation, and a region of close apposition of the conductors that provides the dominant contribution to the driving force of the device.

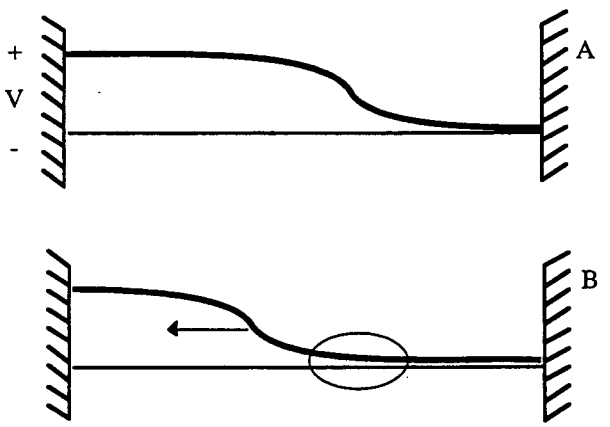


Figure 1 A Schematic example of a "zipping" actuator. Increasing the voltage V drives the mid-point of the "S" to the left. In all states of the actuator its driving force is dominated by the region where the electrode and ground are most closely in contact, but not yet touching-this region is circled in (B). The dielectric layer over the ground plane is not shown.

TWO-PHASE ACTUATOR DESIGN

The two-phase (or multiple-phase) actuator concept is to produce the curved moving electrode by the operation of a "starter" structure, instead of fabricating that curve directly.

A proto-typical two-phase actuator has one movable conducting part, the "plate", and two or more fixed electrodes. Either the plate or both fixed electrodes are insulated to prevent shorts upon contact. The plate is held at some fixed voltage V_F. One electrode, the "starter", deforms the plate into initial contact using a voltage, V_S. The other electrode, the "driver", is then in close contact with the plate, and may drive it through stable zipping motion by ramping its drive voltage, V_D, from 0 to some V_{MAX}. As an example of such a two phase actuator we will discuss a voltage controlled mirror (VCM). Figure 2 shows a 3D view of the VCM. In Figure 3 its dimensions are sketched. The "plate" is modeled in this study as aluminum, with tensile residual stress of 40MPa. The plate has a thick mirror region, and a long thin neck. At the end of the neck there is a wider area used to reduce the initial starting voltage. The mirror region is suspended on two tethers which are labeled in Figure 2. These

0-9640024-1-8 $3.00 © 1996 TRF

tethers are purposely set off center in the mirror to allow the mirror to tilt to almost 40 degrees, before it will touch down at the point where it joins the neck. There is an insulating contact layer 0.05μm above the Driver and Starter electrodes. In operation the thickness of this separation layer determines the scale of the driving force of the device.

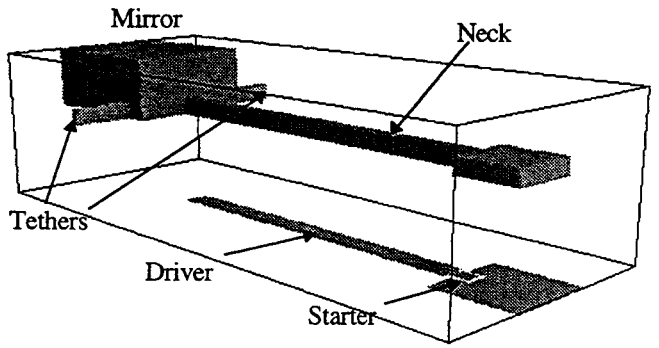


Figure 2 A 3D view of the VCM structure. The "Plate", with labeled regions: Mirror, Neck, and Tethers, is 3.0μm above the Driver and Starter electrodes. There is an insulating contact layer 0.05μm above the Driver and Starter electrodes.

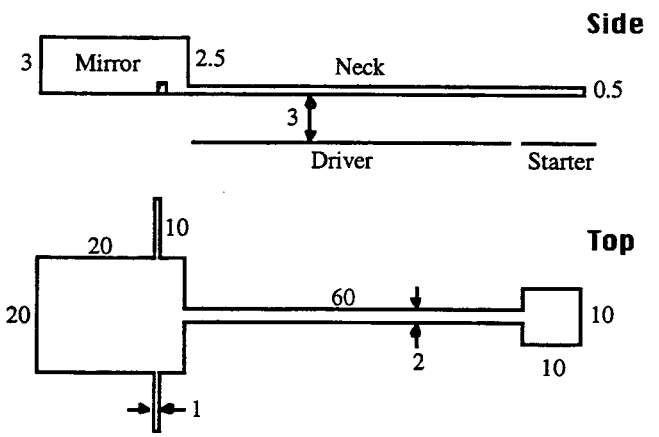


Figure 3 Top and Side views of the VCM. All numbers in μm.

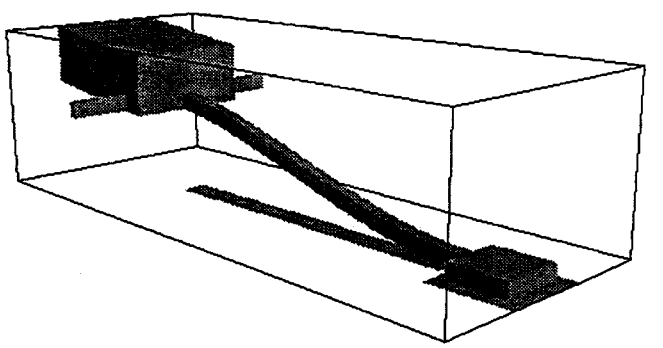


Figure 4 The Model at drive voltage V_D=0. The starter has pulled in the mirror plate, establishing the basic curve for the operation of the VCM. The mirror has an initial tilt angle of 1.4 degrees.

TWO-PHASE ACTUATOR SIMULATION

We built a detailed 3D model of the VCM, using the MEMCAD system. Its quasi-static behavior was then simulated using CoSolve-EM v1.0.

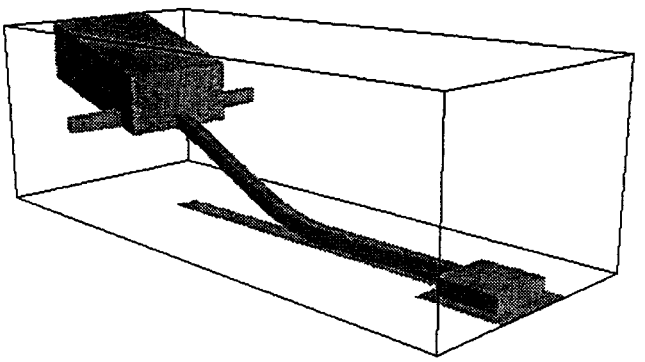


Figure 5 The Model at drive voltage V_D=20V. The drive is zipping in the neck of the mirror plate, causing the mirror to tilt. The tilt angle is 3.5 degrees.

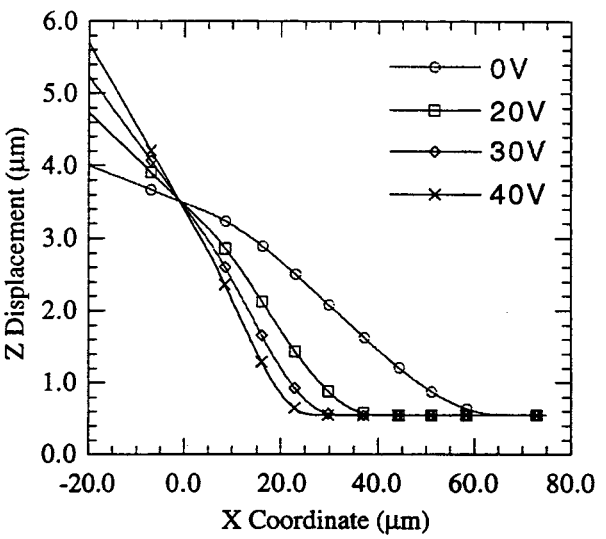


Figure 6 X-Z cross-sections through y=0 on the bottom surface of the plate at four different V_D. The mirror itself extends from -20μm to +4μm and rotates rigidly as the device zips in. The cross-section of the undisplaced surface is constant at z=3.5μm.

In simulation we apply V_S=20V to the starter electrode, and the long neck of the plate pulls in to its initial operating position, as shown in Figure 4. A curved shape has now been established by electrostatic actuation and the driver electrode is in contact with the plate at the location nearest the starter. The shape in Figure 4 is considered the initial operating state of the VCM. In operation we control the driving voltage V_D. As we ramp V_D from 0-80V the neck of the movable conductor zips in, causing the mirror to tilt. Figure 5 shows the VCM at V_D=20V. The transition region of the zipper has moved towards the mirror. Figure 6 plots cross-sections along the length of the

bottom surface of the plate at four drive voltages, 0V, 20V, 30V, and 40V. These cross-sections show the transition region moving in towards the mirror, and the mirror tilting, while remaining undistorted. Figure 7 shows the dependence of mirror tilt angle on V_D over the operating region of the VCM, V_D=0-80V, for which ϕ=1.4-10.3 degrees. The angular dependence is surprisingly linear for so complex a structure. The mirror is designed to tilt up to 40 degrees, though we have not yet simulated it that far. In a more standard electrostatic torsion mirror design, allowing for the possibility of such high angles would substantially reduce the available area and force for driving the mirror. The VCM is stable over its operating region: that is, the mirror exhibits no hysteresis. This is a distinct advantage of operation in the zipping mode, essentially beyond the pull-in point. A more standard torsional micromirror might have stable motion only above pull-in, over approximately one third of its range of tilt.

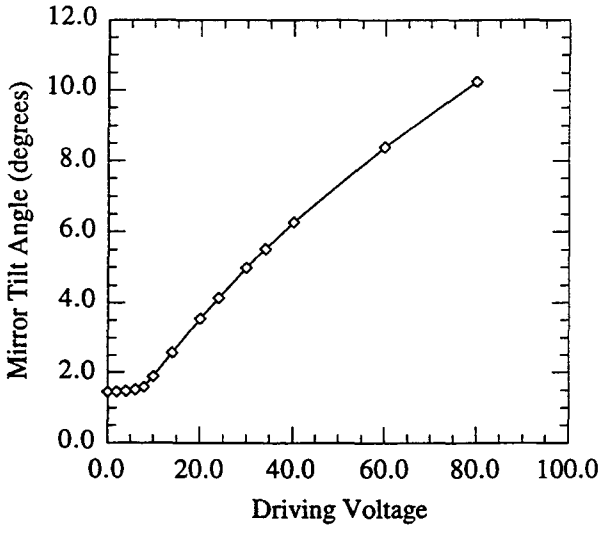


***Figure 7** Driving voltage V_D vs. the mirror tilt angle. We obtain near linear control of tilt angle over most of the VCM's operating range of 1.4-10.3 degrees.*

CONCLUSION

We have proposed a class of electromechanical actuators, two-phase actuators, that have the benefits of zipper actuators, without the need for fabrication of controlled curved surfaces. We have simulated one particular example of a two-phase actuator, the Voltage Controlled Mirror (VCM). The VCM that is simulated here is compatible with fabrication by surface micromachining techniques. The underlying concept, however, is broadly applicable to a variety of MEMS fabrication technologies. The VCM exhibits continuous stable control of its tilt angle over a full 10 degrees, a result which is difficult to duplicate in a surface micromachined process, if the mirror must operate above pull-in. The modeling and simulations of the VCM were done using the MEMCAD system. This work is also a demonstration of the capability to do complex design and analysis tasks in advance of fabrication using MEMCAD.

ACKNOWLEDGMENTS

This work has been supported in part by DARPA contract J92-FBI-196 and by equipment grants from Hewlett Packard and Digital Equipment. The authors would like to thank Rob Legtenberg for many valuable discussions and insights.

1 Branedbjerg and Gravesen, "A New Electrostatic Actuator Providing Improved Stroke Length and Force", Proc. MEMS 1992, Travemünde, Germany, pp. 6-11.

2 K. Sato and M. Shikida, "Electrostatic Film Actuator with a Large Vertical Displacement", Proc. MEMS 1992, Travemünde, Germany, pp. 1-5.

3 R. Legtenberg, et al., "Electrostatic Curved Electrode Actuator", Proc. MEMS 1995, Amsterdam, pp. 37-42.

4 M. Yamaguchi, S. Kawamura, K. Minami, and M. Esashi, "Distributed Electrostatic Micro Actuator", Proc. MEMS 1993, Ft. Lauderdale, pp. 18-23.

5 J.R. Gilbert, R. Legtenberg, and S.D. Senturia, "3D Coupled Electro-Mechanics for MEMS: Applications of CoSolve-EM", Proc. MEMS 1995, Amsterdam, pp. 122-127.

6 J.R. Gilbert, G.K. Ananthasuresh, and S.D. Senturia, "3D Modeling of Contact Problems and Hysteresis in Coupled Electro-Mechanics", Proc MEMS 1996, San Diego.

THICK-FILM GAS MICROSENSORS USING CERMET MATERIALS, CYCLIC VOLTAMMETRY AND NEURAL NETWORKS.

Erika L. Shoemaker, Michael C. Vogt, Tara N. Turner
Energy Systems Division, Argonne National Laboratory
Argonne, IL 60439

ABSTRACT

The sensing characteristics of a thick-film cermet cell, expressed as *MeO_x/Pt/solid electrolyte/Pt* were investigated by applying cyclic voltammetry across the Pt electrodes and measuring changes in ionic current through the solid electrolyte. The sensors were fabricated using thick-film screen printing techniques with high-temperature ceramic-metallic (cermet) materials. In the presence of the cyclic applied voltage, gases react at the three phase, gas/solid electrolyte/electrode boundary (tpb), and cause a change in measured ionic current flow through the sensor. These changes resulted in distinct current-voltage envelopes (voltammograms) for different gases. Neural signal processing was used to characterize and analyze the sampled voltammetry signatures.

INTRODUCTION

In recent years, the need for improved, affordable gas microsensors in the fields of safety, environmental/emissions monitoring, process control, and other applications has grown tremendously. Argonne National Laboratory (ANL) has been developing an economical new type of gas microsensor to evaluate and demonstrate a unique gas sensing technology. The device has demonstrated the potential to address the needs for many industrial, safety, and environmental applications.

The technology involved applying cyclic voltammetry to sandwiched thick-film layers of cermet materials and identifying and quantifying the generated electrical signatures using neural signal processing. The cyclic voltammetry techniques produced complete, broad-spectrum responses. Early attempts at applying voltammetry to electrochemical cells found no elegant way to quantify the responses, but computational neural networks (CNN's) treat the complex sampled voltammetry signatures as simple 1-D patterns and allow training, recognition, quantification, and classification of the signatures.

THEORY

The cell: *Me-MeO_x* ($p(O_2)$) */Pt /solid electrolyte /*Pt was used in this study, with *Me-MeO_x*, a metal/metal-oxide mixture in thermodynamic equilibrium with oxygen gas (O_2), providing a reference oxygen partial pressure, $p(O_2)$. When a cyclic voltage was superimposed across the galvanic cell, ions at the tpb were forced through the solid electrolyte. When the applied potential was negative, oxygen was pumped into the *Me-MeO_x* reference. When the applied potential was positive, oxygen was driven out of the *Me-MeO_x* reference [1]. The number of oxygen ions available at the tpb was greatly reduced for solid-electrolyte systems as compared to aqueous electrolytes. In this cell, only a fraction of the electrode surface was part of the tpb and was electrocatalytically active. The largest portion of the electrode was not in contact with the solid electrolyte and was only catalytically active [2].

The response of the solid-electrolyte sensor studied was primarily dependent on the amount of oxygen anions (O^-) available at the tpb. In this cell, the porous platinum electrodes catalyzed the charge transfer reactions:

$$O(tpb) + 2e^- \rightarrow O^{2-} \quad \text{(cathodic sweep)} \tag{1}$$

$$O^{2-} - 2e^- \rightarrow O(tpb) \quad \text{(anodic sweep).} \tag{2}$$

The oxygen anions were supplied to the solid electrolyte and measured as electronic current through the sensor. The catalytic reactions of gases exposed to the surface of the electrodes either add or remove the oxygen available for transport at the tpb. Different types of reactions were possible. The gas exposed to the electrode could react catalytically, creating oxygen at the tpb as in the reaction:

$$CO_2 + e^- \rightarrow CO + O^- . \tag{6}$$

In addition, adsorbed gases could react with the electrode, freeing bound oxygen as in the reaction:

$$PtO + X(gas) + Ze^- \rightarrow X(gas) + Pt + O^- . \tag{7}$$

Other unreactive gases could hinder the oxygen at the tpb by blanketing the electrode.

The sensors consist of four overlapping thick-film layers (Figure 1) deposited on an aluminum oxide substrate with a platinum heating element on the back. The four layers are two porous catalytic electrodes, an oxygen-ion-conducting solid-electrolyte layer, and a metal oxide oxygen pressure reference layer. The two catalytic electrodes sandwich the solid electrolyte layer. The metal oxide reference layer is deposited beneath the lower electrode. These four layers serve three major functions.

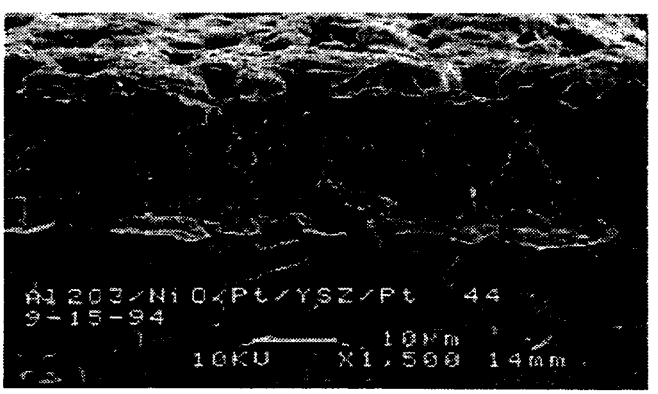


Figure 1. Sensor Micrograph Profile of Al_2O_3/Ni-NiO/Pt/YSZ/Pt Cell at 1500X

The catalytic electrodes promoted electrochemical reactions on their surfaces. In the driven mode, the catalytic electrodes conducted electrons created by the charge-transfer reactions at the tpb. The current measured between the lower and upper electrodes during a voltage scan provided quantitative and

0-9640024-1-8 $3.00

bottom surface of the plate at four drive voltages, 0V, 20V, 30V, and 40V. These cross-sections show the transition region moving in towards the mirror, and the mirror tilting, while remaining undistorted. Figure 7 shows the dependence of mirror tilt angle on V_D over the operating region of the VCM, V_D=0-80V, for which ϕ=1.4-10.3 degrees. The angular dependence is surprisingly linear for so complex a structure. The mirror is designed to tilt up to 40 degrees, though we have not yet simulated it that far. In a more standard electrostatic torsion mirror design, allowing for the possibility of such high angles would substantially reduce the available area and force for driving the mirror. The VCM is stable over its operating region: that is, the mirror exhibits no hysteresis. This is a distinct advantage of operation in the zipping mode, essentially beyond the pull-in point. A more standard torsional micromirror might have stable motion only above pull-in, over approximately one third of its range of tilt.

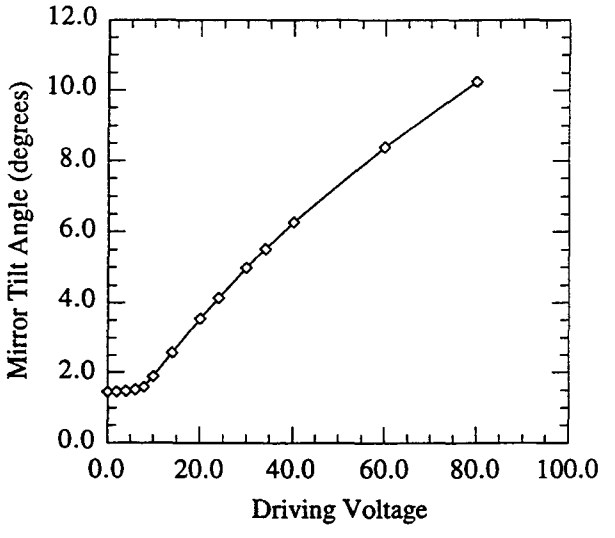


***Figure 7** Driving voltage V_D vs. the mirror tilt angle. We obtain near linear control of tilt angle over most of the VCM's operating range of 1.4-10.3 degrees.*

CONCLUSION

We have proposed a class of electromechanical actuators, two-phase actuators, that have the benefits of zipper actuators, without the need for fabrication of controlled curved surfaces. We have simulated one particular example of a two-phase actuator, the Voltage Controlled Mirror (VCM). The VCM that is simulated here is compatible with fabrication by surface micromachining techniques. The underlying concept, however, is broadly applicable to a variety of MEMS fabrication technologies. The VCM exhibits continuous stable control of its tilt angle over a full 10 degrees, a result which is difficult to duplicate in a surface micromachined process, if the mirror must operate above pull-in. The modeling and simulations of the VCM were done using the MEMCAD system. This work is also a demonstration of the capability to do complex design and analysis tasks in advance of fabrication using MEMCAD.

ACKNOWLEDGMENTS

This work has been supported in part by DARPA contract J92-FBI-196 and by equipment grants from Hewlett Packard and Digital Equipment. The authors would like to thank Rob Legtenberg for many valuable discussions and insights.

1 Branedbjerg and Gravesen, "A New Electrostatic Actuator Providing Improved Stroke Length and Force", Proc. MEMS 1992, Travemünde, Germany, pp. 6-11.

2 K. Sato and M. Shikida, "Electrostatic Film Actuator with a Large Vertical Displacement", Proc. MEMS 1992, Travemünde, Germany, pp. 1-5.

3 R. Legtenberg, et al., "Electrostatic Curved Electrode Actuator", Proc. MEMS 1995, Amsterdam, pp. 37-42.

4 M. Yamaguchi, S. Kawamura, K. Minami, and M. Esashi, "Distributed Electrostatic Micro Actuator", Proc. MEMS 1993, Ft. Lauderdale, pp. 18-23.

5 J.R. Gilbert, R. Legtenberg, and S.D. Senturia, "3D Coupled Electro-Mechanics for MEMS: Applications of CoSolve-EM", Proc. MEMS 1995, Amsterdam, pp. 122-127.

6 J.R. Gilbert, G.K. Ananthasuresh, and S.D. Senturia, "3D Modeling of Contact Problems and Hysteresis in Coupled Electro-Mechanics", Proc MEMS 1996, San Diego.

THICK-FILM GAS MICROSENSORS USING CERMET MATERIALS, CYCLIC VOLTAMMETRY AND NEURAL NETWORKS.

Erika L. Shoemaker, Michael C. Vogt, Tara N. Turner
Energy Systems Division, Argonne National Laboratory
Argonne, IL 60439

ABSTRACT

The sensing characteristics of a thick-film cermet cell, expressed as *MeO_x/Pt/solid electrolyte/Pt* were investigated by applying cyclic voltammetry across the Pt electrodes and measuring changes in ionic current through the solid electrolyte. The sensors were fabricated using thick-film screen printing techniques with high-temperature ceramic-metallic (cermet) materials. In the presence of the cyclic applied voltage, gases react at the three phase, gas/solid electrolyte/electrode boundary (tpb), and cause a change in measured ionic current flow through the sensor. These changes resulted in distinct current-voltage envelopes (voltammograms) for different gases. Neural signal processing was used to characterize and analyze the sampled voltammetry signatures.

INTRODUCTION

In recent years, the need for improved, affordable gas microsensors in the fields of safety, environmental/emissions monitoring, process control, and other applications has grown tremendously. Argonne National Laboratory (ANL) has been developing an economical new type of gas microsensor to evaluate and demonstrate a unique gas sensing technology. The device has demonstrated the potential to address the needs for many industrial, safety, and environmental applications.

The technology involved applying cyclic voltammetry to sandwiched thick-film layers of cermet materials and identifying and quantifying the generated electrical signatures using neural signal processing. The cyclic voltammetry techniques produced complete, broad-spectrum responses. Early attempts at applying voltammetry to electrochemical cells found no elegant way to quantify the responses, but computational neural networks (CNN's) treat the complex sampled voltammetry signatures as simple 1-D patterns and allow training, recognition, quantification, and classification of the signatures.

THEORY

The cell: *Me-MeO_x* ($p(O_2)$) */Pt /solid electrolyte /*Pt was used in this study, with *Me-MeO_x*, a metal/metal-oxide mixture in thermodynamic equilibrium with oxygen gas (O_2), providing a reference oxygen partial pressure, $p(O_2)$. When a cyclic voltage was superimposed across the galvanic cell, ions at the tpb were forced through the solid electrolyte. When the applied potential was negative, oxygen was pumped into the *Me-MeO_x* reference. When the applied potential was positive, oxygen was driven out of the *Me-MeO_x* reference [1]. The number of oxygen ions available at the tpb was greatly reduced for solid-electrolyte systems as compared to aqueous electrolytes. In this cell, only a fraction of the electrode surface was part of the tpb and was electrocatalytically active. The largest portion of the electrode was not in contact with the solid electrolyte and was only catalytically active [2].

The response of the solid-electrolyte sensor studied was primarily dependent on the amount of oxygen anions (O^-) available at the tpb. In this cell, the porous platinum electrodes catalyzed the charge transfer reactions:

$$O(tpb) + 2e^- \rightarrow O^{2-} \quad \text{(cathodic sweep)} \tag{1}$$

$$O^{2-} - 2e^- \rightarrow O(tpb) \quad \text{(anodic sweep).} \tag{2}$$

The oxygen anions were supplied to the solid electrolyte and measured as electronic current through the sensor. The catalytic reactions of gases exposed to the surface of the electrodes either add or remove the oxygen available for transport at the tpb. Different types of reactions were possible. The gas exposed to the electrode could react catalytically, creating oxygen at the tpb as in the reaction:

$$CO_2 + e^- \rightarrow CO + O^- . \tag{6}$$

In addition, adsorbed gases could react with the electrode, freeing bound oxygen as in the reaction:

$$PtO + X(gas) + Ze^- \rightarrow X(gas) + Pt + O^- . \tag{7}$$

Other unreactive gases could hinder the oxygen at the tpb by blanketing the electrode.

The sensors consist of four overlapping thick-film layers (Figure 1) deposited on an aluminum oxide substrate with a platinum heating element on the back. The four layers are two porous catalytic electrodes, an oxygen-ion-conducting solid-electrolyte layer, and a metal oxide oxygen pressure reference layer. The two catalytic electrodes sandwich the solid electrolyte layer. The metal oxide reference layer is deposited beneath the lower electrode. These four layers serve three major functions.

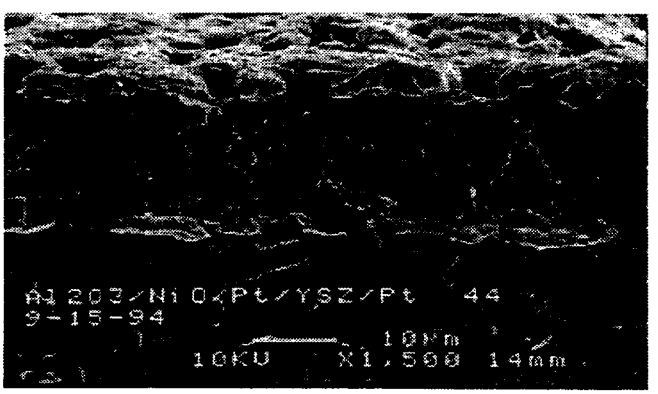


Figure 1. Sensor Micrograph Profile of Al_2O_3/Ni-NiO/Pt/YSZ/Pt Cell at 1500X

The catalytic electrodes promoted electrochemical reactions on their surfaces. In the driven mode, the catalytic electrodes conducted electrons created by the charge-transfer reactions at the tpb. The current measured between the lower and upper electrodes during a voltage scan provided quantitative and

0-9640024-1-8 $3.00

qualitative signature information about specific gases. The catalytic electrodes contain glass frit to keep the layer porous to promote gas migration to the solid electrolyte, forming the tpb.

The solid-electrolyte layer was an oxygen-ion conductor. The current measured through the solid electrolyte was a direct measure of the amount of oxygen anions at the tpb.

The metal oxide reference layer provides a constant partial pressure of oxygen to the tpb. This metal oxide combination provides an oxygen sink/source reservoir described by the equilibrium equation:

$$Me(solid) + \frac{1}{2}O_2(p(O_2) \Rightarrow MeO(solid) \quad (8)$$

where $p(O_2)$ was the oxygen partial pressure. Equilibrium was achieved when $p(O_2)$ has penetrated the metal and oxide. Homogeneity of $p(O_2)$ is maintained at the electrode/electrolyte interface by oxygen anion migration in the electrolyte and electron activity in the metal phase of the *Me/MeO* system. The electron activity reaction is described by:

$$MeO(solid) + 2e^- \Rightarrow Me(solid) + O^-(electrolyte). \quad (9)$$

In the driven mode of operating the gas sensor (applying a cyclic voltage), the preferred metal oxide reference serves as a constant oxygen source which stabilized the sensor response even in oxygen deficient environments.

CNNs function by interconnecting a number of separate computational "neurons." These artificial neurons are the building blocks for the network. Each neuron is modeled in two stages (Figure 2), the summation stage and the activation stage. The summation stage, quite simply, multiplies each input signal by its associated weight and sums them all together. The activation stage uses a squashing function (input ± ∞, output approximately 0 or 1 only) and a threshold value to determine what the output from the neuron will be. This type of processing effectively makes a small decision on the basis of weighted variables. A collection of these small weighted decisions represents a powerful tool for pattern recognition and decision control.

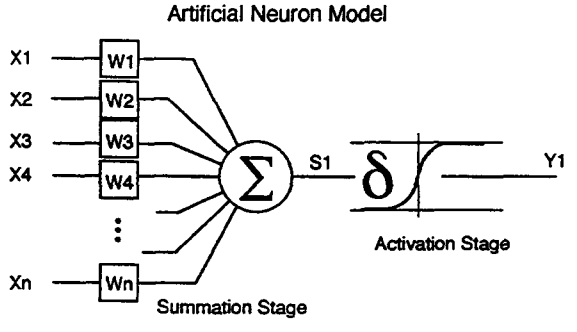


***Figure 2.** Model computational neuron.*

The CNNs can be used simply to recognize and classify the signatures that are sampled, but they also have the capability to approximate the functional response of the sensor signature changes. This can produce an interpolated value attached to the shape change of a signature. Qualitatively, different gases dissociate at characteristic voltages. This manifests itself as a "peak" or "valley" showing up along the voltage axis of a CV signature. An increase in concentration of a given gas increases the probability that that gas will react, and broadening and increasing the peak or valley at the characteristic voltage. The reactions are complex with obvious covariant parameters and many hidden univariant parameters. This complexity normally limits voltammetry to a qualitative measurement technique; it also illustrates the value of being able to identify, and quantify, the univariant parameters with CNNs. The interpolation quality of a network is a function of the completeness of the training and validation sets used. If the signatures for a given gas are smooth and continuous, a relatively simple network and limited set of training samples can produce interpolation errors of only a few percent.

EXPERIMENTAL DETAILS

The sensors were screened onto 0.5 in. x 0.5 in. x .025 in. alumina substrates. The substrates were laser scored into ten 0.1 in. x 0.25 in. individual substrates. After all layers were deposited and fired, the substrates were separated into ten individual devices. This process insured reproducibility from sensor to sensor within each group of ten on a single substrate. NiO was used as the Me-MeO_x reference material. Two oxygen-ion-conducting solid electrolytes, yttria stabilized zirconia (YSZ) (8 mol% Y_2O_3) and tungsten stabilized bismuth oxide (WBO) (WO_3 - Bi_2O_3) (0.25 mol% WO_3), were used. The NiO and solid electrolyte powders were pulverized to pass through a 325 mesh screen and mixed with Heraeus Cermalloy Electronic organic vehicle RV-025 (ethyl cellulose and terpineol) in a 1:3 ratio to form a thick paste. Commercial platinum oxygen sensor paste (Heraeus Cermalloy OS1) was used for the upper and lower catalytic electrodes and the heating element. The prepared platinum paste was thinned with 20 wt% organic vehicle. Each layer was screen-printed separately onto the aluminum oxide substrates by using a Presco™ industrial screen printer with photolithographically stenciled 325 mesh stainless steel screens. The films were fired separately after each screening. The sensors made with the WBO solid electrolyte required a YSZ "buffer" layer between the electrolyte and the upper Pt electrode to keep the Pt from migrating into the WBO layer and shorting the electrodes. The YSZ sensors were fired at 1350°C in air in a controlled cycle to prevent film blistering between screenings. The WO_3-Bi_2O_3 sensors were fired at 850°C in the same way. The film thickness ranged from 3-4 μm for the NiO, 20 μm for the solid electrolyte, and 10 μm for the Pt electrodes and Pt heating element.

The chamber used for gas testing was constructed from a 500-ml bell flask with one sealed spherical end. The stopper that sealed the chamber was channeled, and two, 8-mm stainless steel tubes were inserted to act as feed-throughs for the test and measurement leads. A gap was left in both feed-throughs to allow the incoming gas to escape and to maintain a small positive pressure of the test gas in the chamber so that no ambient air intruded during a test. The sensor was held by a spring-loaded test clip with gold-plated fingers. A micro-mass K-type thermocouple was used to measure the surface temperature of the sensor.

Once the sensor reached the set-point operating temperature in the range of 225°-300°C, the input voltage was applied and the sensor sampled. Three, -1 to +1 to -1 V DC, 30-second long, triangular waveforms were applied to the sensor in succession, and the current passing through the sensor was sampled at approximately 100 points/waveform. These triple (3 X 100 points) voltage-current curves were used to produce the voltammograms that were used in later analyses.

Gas standards of O_2, CO, CO_2, and a variety of hydrocarbons in a balance of N_2 were introduced to the test chamber via a gas-line manifold. A precision flowmeter was used to control the gas entering the chamber in two flushing stages. For gas mixture tests, a two-line gas proportioner was used. In the first stage, the test chamber was flushed with approximately 50 chamber volumes at approximately 5 L/min. for 5 min. In the second stage, the flow rate was lowered to 50 ml/min. to create a stable environment and allow the sensor to attain a predetermined operating temperature. This flow rate was maintained during sensor sampling. Temperature dependence of the sensor was studied by stabilizing the sensor in air and increasing the temperature by 25° increments in the range of 175° - 325°C before sampling.

The neural signal processing of the voltammograms involved introducing the 1-D sampled array to a neural algorithm and producing a resultant value that was an measurement of a particular gas in a mixture. The neural network was developed using NASA NETS3.0 software. The network was "trained" by presenting it with successive sets of known input signatures and algorithmically adjusting the network interconnections until the produced error was minimized. Typical training sets consisted of (3) samples each of (6) gas concentrations. Two of the three samples were presented to the network as training vectors, with the third sample being used for validation. This process typically produced errors of < 5% full scale value across the concentration range.

RESULTS

The dependence of the YSZ sensor on temperature is shown in Figure 3. As the temperature increases, the amplitude and sharpness of the peak current increase with the conductivity of the electrolyte.

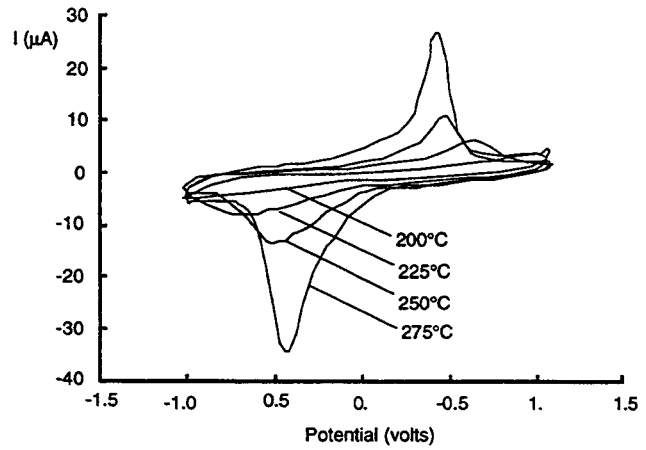


Figure 3. Sensor (YSZ electrolyte) response to temperature.

A plot of peak current vs. temperature shows a similar relationship to the typical conductivity dependence for an oxygen-ion-conducting solid-electrolyte (Figure 4). The conducting range for the driven solid-electrolyte cell is in a much lower temperature range (200-300°C) as compared to a typical non-driven solid electrolyte cell (500-1100°C). The lower temperature range is possible because of the presence of the applied voltage, which forces the ions through the electrolyte, as opposed to a strictly diffusion-controlled process in the non-driven cell.

Figure 5 shows a YSZ sensor response at 250°C to various concentrations of O_2 in a balance of N_2. The anodic and cathodic peaks increase with oxygen concentration. Since response is directly related to the amount of oxygen available at the tpb, increased peak magnitude was expected as higher oxygen concentrations directly supply more oxygen for transport through the solid electrolyte. The small O_2 peaks in the 100% N_2 voltammogram are due to the partial pressure of oxygen supplied to the tpb by the metal oxide reference layer.

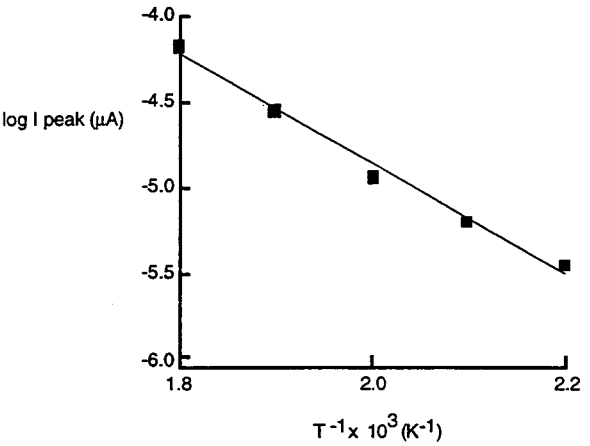


Figure 4. Log of peak current vs. inverse temperature.

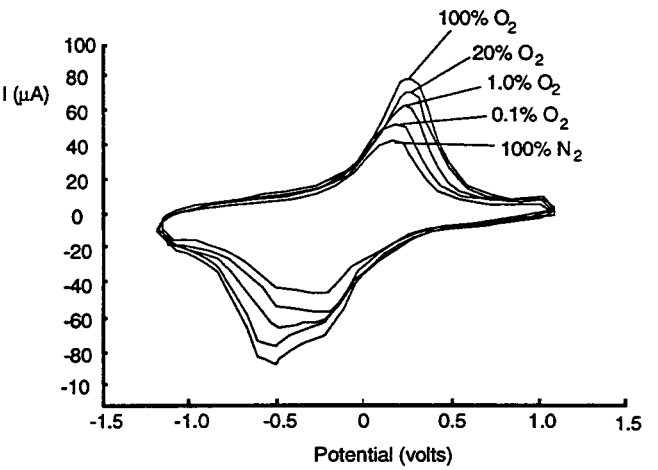


Figure 5. YSZ sensor response to oxygen.

Figure 6 shows a WBO sensor response at 250°C to various O_2 concentrations. The WBO sensor shows a greater sensitivity to change in O_2 concentrations than did the YSZ sensor. At higher O_2 concentrations, the voltammogram changes shape as the current increases significantly. The sensitivity to change in concentration does not increase with temperature for either sensor type. Although the magnitude of the peak current increased with temperature, the relative change in peak current with gas concentration did not increase with temperature:

$$\frac{\partial\left(\frac{dI}{d[c]}\right)}{\partial T} = 0 . \qquad (3)$$

The sensors appear to operate best in a window of about 225°-300°C. Above the "ideal" temperature range, the sensitivity to concentration change decreased.

Figure 7 shows the WBO sensor response to various concentrations of CO_2 in a balance of N_2. As expected, the voltammogram for CO_2 in N_2 was significantly different than that of O_2 in N_2. A unique CO_2 peak was formed that increased with concentration.

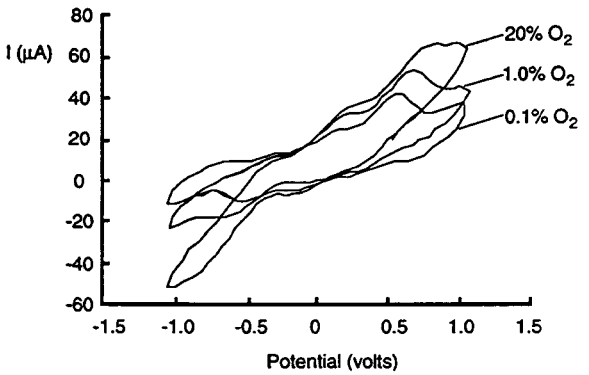


Figure 6. WBO sensor response to oxygen.

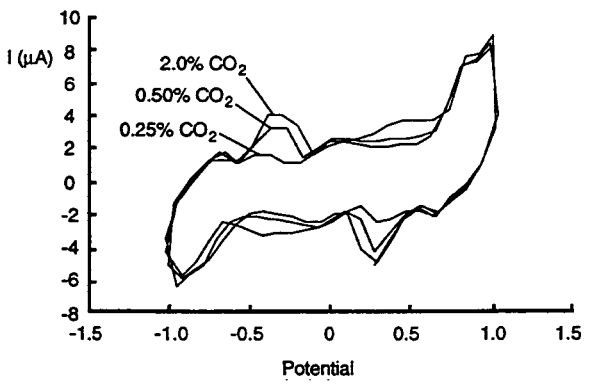


Figure 7. WBO sensor response to carbon dioxide.

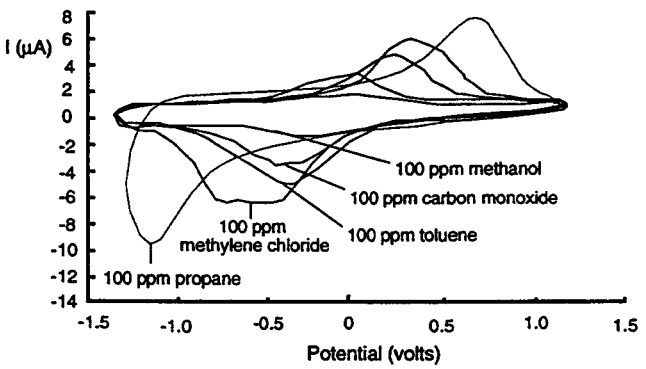


Figure 8. YSZ sensor response to hydrocarbons.

This response implied that catalytic CO_2 reactions were occurring with a subsequent increase in O^- ionic current through the bismuth oxide solid electrolyte as described by equation (6).

The YSZ sensor response to a variety of hydrocarbons is represented by the voltammograms in Figure 8. Between testing the gases, a standard of 20% O_2 in N_2 was run to bring the sensor back to an equilibrated state so that mixed responses were avoided. The voltammogram for each gas at 100 ppm was distinct. A change in size and shape of the voltammograms was exhibited for each gas tested.

The typical analog CV sweep was 30 seconds long and subsequent voltammograms decreased in peak measured current when this was shortened. The reactions observable from the analog CV were clear and illustrated the reproducibility of the applied CV technique. The first discrete CV waveforms, used for the neural signal processing, modeled the analog and used a large number of steps and similar 30 s sweeps. Later changes were made to reduce the number of steps and shorten the sweep duration to only a few seconds.

The CNN algorithms allowed training of simple 24 input-5 hidden-1 output node networks in about (10) minutes. The networks were typically trained to ±5% FSV error and produced acceptable error from the validation vectors. The network algorithms for all the gases in this work were identical. They differed only by a table of interconnection weight values. This allowed very generic and optimized algorithms to be written for "gas analysis", with only table values changed to alter what gases were being analyzed "for". These algorithms were coded and implemented on microcomputers and single-chip microcontrollers.

CONCLUSIONS

Cyclic voltammetry applied to a thick-film electrochemical cell using cermet materials were used as a new class of active gas microsensors. The microsensors detected a variety of gases, including O_2, CO_2, CO, and hydrocarbons, and showed potential to distinguish individual gases in a mixture. The sensors exhibited a temperature dependence similar to that of typical oxygen-ion-conducting solid-electrolytes but at significantly lower temperatures due to the applied voltage. Advantages of thick-film cermet sensors include low cost, small size, versatility, and high-integrity materials and construction. The rugged materials allows them to withstand considerable mechanical, chemical, and thermal stress. Thick-film technology provides simple, low-cost fabrication without the need for exotic equipment or production methods. The ability of the CNNs to identify gas signatures without dissecting them gave the system the ability to be *trained* to recognize thresholds of simple mixtures, and the potential existed for recognition of very complex mixtures. This capability made the sensors attractive for applications that require multi-gas/mixture monitoring. Areas of future development include reducing the sensors' power consumption, improving their sensitivity, further improving the signal processing, and miniaturizing the support electronics.

REFERENCES

1. C.G. Vayenas, A. Ioannides, S. Bebelis, "Solid Electrolyte Cyclic Voltammetry for in Situ Investigation of Catalyst Surfaces", *Journal of Catalysts*, 129, 67-87 (1991).

2. K. Huang, S. Tang, and Q. Liu, "The diffusitivity and solubility of oxygen in reference electrodes Ni-NiO, Co-CoO, Fe-FeO, and MO-MoO_2 systems," *Journal of Solid State Ionics* 61, 355 (1993).

LOW REYNOLDS NUMBER MICRO-FLUIDIC DEVICES

James P. Brody and Paul Yager
Center for Bioengineering, University of Washington
Seattle, WA 98109-2141

ABSTRACT

The fluid dynamics in microfabricated fluidic devices is inherently in the low Reynolds number regime. By recognizing this, efficient micro-devices can be designed. Two such devices are described, a diffusion-based filter and a virtual valve. The diffusion-based filter differentially separates particles and molecules based on diffusion coefficient while diluting their concentration. Results and a simple analysis indicate that the dilution of the particles is exponentially related to the diffusion coefficient. The virtual valve functions as a switch, directing flow at the intersection of three channels by manipulating the pressure driving the flow. Switching times as fast as 100 ms are demonstrated. These devices are unique to the micro-scale—they would not work as well if scaled up.

INTRODUCTION

It has recently become possible to fabricate intricate fluid handling systems with channel sizes as small as 1 μm using microfabrication techniques. These devices can be mass produced inexpensively and some may soon be in widespread use for simple analytical tests [1, 2]. There are many advantages these miniaturized devices could offer, including the promise of performing some analytical functions more rapidly and on smaller samples than with macroscopic systems.

However, miniaturization of analytic instruments is not a simple matter of reducing their size. At small scales different effects become important, rendering some processes inefficient and others useless. It may also be difficult to replicate smaller versions of some devices because of material or process limitations. For these reasons it may prove necessary to develop new methods for performing common laboratory tasks on the micro-scale.

This paper presents two individual micro-devices which function as the equivalent of macroscopic-scale filters and valves, but which operate on entirely different principles. These devices would be inefficient on the macroscopic scale, but function quite well on the microscopic scale.

These devices are able to function because of several important physical principles. Typical micro-fluidic devices operate in the low Reynolds number regime where turbulent mixing and inertia are absent. Mixing can only be done by diffusion,j and diffusion of small molecules occurs rapidly over distances of 1–10 μm.

FLUID DYNAMICS

Microfabricated fluid systems often operate under low Reynolds number conditions. The Navier-Stokes equations governing fluid flow can be written as

$$-\nabla P + \eta \nabla^2 u = \rho \left(\frac{\partial u}{\partial t} + (u \cdot \nabla) u \right), \qquad (1)$$

where P is the applied pressure, u is the velocity vector, ρ is the fluid density and η is the viscosity of the fluid. The left side of this equation represents the forces on the fluid due to applied pressure and viscosity and the right hand side represents "inertial forces". A useful quantity is the ratio of the inertial forces to the viscous force. The magnitude of this ratio,

$$R = \frac{\rho \left(\frac{\partial u}{\partial t} + (u \cdot \nabla) u \right)}{\eta \nabla^2 u} = \frac{\rho r^2}{\eta \tau} + \frac{\rho u r}{\eta}, \qquad (2)$$

is written in terms of the characteristic size (typically the radius) of the channel, r, and the time scale over which the velocity is changing, τ (where $u/\tau = \partial u/\partial t$.)

Fluid flow behavior in the steady state, $\tau \rightarrow \infty$, is characterized by the Reynolds number, $R_e = \rho u r/\eta$. Because of the small sizes and slow velocities, microfabricated fluid systems are often in the low Reynolds number regime ($R_e \ll 1$). In this regime, inertial effects, which cause turbulence and secondary flows, are negligible; viscous effects dominate the dynamics. (See [3] for a good introduction to low Reynolds number fluid dynamics.) Because of the absence of inertial effects, at low Reynolds number all mixing is done by diffusion.

DIFFUSION

Diffusion of small molecules occurs radiply over typical microfabricated dimensions. The relationship between the size of a particle, R_o, the diffusion coefficient, D, and temperature, T, is due to Einstein and for the simplest case, spherical particles, this can be written as[4],

$$D = \frac{k_b T}{6 \pi \eta R_o}. \qquad (3)$$

The characteristic distance, l, which a particle with diffusion coefficient D will diffuse in time, t, is $l = \sqrt{Dt}$. Table 1 gives some typical diffusion coefficients and characteristic times.

Table 1: *Some typical values for different sized particles and molecules. The characteristic time to diffuse 10 μm is given.*

Particle	D (20° C)	t
0.5 μm sphere	5×10^{-9} cm^2/sec	200 sec
Protein (hemoglobin)	7×10^{-7} cm^2/sec	1 sec
Small Molecule (fluorescein)	5×10^{-6} cm^2/sec	0.2 sec

SILICON PROCESSING

Silicon micromachining was done in the Washington Technology Center microfabrication facility. The process was modified from previous work [5] to avoid the problem of coating deep features uniformly with photoresist. In our experience, spin coating photoresist on a wafer that had features greater than a few microns deep led to a non-uniform coating that either left undeveloped areas after exposure or areas uncoated with photoresist. These problems were avoided with this process.

This process required two mask levels. One level defined connection ports, which were etched completely through the wafer, and the second level defined the fluid transport channels on the face of the wafer, typically only 10 μm deep.

0-9640024-1-8 $3.00 © 1996 TRF

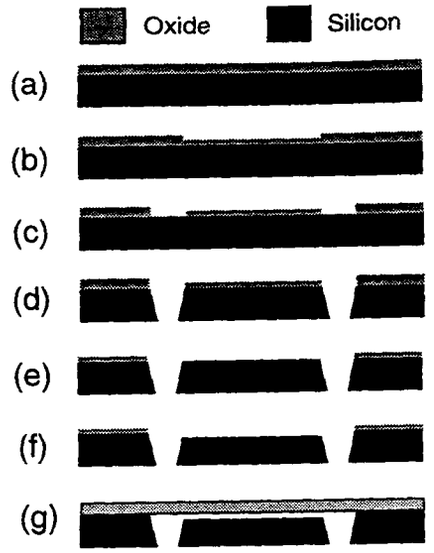


Figure 1: *A schematic diagram of the silicon processing. (a) A 1 μm thick wet thermal oxide is grown on a 3" silicon wafer. (b) This oxide is photolithographically patterned with the flow channels and etched to a depth of 600 nm. (c) The wafer is recoated with photoresist and patterned with the through-hole connections. This oxide is completely removed from this pattern. (d) EDP etching is done to etch completely through the wafer (approximately 400 μm). (e) An oxide etch is performed to uniformly remove 400 nm of oxide from the wafer. (f) The flow channels are etched into the silicon approximately 10 μm deep. (g) The wafer is anodically bonded to a 3" disk of Pyrex glass.*

We started with 4" chrome masks made to our specifications by Photo Sciences, Inc. (Torrance, CA) and 3" silicon wafers (100, n-type) with 1000 nm of wet thermal SiO_2 grown on them. A primer (P-10M, MicroSi, Phoenix, AZ) was spun on at 3000 rpm. About one micron of AZ-1512 (Hoechst) photoresist was deposited by spin coating (3000 rpm), and this was followed by a soft bake (30 min at 90 C).

A contact aligner was used to align and expose wafers. Exposure time (nominally 15 seconds) was varied to yield best results. No post-exposure bake was done. Wafers were developed in AZ-351 (Hoechst) (diluted 4:1) for one minute, and rinsed in DI water. Blue tack tape (Semiconductor Equipment Corporation, Moorpark, CA) was applied to the backsides of the wafers to protect the oxide from the oxide etch.

The fluid channel mask was exposed first. After developing, the wafers were immersed in a buffered oxide etch (BOE, 10:1 HF (49%) and NH_4F (10%)) to etch away 600 nm of the unprotected oxide. The blue tack tape was removed by hand, and the photoresist was removed in an acetone rinse.

Wafers were again coated with photoresist and the fluid connection mask was aligned and exposed on the wafers. These wafers were developed and blue tack tape was once again applied to the backsides of the wafers to protect the oxide, and then immersed in the BOE to remove all the unprotected oxide. Blue tack tape and photoresist were removed as before.

Silicon etching was done in a mixture of ethylene-diamine, pyrocatechol, and water (EDP F-etch) set up in a reflux boiling flask. This etch attacks the {100}-planes of silicon at a rate of about 100 μm per hour[6]. First fluid connection ports were etched through the wafer, then 400 nm of the oxide were removed in a BOE. Finally, flow channels between fluid connection ports were etched about 10 μm deep. After final processing the wafers were cleaned in a Piranha bath and rinsed in DI water.

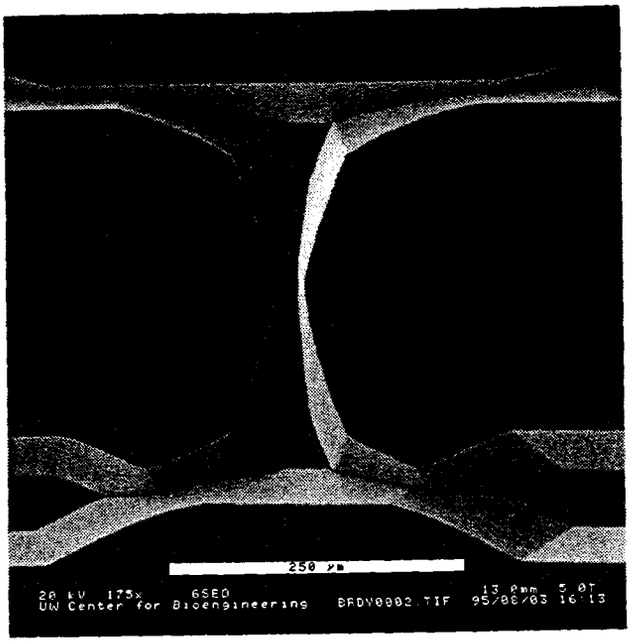


Figure 2: *A scanning electron micro graph showing the central portion of a diffusion-based filtration device.*

We used anodic bonding [7] to attach Pyrex glass to the silicon devices. We obtained 3" diameter pieces of Pyrex glass (100 μm thickness) from Esco Products Inc (Oak Ridge, New Jersey). First, the silicon and Pyrex glass were immersed in a solution of H_2O_2, NH_4OH, and H_2O (1:4:6) heated to 50 C. This process removes any organic matter on the surfaces and also makes the surfaces hydrophilic. After 20 minutes in this solution, the silicon and Pyrex were placed in a rinser-dryer. Upon removing, the Pyrex glass was immediately placed into contact with the silicon wafer and set on a hot plate for anodic bonding. Bonding was done at 485 C with 800 V applied between the Pyrex and the silicon.

The Pyrex-silicon structures were diced into individual 1 cm by 1 cm devices by manually marking the silicon with a diamond tipped scribe and gently applying stress along the scratch. The Pyrex-silicon fractured along the silicon fault without problems.

EXPERIMENTAL SETUP

Flow visualization was done by flowing dilute suspensions of fluorescent micro-spheres through the device. We added one drop of 0.5 μm diameter fluorescent micro-spheres (1% solids, from Duke Scientific, Palo Alto, CA) to 5 ml of DI water.

This mixture was introduced into the device, and observations were made using a Zeiss ICM-405 inverted microscope equipped with an XF34 fluorescent set (535 nm excitation, 590 nm emission) from Omega Optical (Burlington VT). We observed the fluorescence of the spheres using both a silicon intensified target camera (SIT-66x, from Dage-MTI) and a Pulnix monochrome CCD video camera. We recorded the observations on video tape. Some images were also digitized using a frame grabber (Data Translation or Power Macintosh 7100AV).

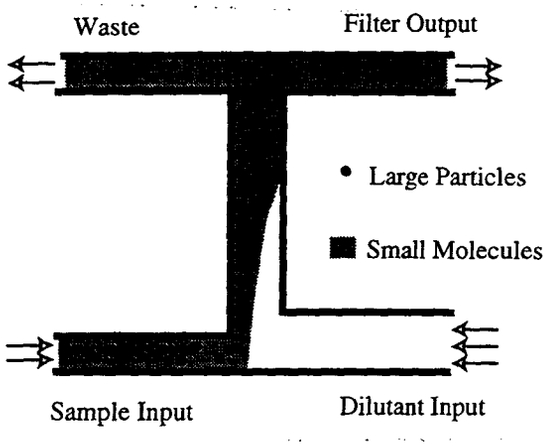


Figure 3: *The concept of diffusion-based filtration. Two flow streams are brought into contact. Because the flow is at low Reynolds number conditions, no mixing occurs at the intersection. As the two flow streams move down the central channel, some molecules diffuse across. At the end of the central channel, the flow is again split (not necessarily at the same place) into two streams. The output is preferentially enriched with smaller molecules over larger ones.*

DIFFUSION BASED SEPARATION

The concept of diffusion-based separation is described in Figure 3. The process relies on the non-mixing nature of low Reynolds number flow along with the fast diffusion times for small molecules. The idea was originally implemented in a system with two glass plates separated by a shim by Giddings[8]. We present here a greatly modified version of this implemented in a microfabricated fluid system.

A detailed analysis will be published elsewhere, but the main result is that a molecule (with diffusion coefficient, D) present in the input stream at a concentration C_o will be present in the output stream at a concentration C,

$$C \sim C_o \exp(-D_f/D), \tag{4}$$

where the parameter, D_f, quantifies the device, $D_f = \lambda_b^2/t_R$. The time a particle spends in the central channel is t_R, and the shortest distance a particle must diffuse to go from the input stream to the output stream is λ_b (the distance between the inlet splitting plane and the outlet splitting plane in [8].) The resolution of the separation is set by the minimum concentration that can be detected.

The pressures at the sample input port and at the dilutant port were set by filling a tube which fed the ports to the appropriate height with liquid. Typically a pressure difference of about 2 cm of water gave a maximum velocity in the central channel of about 100 μm/sec. The waste and filter output ports were held constant. Measurements were made with two different devices (with different central channel lengths) and with varying flow rates to vary the parameters λ_b and t_R and hence D_f. The data roughly follow the predicted behavior.

VIRTUAL VALVES

Virtual valves have been demonstrated in electrokinetically driven flow[9], but apparently not in pressure driven flow. The same principles apply in each case—by controlling the pressure (as opposed to the voltage for electrokinetics) at each port of a "T" shaped device, the flow can be directed one way or another at the intersection of channels.

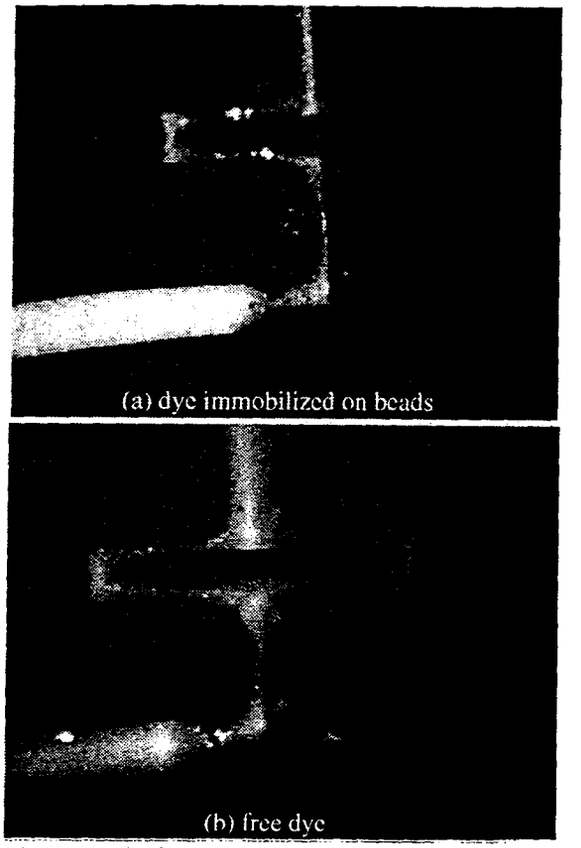


Figure 4: *Digitized video images from the diffusion-based filter. As in Figure 3, flow is from the lower left and right up through the central channel and splits into the waste (upper left) and output (upper right). In (a) the fluorescent dye is immobilized on 0.5 μm spheres which diffuse too slowly to end up in the output stream in any significant concentration. In (b), the dye (fluorescein), is free to diffuse, and it readily diffuses throughout the central channel. In this device the output and waste channel are tied together to simplify the pressure control.*

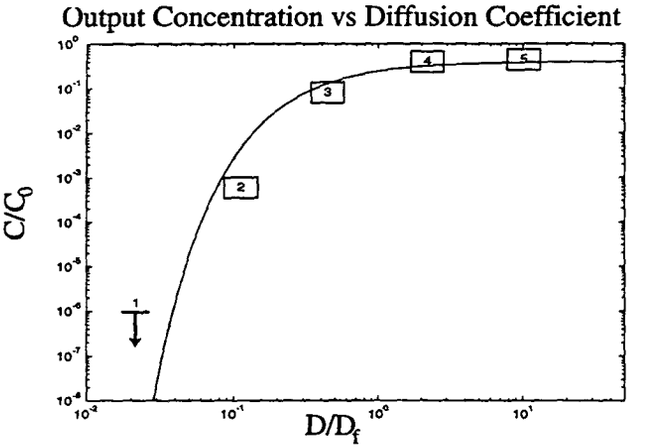


Figure 5: *Data from the diffusion-based filter. The ratio of concentration in the output to concentration in the input (C/C_o) is plotted versus the diffusion coefficient of the molecule of interest (D). (The diffusion coefficient is measured in terms of the device parameter, D_f, which is a function of flow rate, central channel length, and diffusion distance, see text.)) Concentrations were measured by either counting fluorescent spheres or measuring fluorescence intensity of fluorescein with a photo diode. The solid line is the expected behavior from equation (4). Errors are or order the sizes of the boxes.*

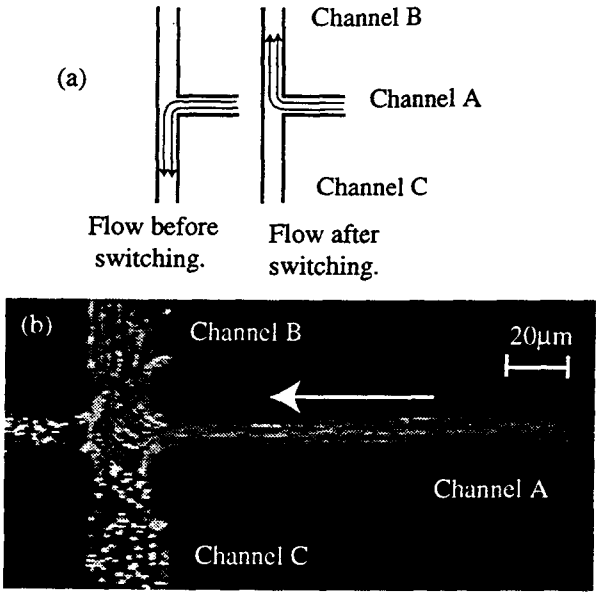


Figure 6: *A virtual valve. (a) The flow at the intersection of three channels can be directed by controlling the pressures behind Channels A, B, and C. (b) A fluorescent picture of the virtual valve in action. Flow is from right to left in channel A and up channel B. Note that the fluid is stationary in channel C.*

The fastest rate at which one can expect to switch the flow without significant mixing is given by the requirement that $R < 1$ from equation (2). Faster switching may be possible with careful control of the geometry, but as long as $R < 1$ the geometry is irrelevant. Table 2 gives some typical examples for water in different sized channels.

Table 2: *Approximate fastest switching allowed while still maintaining low Reynolds number conditions. (Time given for water in different channel diameters.*

Diameter of Channel	Switching time
10 μm	100 μs
100 μm	10 ms
1 mm	1 s

These experiments were performed by driving the fluid through 10 μm deep channels with a pressure of about 1 PSI over a total length of 1 cm. The pressure was provided from a compressed nitrogen cylinder. A small pressure regulator was built which allowed control over pressure at two of the ports (the third was allowed to float at atmospheric pressure.) A valve was installed between the pressure regulator and the device which allowed rapid switching between the ports. For instance as in figure 6a, before switching the relative pressures were Channel A, 1 PSI; Channel B, 0.5 PSI; Channel C, 0 PSI. After switching, they were Channel A, 1 PSI; Channel B, 0 PSI; Channel C, 0.5 PSI.

CONCLUSIONS

Micro-fluidic devices operate in the low Reynolds number regime. The fluid dynamics in this regime is dominated by viscous forces, rather than inertial ones. Because of this, micro-fluidic devices are inherently different than macroscopic ones. By understanding the fluid dynamics and diffusion processes which occur at the micro-scale, efficient and novel devices can be designed to perform common laboratory functions needed in any lab-on-a-chip.

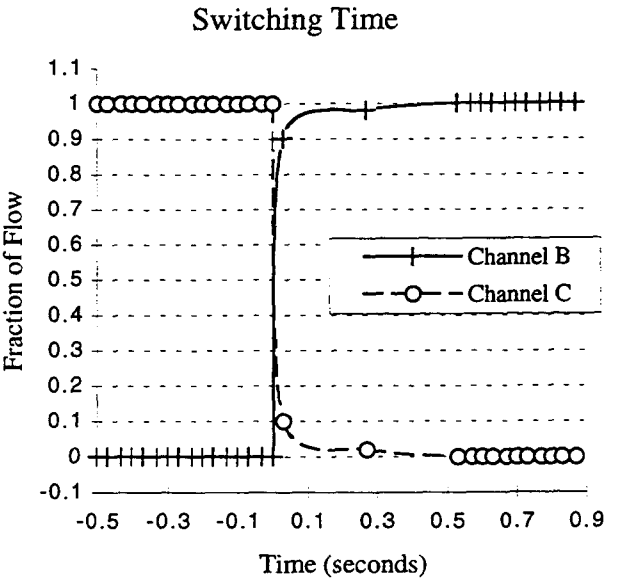


Figure 7: *Data from the switching of the virtual valve. Video images were analyzed frame-by-frame and the relative flow rates in two channels were recorded. The switching is completely done in ~100 ms. An analysis indicates that this could be done as quickly as 100 μs in a 10 μm deep channel. In our experiment, the switching time is limited by our pressure control system.*

ACKNOWLEDGMENTS

We thank H. Sho Fuji and D. Schutte for expert advice on microfabrication and M. Holl for help with the pressure control system. This work was supported by the Washington Technology Center, ARPA (grant # DAMD17-94-J-4460) and Senmed Medical Ventures, Inc.

REFERENCES

[1] J. Michael Ramsey, Stephen C. Jacobson, and Michael R. Knapp. Microfabricated chemical measurement systems. *Nature Medicine*, 1(10):1093–1096, 1995.

[2] D. Jed Harrison, Karl Fluri, Kurt Seiler, Zhonghui Fan, Carlos S. Effenhauser, and Andreas Manz. Micromachining a miniaturized capillary electrophoresis-based chemical analysis system on a chip. *Science*, 261:895–897, August 1993.

[3] E. M. Purcell. Life at low Reynolds number. *American Journal of Physics*, 45(1):3–11, 1977.

[4] E. L. Cussler. *Diffusion: Mass Transfer in Fluid Systems*. Cambridge University Press, 1984.

[5] James P Brody, Thor Osborn, Fred K Forster, and Paul Yager. A planar microfabricated fluid filter. In *Transducers '95*, 1995.

[6] Kurt E. Petersen. Silicon as a mechanical material. *Proceedings of the IEEE*, 70(5):420–457, 1982.

[7] George Wallis and Daniel I. Pomerantz. Field assisted glass-metal sealing. *Journal of Applied Physics*, 40:3946–3949, 1969.

[8] Shulamit Levin and J Calvin Giddings. Continuous separation of particles from macromolecules in split-flow thin (SPLITT) cells. *Journal of Chemical Technology in Biotechnology*, 50:43–56, 1991.

[9] Andreas Manz, Carlo S Effenhauser, Norbert Burggraf, D Jed Harrison, Kurt Seiler, and Karl Fluri. Electroosmotic pumping and electrophoretic separations for miniaturized chemical analysis systems. *J. Micromech. Microeng*, 4:257–265, 1994.

LINEAR VIBROMOTOR-ACTUATED MICROMACHINED MICROREFLECTOR FOR INTEGRATED OPTICAL SYSTEMS

Michael J. Daneman, Norman C. Tien, Olav Solgaard, Kam Y. Lau, Richard S. Muller
University of California at Berkeley
211-70 Cory Hall, #1772
Berkeley, CA 94720-1772

ABSTRACT

We present a linear vibromotor-driven micromachined microreflector designed for integrated microphotonic applications. This device has two degrees of freedom, with an angular travel range of over 90° and a translational travel range of 60 μm. The precision and speed of the microreflector are characterized, and it's application to beam steering and fiber coupling systems is demonstrated.

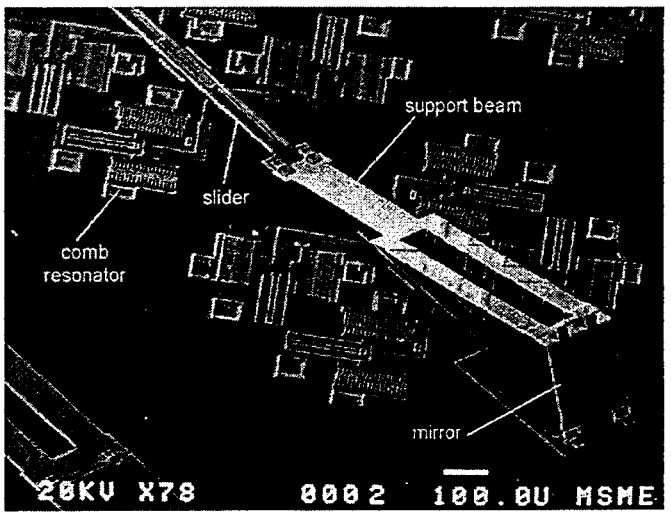
INTRODUCTION

The growing optical communications and measurement market requires low-cost, high-performance optoelectronic modules such as laser-to-fiber couplers, tunable lasers, scanners, interferometers, etc. Unlike integrated electronic circuits, integrated optical systems require precise alignment of components. A misalignment of less than 1 μm can dramatically decrease system performance. While silicon-optical-bench (SOB) technology provides for hybrid integration of semiconductor lasers, lenses, and optical fibers on a silicon chip, it is typically limited to ±1 μm alignment tolerances (without external adjustment).. Another limitation is the lack of on-chip actuated optical components (such as mirrors, gratings, lenses, etc.) As a result, SOB technology applications have been limited to simple systems with no more than three or four components. To overcome these limitations a micromachined free-space optical platform has been proposed [1,2]. By combining micromachined movable optical components with lasers, lenses, and fibers on a silicon substrate, we can produce complex self-aligning optical systems on a chip. In order to function in a self-contained optical module, the micromachined components must have sufficient precision and, where required, must have on-chip actuation. In this paper, we present an actuated micromachined microreflector with two degrees of freedom and on-chip resonant electrostatically driven impact actuators.

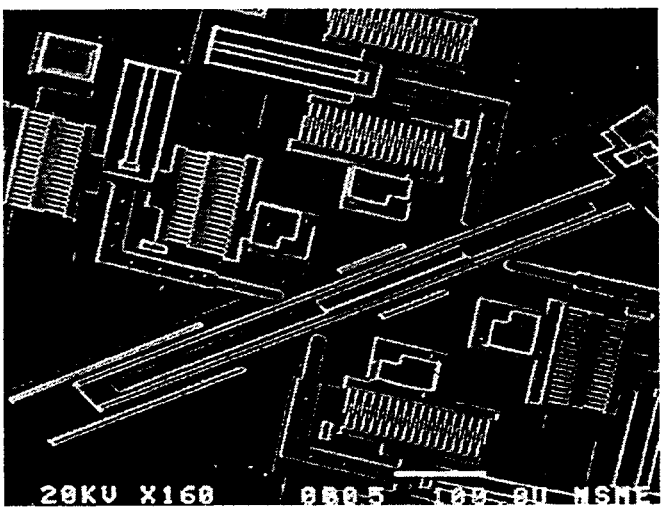
DESIGN

The actuated microreflector (Figure 1a) was designed for laser-to-fiber coupling and external-cavity-laser applications. It consists of a polysilicon mirror plate hinged to a support beam. Each of these is, in turn, hinged to a vibromotor-actuated slider. Microhinge technology [3] allows the joints to rotate out of the plane of the substrate to achieve large aspect ratios. Common-mode actuation of the sliders results in translational motion, while differential slider motion produces out-of-plane mirror rotation. These motions permit the microreflector to position an optical beam in two dimensions.

Each of the two sliders is actuated with an integrated microvibromotor (Figure 1b). The vibromotor consists of four electrostatic-comb resonators with attached impact arms driving a slider through oblique impact. To balance the forces, two opposing impacters are used for each direction of travel. The resonator is a capacitively driven mass anchored to the substrate through a folded beam flexure. The spring constant of the flexure determines the resonant frequency and travel range of the resonator. The force exerted by the comb drive is proportional to the square of the applied voltage



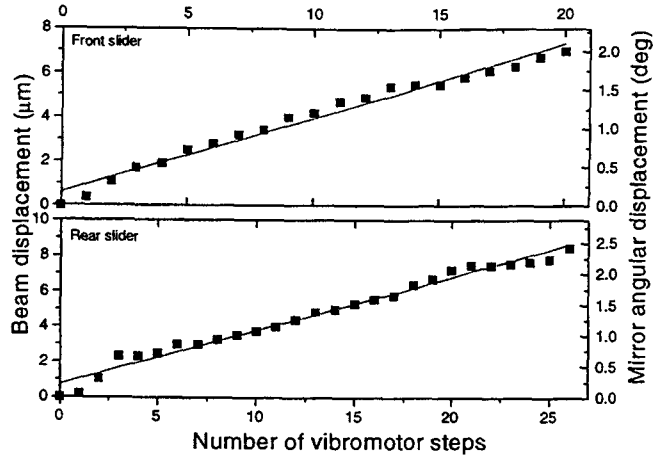
(a)



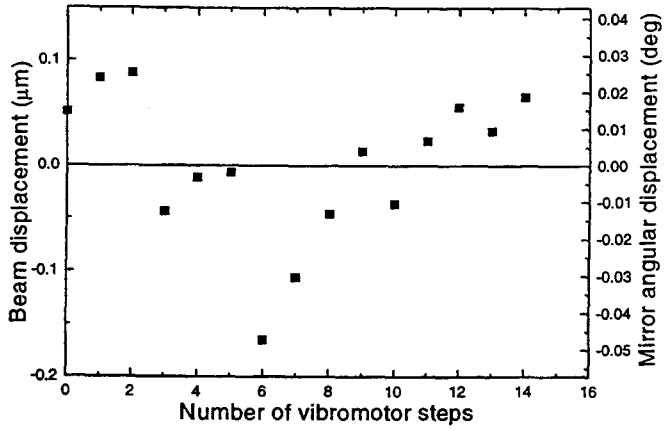
(b)

Figure 1. (a) Scanning electron micrograph of a released and assembled actuated microreflector. (b) A close up view of the integrated vibromotor actuator.

0-9640024-1-8 $3.00 © 1996 TRF



(a)



(b)

Figure 2. Vertical (a) and horizontal (b) microreflector precision. (a) The vertical position data show standard deviations of 0.32 μm and 0.6 μm (for the front and rear sliders respectively) from a linear response. (b) The horizontal beam position has an average 0.05 μm of in-plane wobble.

$$F = \frac{1}{2}\frac{\partial C}{\partial x}V^2$$

where V typically has both dc and ac components. This quadratic response produces a primary frequency driving term proportional to the product of the dc and ac voltages, effectively linearizing the resonator and increasing the impact force. The comb structures are driven at their resonance frequency (roughly 7.5-8.5 kHz) thereby achieving an amplification of the electrostatic force by the resonator quality factor (typically 30 to 100 in air [4,5]). Since energy is transferred to the slider only during impact (typically lasting only a few microseconds), the impacters can deliver short-duration forces that are large enough to overcome static friction in the sliders and hinges. Due to the damping (primarily due to airflow [4,5]) in the comb structure, the resonators require a few initial cycles to build up sufficient amplitude and momentum for impact. Therefore, in air, slider motion is observed only after three or more voltage cycles. The number of initiation cycles depends on the ambient atmosphere, and will decrease at lower pressures. When driven with a free-running resonant oscillation, the slider reaches a maximum

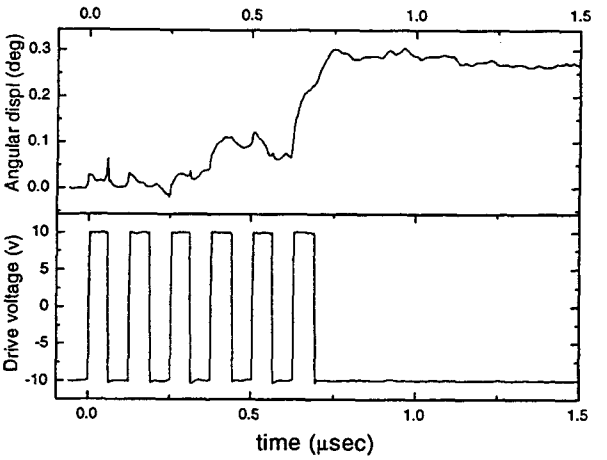


Figure 3. Time domain response of the microreflector to a square wave drive voltage. An impact is first achieved during the third cycle at which point the mirror angle begins to change. The motion concludes shortly after the last cycle resulting in a total angular displacement of about 0.3°.

velocity of over 1 mm/s. Slider velocity can be controlled by driving the comb drives with gated bursts of 4 to 5 cycles of the resonant waveform. Once the slider is in position, it is kept in place by static friction until further actuation. Similar structures were previously subjected to shock and vibration tests and showed no detectable slider motion at forces up to 500 G's [1].

FABRICATION

The actuated microreflector was fabricated on a silicon substrate using silicon-surface micromachining technology. The fabrication process is described in detail in [1,8]. An n+ polysilicon layer defines a ground plane. Three additional polysilicon structural layers are used to define the comb-drives, sliders, hinges, and mirror beams. Phosphorous doped silicon dioxide is used for sacrificial spacer layers between polysilicon layers. A special pre-release etch in 5:1 HF followed by a vigorous rinse is used to eliminate stringers. Then the structure is released for 10 minutes in concentrated HF to dissolve the oxide and dried using a critical-point CO_2 drier [7] to avoid stiction. A 400 Å gold layer is evaporated onto mirror surface to increase reflectivity.

CHARACTERIZATION

Characterization of the vibromotor alone has been described in a previous publication and has shown the slider motion to have a step resolution of less than 0.3 μm [8]. In this paper, vibromotor performance in an integrated microreflector system (Figure 1) is characterized. To measure the precision of the actuated microreflector, a HeNe laser is reflected from the micromirror surface onto a CCD camera. As the front and rear sliders are actuated, the beam position along each of the two axes is measured on the CCD and extrapolated to a location 200 μm in front of the mirror (where a fiber would typically be positioned for a laser-to-fiber coupling application). The standard deviation of the vertical beam position data from a linear response depends on the selected slider step size. For the front slider this deviation is roughly equal to the step size, while for the rear slider the deviation exceeds the step size by about 50-60%. The selected average step sizes of 0.35 μm for the front slider and 0.42 μm for

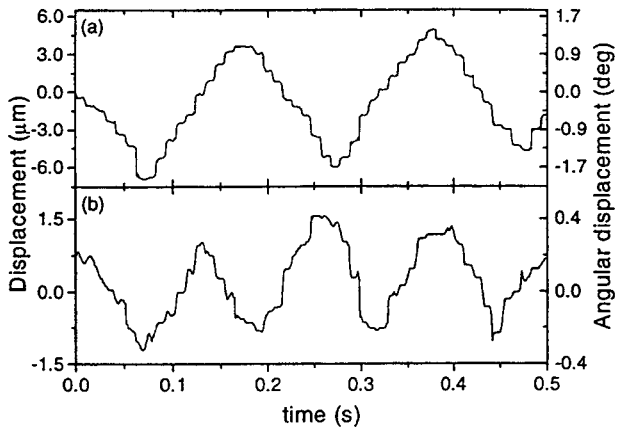


Figure 4. Beam scanning with the actuated micro-reflector. In (a) a 40 V dc offset was used, resulting in 0.7-1.1 μm step size on the plane 200 μm from the micromirror, while in (b) a 38 V dc offset produces 0.3-0.6 μm steps.

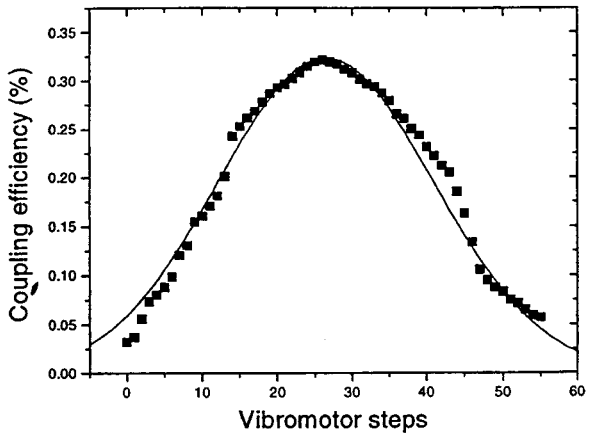


Figure 5. Coupling efficiency from semiconductor laser to single-mode fiber using the actuated micromirror. An efficiency of 32% was achieved.

the rear slider produce standard deviations of 0.32 μm and 0.60 μm respectively (Figure 2a). These deviations are due primarily to the "play" in the hinges and the wobble in the slider structure. The greater length of the rear slider results in increased wobble and leads to a greater standard deviation. The horizontal beam deviation is 0.05 μm (Figure 2b) and is comparable to the 0.07 μm deviation measured in externally actuated structures. This precision is sufficient for laser to single-mode fiber coupling applications where, due to lens magnification, the beam only needs to be within approximately ±1 μm for high coupling. However, for external cavity laser applications a higher angular precision is necessary. Since in earlier experiments a microreflector with no on-chip actuators and an alternate hub design has shown a vertical precision of 0.17 μm [1], improving the design of the actuated slider should greatly improve the mirror precision.

In order to observe the actuation dynamics of the microreflector, a reflected HeNe beam was imaged on a position-sensitive detector while the angular position of the mirror was swept in real time. The rear vibromotor was biased at 40 V dc and driven with 20 V (p-p) resonant (8.2 kHz) square-wave pulses. The position of the beam and the drive voltage were then monitored on a digital oscilloscope. The resulting trace is shown in Figure 3. As predicted and observed in [8], the comb structures require a few cycles (3 in this case) to build up sufficient energy for impact. The first significant impact occurs during the third cycle at which point the mirror angle begins to change. The mirror is further deflected during the fourth cycle; however, sufficient energy is lost by the resonator to keep the next impact from producing significant displacement. Finally, the sixth cycle provides the greatest impact, resulting in a total mirror rotation of 0.3°. The small back and forth motion apparent in the response is most likely due to slight deformation of the hinge joints during impact as the square pin is forced against the staple. The remaining "roughness" of the response is due to noise in the detector.

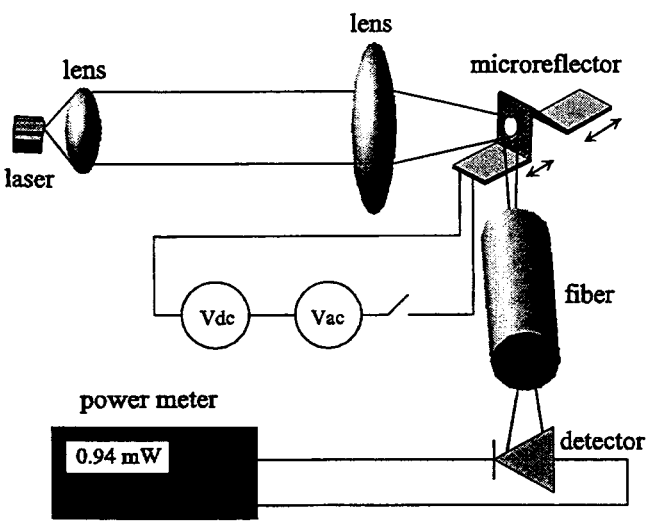


Figure 6. Experimental setup for the laser-to-fiber coupler system. The microreflector is used to optimize optical alignment into the single-mode fiber.

To demonstrate the use of the actuated microreflector in scanning and beam-positioning applications, the mirror was used to continuously scan a laser beam across the detector. The rear vibromotor was driven with bursts of four 20 V (p-p) resonant (8.2 kHz) square-wave cycles spaced 10 ms apart, with a 40 or 38 V dc offset. The resulting output (Figure 4) clearly shows the stepwise nature of the mirror motion. The average speed of the sweep can be changed by varying the spacing of the bursts. Figures 5a and 5b also demonstrate that the size of the step itself depends on the applied voltage and can be controlled. With an average angular step size of 5 mrad, the microreflector has a maximum estimated scan rate of 10.2 rad/s or a beam speed of 2 mm/s on a plane 200 μm in front of the reflector.

Next the microreflector was used in a fiber coupler configuration to provide fine alignment between a semiconductor laser and a single-mode optical fiber (9 μm core) . The light from a standard telecommunications grade 1.3 μm DFB laser was imaged by a lens system onto the fiber facet. All components were kept on separate alignment stages, and the actuated microreflector was positioned between the lens and the fiber to provide fine alignment (Figure 6). The microreflector was actuated to provide a maximum coupling efficiency of 32%. Figure 5 shows the dependence of coupling efficiency on the number of rear vibromotor steps. The step size can be varied by changing the drive voltage to provide various degrees of control.

The data show a good fit to a Gaussian curve with deviations being due to slider wobble and angular misalignment of the components. The coupling efficiency was limited to 32% in this experiment due to misalignment of the bulk components. However, efficiencies up to 45% were previously demonstrated using externally actuated devices [1,6], and, with all on-chip components and a higher placement precision, should reach 70%.

CONCLUSIONS

In conclusion, we have fabricated and characterized a vibromotor actuated micromachined microreflector for use in integrated optical systems on a silicon chip. The microreflector has two degrees-of-freedom and is able to position an optical beam in two dimensions. It is designed for use in self-contained laser-to-fiber coupling modules, integrated external-cavity devices, as well as other silicon-optical-bench applications requiring precise optical alignment. The reflector is capable of positioning an optical beam a fiber with a resolution of 0.81 μm, and a significantly higher precision is expected in future designs. We have demonstrated repeatable beam scanning, with the scanning speed and step size determined by the applied electrical signal. We also used the microreflector for fine alignment in a laser to single-mode fiber coupler, achieving precise control of the beam and a 32% coupling efficiency. The actuated microreflector and devices like it mark an important step towards making useful and cost-effective miniaturized integrated optical system.

REFERENCES

1. O. Solgaard, M. Daneman, N.C. Tien, A. Friedberger, R.S. Muller, K.Y. Lau, "Optoelectronic packaging using silicon surface-micromachined alignment mirrors," IEEE Photonics Technology Letters, vol. 7, no. 1, 1995, pp. 41-43.
2. L.Y. Lin, S.S. Lee, K.S.J. Pister, M.C. Wu, "Micro-machined three-dimensional micro-optics for free-space optical systems," IEEE Photonics Technology Letters, vol. 6, no. 12, 1994, pp. 1445-1447.
3. K.S.J. Pister, M.W. Judy, S.R. Burgett, R.S. Fearing, "Microfabricated hinges," Sensors and Actuators (A), vol. 33, no. 3, 1992, pp. 249-256.
4. W.C. Tang T.-C. H. Nguyen, M.W. Judy, R.T. Howe, "Electrostatic-comb drive of lateral polysilicon resonators, " Sensors and Actuators, vol. A21-23, 1990, pp. 328-331.
5. Y.H. Cho, A.P. Pisano, R.T. Howe, "Viscous Damping Model for Laterally Oscillating Microstructures," Journal of Microelectromechanical Systems, vol. 3, no. 2, 1994, pp.81-87.
6. M.J. Daneman, O. Solgaard, N.C. Tien, K.Y. Lau, R.S. Muller, "Integrated Laser-to-Fiber Coupling Module Using a Micromachined Alignment Mirror," Photonics Technology Letters, vol. 8, no. 3, March, 1996.
7. G.T. Mulhern, D.S. Soane, R.T. Howe, "Supercritical Carbon Dioxide Drying of Microstructures," 7th International Conference on Solid State Sensors and Actuators, 1993, pp. 296-299.
8. M.J. Daneman, N.C. Tien, O. Solgaard, A.P. Pisano, K.Y. Lau, R.S. Muller, "Linear Microvibromotor for Positioning of Optical Elements," Microelectromechanics Structures Conference, Amsterdam, The Netherlands, January 1995, pp. 55-60.

CHARACTERISTICS OF A POLYIMIDE MICROVALVE

Yu-Cheng Lin, Peter J. Hesketh, James G. Boyd[+]
Center for Applications of Microfabrication in Manufacturing,
Department of Electrical Engineering and Computer Science,
[+]Department of Mechanical Engineering,
University of Illinois at Chicago, Chicago, Illinois 60607

Susan M. Lunte*, and George S. Wilson*
*Center for Bioanalytical Research, University of Kansas, Lawrence, Kansas 66047

ABSTRACT

A new micromachined microvalve has been designed and fabricated for a miniature micropump for nanodialysis[1]. Nanodialysis is an *in vivo* technique that permits continuous sampling of drugs and endogenous compounds in awake animals[2]. Twenty microvalves are fabricated on each 2" silicon wafer. These microvalves have been integrated into a micropump for on-line nanodialysis.

INTRODUCTION

Micromachining and microfabrication techniques are now established processes for the fabrication of miniature sensors, microactuators, and microfabricated devices for microsystems.

More recently, micromachining techniques have been applied to the fabrication of chemical and biochemical sensors, and more significantly to micro chemical systems[3,4]. Micropumps and microvalves are one of the key elements in these analytical systems. Especially, microvalves play an important role for micropumps to function properly.

The microvalves can be categorized in two types: active and passive[5-20]. The passive microvalves are function as check valves. The valves open under forward applied pressure and have small leakage under backward applied pressure. The polyimide microvalve discussed in this paper is a passive microvalve.

FABRICATION

The cavities and channels are 50 μm deep and made by anisotropic etching of silicon in 50% CsOH at 60°C for 4.5 hours, using SiO_2 masking layers on both sides of the wafer. The silicon wafer was anodically bonded to 2¼" diameter Pyrex glass disc by using 700-1000 V at 350°C in a nitrogen purged cabin to avoid oxidation of polyimide. The glass surface was damaged in the bonding process when using one contact point probe, however, there was no damage on the glass surface when using a gold coated surface beneath the contact. The glass opening is etched in 50% HF with Cr/Au masking layer to form the valves.

The polyimide microvalve, shown in Figure 1, has an opening diameter of 300 μm and 50 μm wide cross reinforcement strips.

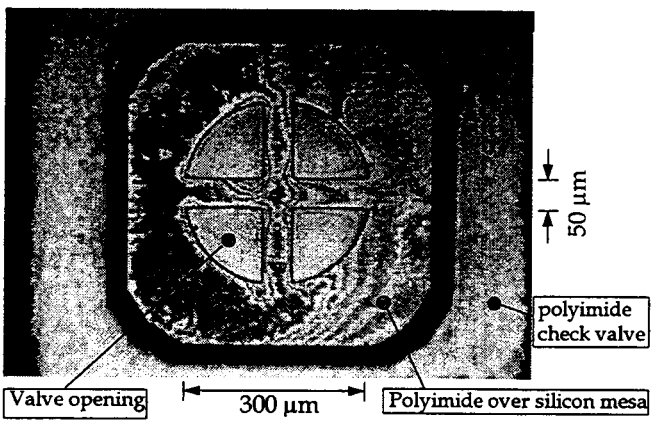


Figure 1. Photograph of the polyimide check valve with opening diameter of 300 μm and 50 μm wide cross reinforcement strips.

EXPERIMENT

The flow rate of the microvalve was measured by using the setup shown in Figure 2. The different applied pressures were obtained by changing the water height. The height remain the same by adding the DI water or alcohol with a syringe. The flow volume is obtained by scaling the syringe every five minutes. The pressure loss of this measurement setup has been calculated by using Poiseuille's equation and the Bernoulli equation for viscous flow, and found negligible compared to the applied pressure.

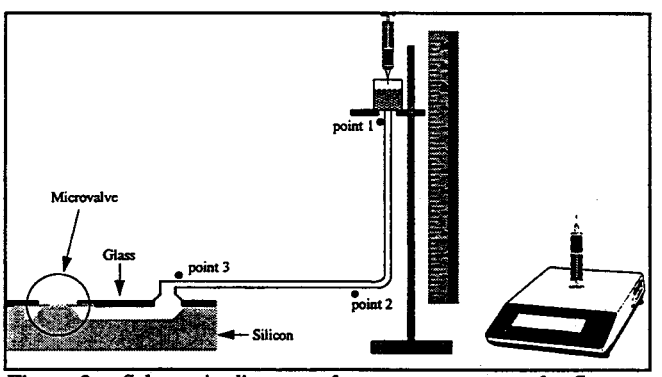


Figure 2. Schematic diagram of measurement setup for flow rate at different pressures.

0-9640024-1-8 $3.00 © 1996 TRF

Solid-State Sensor and Actuator Workshop
Hilton Head, South Carolina, June 2-6, 1996

All valves tested show an aging effect. The fresh valve has a lower flow rate than after aging. Figure 3 shows the aging flow rates of DI water through a fresh 4.6 μm thick polyimide check valve with two 50 × 200 × 1500 μm channels under 20 cm-H_2O pressure at 22°C.

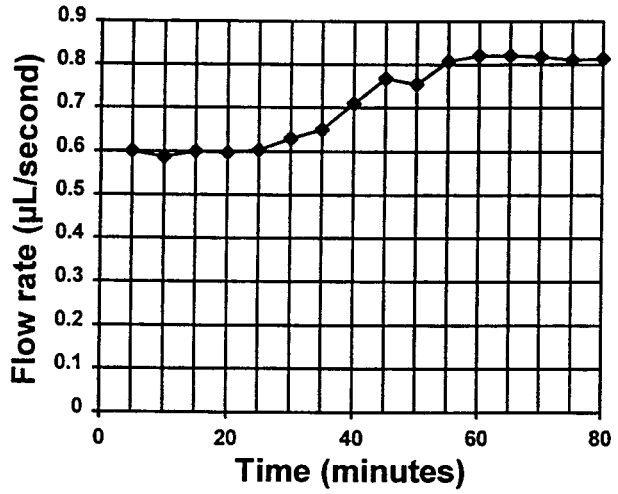


Figure 3. Valve aging and saturation with 20 cm-H_2O pressure applied. DI water flows through a fresh 4.6 μm thick polyimide check valve with two 50 × 200 × 1500 μm channels under 20 cm-H_2O pressure at 22°C.

The first check valve was made with a 2.6 μm thick layer of polyimide. The flow rates of DI water under different forward pressures through the 2.6 μm thick polyimide check valve with a 58 × 500 × 1500 μm channel are shown in Figure 4.

The 4.6 μm check valve is stronger than 2.6 μm one for higher pressure applications. Figure 5 shows the flow rates of DI water through a 4.6 μm thick check valve and two 50 x 200 x 1500 μm channels at different forward pressures after aging. The dashed line indicates the flow rate is proportional to the applied pressure, and the minimum pressure to start flow is less than 5 cm-H_2O pressure.

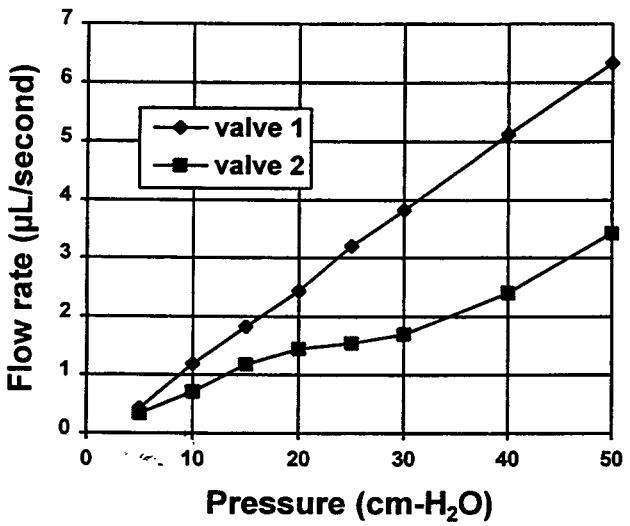


Figure 4. Comparison of the flow rates of DI water through two different 2.6 μm thick check valves and the 58 × 500 × 1500 μm channel under different forward pressures.

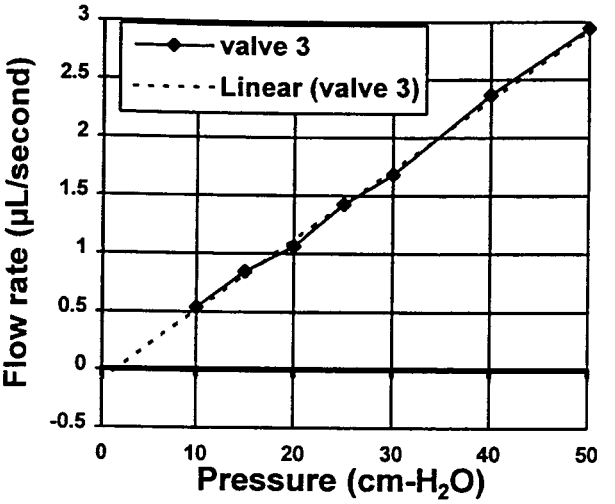


Figure 5. Flow rate of DI water through a 4.6 μm thick check valve and two 50 x 200 x 1500 μm channels at different forward pressures after aging.

The flow rate ratios of DI water for equal forward and backward pressures of the 4.6 μm thick microvalve are above 200, shown in Figure 6.

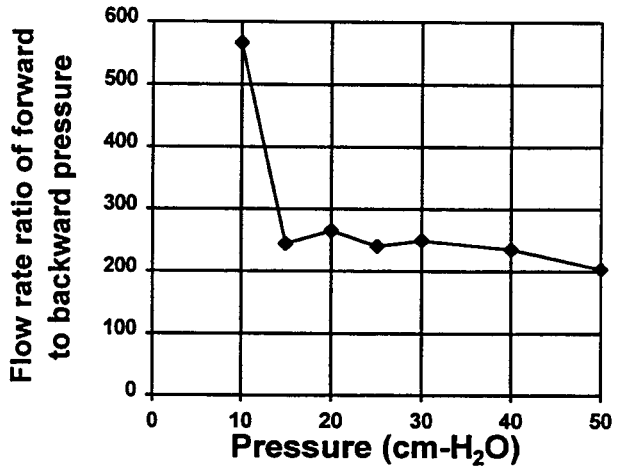


Figure 6. The ratio of flow rate of DI water for equal forward and backward pressures through a 4.6 μm thick polyimide check valve and two 50 x 200 x 1500 μm channels at 22°C.

The flow rates of isopropyl alcohol through a 4.6 μm check valve and three 50 × 200 × 1500 μm channels under different forward pressures are shown in Figure 7. Each measurement is 5 minutes at 22°C.

The isopropyl alcohol leakage through 4.6 μm thick check valves with three 50 × 200 × 1500 μm channels under different backward pressures are shown in Figure 8. The dashed lines of the flow rates indicate the leakage is proportional to the applied pressure.

The ratios of flow rate of isopropyl alcohol for equal forward and backward pressure through a 4.6 μm thick check valves with three 50 × 200 × 1500 μm channels for valve 6 and valve 7 are less than 15:1.

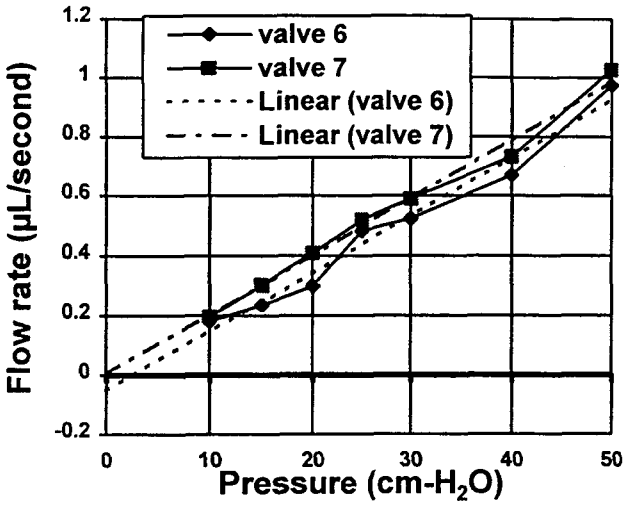


Figure 7. Comparison of two different 4.6 µm check valves with isopropyl alcohol as the working fluid. The flow runs through a 4.6 µm check valve and three 50 × 200 × 1500 µm channels under different forward pressure.

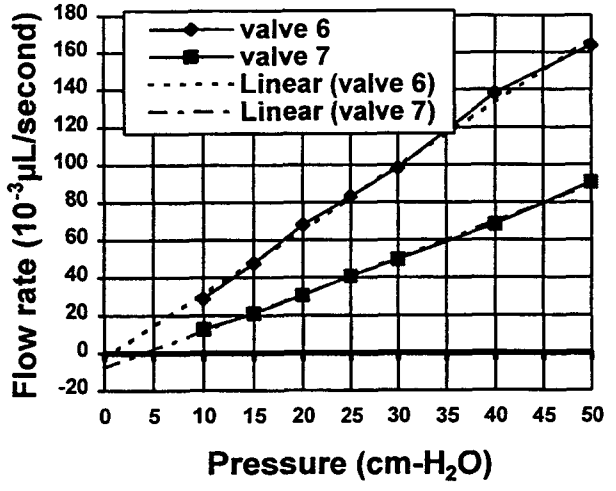


Figure 8. The isopropyl alcohol leakage through a 4.6 µm check valve and three 50 × 200 × 1500 µm channels under different back pressures. The dashed lines are linear and indicate the leakage is proportional to the pressure, and the minimum pressure to start leakage is less than 5 cm-H_2O pressure.

CONCLUSIONS

All valves have an aging effect. The fresh valve has a lower flow rate than after aging. The flow rate of this valve is proportional to the applied pressure, and the minimum pressure needed to start flow is quite small. The ratios of forward flow rate to backward leakage are above 200 with DI water as the working fluid. However, the ratios are less than 15 with isopropyl alcohol as the working fluid.

REFERENCES

1. Yu-Cheng Lin, Peter J. Hesketh, "A Micromachined Diaphragm Micropump", in *Microstructures and Microfabricated Systems II*, Proceedings of Electrochemical Society, Inc. p. 67-72, Vol. 95-27, Pennington, NJ, 1995.

2. Barry L. Hogan and Susan M. Lunte, "On-line Coupling of In Vivo Microdialysis Sampling with Capillary Electrophoresis", *Analytical-Chemistry*, pp. 596-602, V66, N5, March, 1994.

3. S. C. Jacobson, R. Hergenroder, A. W. Jr. Moore, and J. M. Ramsey, "Precolumn reactions with electrophoretic analysis integrated on a microchip", *Analytical Chemistry* 66:4127-4132, 1994.

4. D. J. Harrison, K. Fluri, N. Chiem, T. Tang, and Z. Fan, "Micromachining chemical and biochemical analysis and reaction systems on glass substrates", *Transducers '95, Eurosensors IX*, pp. 752-755, 1995.

5. S. C. Terry, J. H. Jerman and J. B. Angell, "A gas chromatographic air analyzer fabricated on a silicon wafer", *IEEE Trans. Electron Device*, 26:1880-1886, 1979.

6. M. J. Zdeblick, and J. B. Angell, "A microminiature electric-to-fluidic valve", *IEEE Technical Digest of Transducers*, 87:827-929, 1987.

7. M. J. Zdeblick, R. Anderson, J. Jankowski, B. Kline-Schoder, L. Christel, R. Miles, and W. Weber, "Thermopneumatically actuated microvalves and integrated electro-fluidic circuits", *Technical Digest IEEE Solid-State Sensors and Actuator Workshop*, pp. 251-255, 1994.

8. S. Shoji, M. Esashi, and M. Matsuo, "Prototype miniature blood gas analyzer fabricated on a silicon wafer", *Sensors and Actuators* 14:101-107, 1988.

9. M. Esashi, "Integrated microflow control systems", *Sensors and Actuators* A14-A23:161-167, 1990.

10. S. Shoji, and M. Esashi, "Micromachining for chemical sensors", *Chem. Sensor Techn.* 1:179-193, 1988.

11. T. Ohnstein, T. Fukiura, J. Ridley, and U. Bonne, "Micromachined silicon microvalve", *Proc. IEEE Micro Electro Mechanical Systems Workshop*, pp. 95-98, 1990.

12. H. Jerman, "Electrically-activated micromachined diaphragm valves", *Technical Digest IEEE Sensors and Actuator Workshop*, pp. 65-69, 1990.

13. K. Yanagisawa, H. Kuwano, and A. Tago, "An electromagnetically driven microvalve", *IEEE Technical Digest of Transducers* 93:102-105, 1993.

14. D. Bosch, B. Heimhofer, G. Muck, H. Seidel, U. Thumser, and Welser, W., "A silicon microvalve with combined electromagnetic/electrostatic actuation", *Sensors and Actuators* A37-A38:684-692, 1992.

15. H. T. G. Van Lintel, and F. C. M. Van de Pol, "A piezoelectric micropump based on micromachining of silicon", *Sensors and Actuators* 15:153-167,(1988.

16. J. Tiren, L. Tenerz, and B. Hok, "A batch-fabricated non-reverse valve with cantilever beam manufactured by micromachining of silicon", *Sensors and Actuators* 18:389-396, 1989.

17. M. Esashi, S. Shoji, and A. Nakano, "Normally closed microvalve and micropump fabricated on a silicon wafer", *Sensors and Actuators* 20:163-167, 1989.

18. L. Smith, and B. Hok, "A silicon self-aligned non-reverse valve", *IEEE Technical Digest of Transducers*, 1049-1051, 1991.

19. Schomburg and Scherrer, "3.5 μm thin valves in titanium membranes", *J. Micromech. Microeng.* 2:184-186, 1992.

20. S. Shoji, M. Esashi, B. Schoot, and N. Rooij, "A study of a high-pressure micropump for integrated chemical analyzing systems", *Sensors and Actuators A* 32:335-339, 1992.

SEQUENTIAL SOLUTIONS OF FIELD EQUATIONS USING A SINGLE BEM MODEL

Per B. Ljung
Berkeley Sensor & Actuator Center
497 Cory Hall
University of California Berkeley
Berkeley, CA 94720

ABSTRACT

A new portable CAD tool is introduced allowing a single boundary element model to solve coupled nonlinear Laplace, Poisson and Navier field equations. The compact formulation of the model and discretized differential equations allows the software to be implemented on standard personal computers in vectorized Matlab code. Several 2D and 3D examples are given using common micromachined structures.

INTRODUCTION

Designers of microelectromechanical systems need to be able to analyze 3D coupled multi-energy domains within a single integrated CAD environment. Currently separate analysis tools [1]-[7] are used for thermal, electrostatic, elastostatic, thermoelastic, and magnetostatic field solutions, iterating between tools until results converge. This requires redundant data preparation since the domains must be modeled appropriately for each individual CAD tool. Further, this results in inefficient operation due to the overhead of translating intermediate results between different CAD tools. Nontrivial models become unwieldy requiring up to 1 month processing time [5] to solve coupled electrostatic and elastostatic problems using iteration between separate CAD packages. These large coupled FEM problems have also been solved with parallel computation [8] of linked workstations.

DUAL RECIPROCITY BEM

In contrast to earlier work, this paper presents an integrated CAD tool capable of solving coupled field equations using a single boundary element method (BEM) mesh. Since only the boundaries of the system are modeled, this results in simple data preparation as well as small data structures and fast simulation. Further, the single BEM mesh is used for each desired field analysis resulting in efficient numerical interaction for large nontrivial coupled problems as depicted schematically in Figure 1.

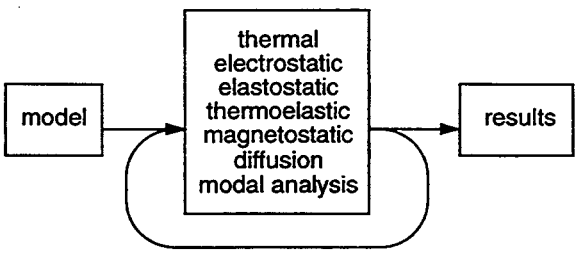


Figure 1. A single BEM model suffices to solve an arbitrary sequence of coupled field equations.

Compared to FEM which only solves for the desired state variable, BEM simultaneously solves for both the state variable and flux. While traditional BEM programs can solve static 2D and 3D Laplace and Navier field equations, the dual reciprocity boundary element method (DR-BEM) [9] used here can solve:
Laplace

$$\nabla^2 u = 0; \tag{1}$$

time-dependent Poisson

$$\nabla^2 u = b(x, y, z, u, t); \tag{2}$$

and time-dependent Navier

$$G u_{j,kk} + \frac{G}{1-2\nu} u_{k,kj} + b_j = \rho \ddot{u}_k \tag{3}$$

field equations. The DR-BEM is implemented in portable vectorized Matlab [10] code using isoparametric surface elements to solve Poisson and Navier equations [11] enabling the calculation of distributed LCR-networks, electrostatics, elastostatics, thermoelastics, magnetostatics and modal analysis. Figure 2 shows a diagram where the field equation, valid in the domain Ω, is evaluated on the discretized boundary elements and internal nodes.

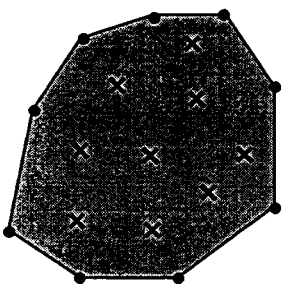


Figure 2. DR-BEM boundary discretization using linear elements consisting of N external nodes and L internal nodes.

Problems with nonhomogeneous time-varying forcing functions and material nonlinearities are fully supported by DR-BEM while problems with geometric nonlinearities are solved using iteration. Coupled systems such as electrostatic-elastostatic problems are solved by iteration between the DR-BEM field solutions, whereas coupled systems such as thermoelastic-modal analysis problems are solved directly using the first field solution.

PORTABLE MATLAB IMPLEMENTATION

Solving a discretized boundary element problem requires matrix manipulation, numerical integration and solving linear systems. For this reason, it was expedient to use Matlab as the platform for the boundary element software. Consider

$$\nabla^2 u = b \tag{4}$$

which allows the modelling of:

(i) Laplace problems, where $b = 0$;
(ii) diffusion problems, where $b = b(\partial u/\partial t)$;
(iii) convection problems, where $b = b(\partial u/\partial x)$;
(iv) nonlinear material problems; and
(v) modal analysis problems, where $b = -u$.

0-9640024-1-8 $3.00 © 1996 TRF

A weighted residual formulation of Eq. 4 produces

$$\int_{\Omega}\left(\nabla^2 u - b\right)\underline{u}\,d\Omega = \int_{\Gamma_2}(q-\bar{q})\,\underline{u}\,d\Gamma - \int_{\Gamma_1}(u-\bar{u})\,\underline{q}\,d\Gamma \qquad (5)$$

where Ω is the solution domain, $\underline{u}$ is the fundamental solution [11] such that $\nabla^2\underline{u} = -\Delta$, and where Δ is the Dirac delta function. The specified boundary conditions are u_1 on Γ_1 and $q_2 = \partial u/\partial n$ on Γ_2 where $\Gamma = \Gamma_1 + \Gamma_2$ is the total boundary and n is the normal to the boundary. Eq. 5 is integrated twice by parts and grouping all the boundary terms together, one obtains for a point source i

$$c_i u_i + \int_{\Gamma} u\underline{q}\,d\Gamma - \int_{\Gamma} q\underline{u}\,d\Gamma = \int_{\Omega} b\underline{u}\,d\Omega\,. \qquad (6)$$

Standard BEM schemes [11] cannot evaluate the domain integral on the right-hand side of Eq. 6. However, using the dual reciprocity approximation [9], the domain term is approximated as

$$b \cong \sum_{j=1}^{N+L}\alpha_j f_j \qquad (7)$$

where N are the number of boundary nodes, L are the number of internal nodes, α_j are a series of unknown coefficients and f_j are a set of approximating functions, which for the case of the Laplace operator satisfy

$$\nabla^2\hat{u}_j = f_j\,. \qquad (8)$$

Substituting Eq. 8 into Eq. 7 results in

$$\nabla^2 u = b \cong \sum_{j=1}^{N+L}\alpha_j\left(\nabla^2\hat{u}_j\right). \qquad (9)$$

Multiplying by the fundamental solution $\underline{u}$ and integrating over the domain gives

$$\int_{\Omega} b\underline{u}\,d\Omega = \sum_{j=1}^{N+L}\alpha_j\int_{\Omega}\left(\nabla^2\hat{u}_j\right)\underline{u}\,d\Omega\,. \qquad (10)$$

Integrating twice by parts for each source point i leads to

$$c_i u_i + \int_{\Gamma} u\underline{q}\,d\Gamma - \int_{\Gamma} q\underline{u}\,d\Gamma = \sum_{j=1}^{N+L}\alpha_j\left(c_i\hat{u}_{ij} + \int_{\Gamma}\underline{q}_i\hat{u}_j\,d\Gamma - \int_{\Gamma}\underline{u}_i\hat{q}_j\,d\Gamma\right) \qquad (11)$$

Discretizing over N boundary elements and evaluating the integrals using gaussian numerical integration produces

$$c_i u_i + \sum_{k=1}^{N} H_{ik}u_k - \sum_{k=1}^{N} G_{ik}q_k = \sum_{j=1}^{N+L}\alpha_j\left(c_i\hat{u}_{ij} + \sum_{k=1}^{N} H_{ik}\hat{u}_{kj} - \sum_{k=1}^{N} G_{ik}\hat{q}_{kj}\right) \qquad (12)$$

After evaluation at all boundary nodes i, the following matrix equation results

$$Hu - Gq = \sum_{j=1}^{N+L}\alpha_j\left(H\hat{u}_j - G\hat{q}_j\right). \qquad (13)$$

Collecting the vectors $\hat{u}_j$ and $\hat{q}_j$, the nonhomogeous terms form the vector

$$C = \left(H\hat{U} - G\hat{Q}\right)F^{-1}b \qquad (14)$$

This can be rearranged to reflect that either the state variable u_1 or flux q_2 is known at each node

$$\begin{bmatrix}H_{u_1} & H_{u_2}\end{bmatrix}\begin{bmatrix}u_1\\ u_2\end{bmatrix} - \begin{bmatrix}G_{q_1} & G_{q_2}\end{bmatrix}\begin{bmatrix}q_1\\ q_2\end{bmatrix} = C\,. \qquad (15)$$

Rearranging so that all the unknowns are on the left-hand side results in

$$\begin{bmatrix}-G_{q_1} & H_{u_2}\end{bmatrix}\begin{bmatrix}q_1\\ u_2\end{bmatrix} = \begin{bmatrix}-H_{u_1} & G_{q_2}\end{bmatrix}\begin{bmatrix}u_1\\ q_2\end{bmatrix} + C\,. \qquad (16)$$

which is in the form of $Ax = B$ which can be solved using the Matlab linear system solver. Note that the matrices G and H are fully populated and nonsymmetric. Typically 95% of the simulation time is spent forming the matrices G and H and the remaining 5% in solving the linear system of Eq. 16.

Note that only boundary elements are required for homogeneous equations, and that no approximations were made other than the discretization of the boundary and numerical integration. Internal nodes are required for nonhomogeneous equations using DR-BEM which introduces an approximation in the evaluation of the volume integral. In contrast, the FEM introduces approximations in the solution over each element as well as discretization and numerical integration approximations. Even with a coarse boundary mesh and few internal nodes, the DR-BEM is extremely accurate compared with a much finer volume mesh used in FEM. Using DR-BEM allows the problem to be easily discretized with coarse meshes leading to low memory requirements allowing its implementation on standard personal computers.

Similar DR-BEM expansions for the Navier fundamental solution [11] are readily generated such that

$$Ad - Bt = C \qquad (17)$$

where d_i is the displacement and t_i is the traction at node i.

Thermoelastic problems can be written as two matrix equations

$$\begin{aligned} Hu - Gq &= C_1 \\ Ad - Bt + Eu - Fq &= C_2 \end{aligned} \qquad (18)$$

where the temperature u and flux q are solved from the thermal Laplace equation and d and t are the resulting elastostatic deflection and traction, respectively. If the thermal variables u, q do not rely on the elastostatic deflection t, then the thermal and thermoelastic equations can be solved simultaneously.

Coupled electrostatic and elastostatic field equations can be written as two matrix equations

$$\begin{aligned} Hu - Gq &= C_1 \\ Ad - Bt - Fq &= C_2 \end{aligned} \qquad (19)$$

where the electrostatic force is calculated in the F matrix. Due to the strong coupling between the Laplace solution the elastostatic deflection, these coupled equations must be solved using iteration.

For modal analysis, the problem to be solved is

$$\nabla^2 u + \omega^2 u = 0 \qquad (20)$$

which is of the form Eq. 4 where $b = -\omega^2 u$. In this case Eq. 14 can be rewritten as

$$C = \omega^2\left(H\hat{U} - G\hat{Q}\right)F^{-1}u \qquad (21)$$

which can be repartitioned [12] to form a generalized eigenvalue problem

$$Ku = \omega^2 Mu \qquad (22)$$

where Matlab can solve for the eigenvalues and eigenvectors. Note that since the matrices are full and nonsymmetric, complex eigenvalues may result due to numerical discretization.

ELECTROSTATIC EXAMPLES

3D electrostatic fields must be addressed even for nominally planar geometries such as parallel-plate capacitors. While the BEM program FASTCAP [6] can solve Laplace's equation with constant dielectrics, it has several disadvantages including the use of constant elements, no provision for symmetry planes to simplify modeling, and no provision for groundplanes. The present program can solve the more general Poisson's equation allowing for continuously varying dielectrics. Further, it fully supports the use of symmetry planes and groundplanes. Finally, it uses isoparametric elements allowing linear or quadratic interpolation of the potential and flux over the element resulting in much higher accuracy with fewer elements.

Simple parallel-plate example

The potential and flux are calculated for the 3D parallel-plate over a groundplane as shown in Figure 3. The model consists of the surfaces of the plate (width=4, thickness=2) and an underlying groundplane (gap=2) forming a capacitor. The arrows indicate the flux direction and magnitude. Note the flux concentration along the bottom surface edges. This concentration leads to a higher capacitance than the ideal parallel-plate model indicates.

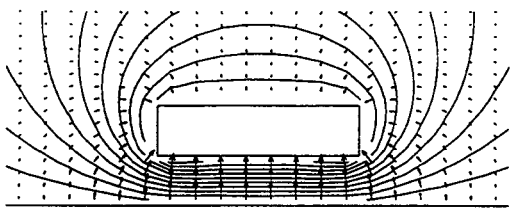


Figure 3. Cross-section of 3D parallel-plate over a groundplane showing equipotential lines and flux. Note the flux concentration along the corners.

The fringing field capacitance associated with each surface is calculated and shown in Figure 4 for the 2D case and in Figure 5 for the 3D case.

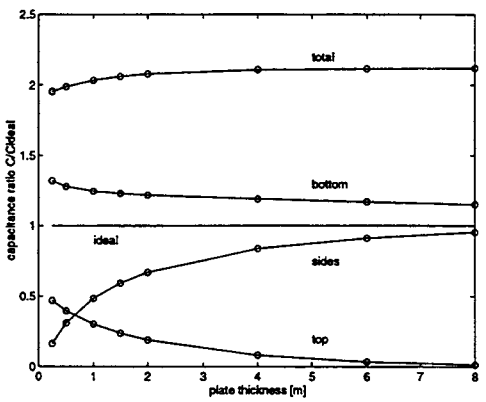


Figure 4. Capacitance of 2D parallel plate showing the fringing field contributions from the bottom, side and top surfaces.

The 2D BEM simulations show that the capacitance is about 2x larger than the ideal 2D parallel-plate capacitor. As shown in Table 1, the 2D BEM shows relative errors >1% compared to the exact 2D conformal mapping capacitance. While no analytical solution for the 3D parallel-plate capacitance exists, simulations show that the converged 3D BEM capacitance is about 6x larger than the ideal 3D parallel-plate.

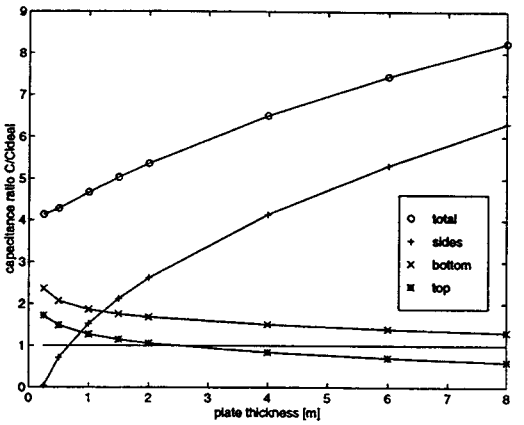


Figure 5. Capacitance of 3D parallel plate showing the fringing field contributions from the bottom, side and top surfaces.

For large plate thicknesses, the total sidewall area rapidly becomes larger than the bottom surface area. The flux concentration along the bottom and side edges increases the capacitance of the plate, accounting for the large capacitance difference between the calculated 2D and 3D models. The simulation times on a Macintosh PowerPC portable computer is also shown in Table 1.

TABLE 1: 2D WIRE CAPACITANCE RATIO C/C_{IDEAL}

thickness gap=2 width=4	2D BEM 26 elements (5 sec)	2D BEM 104 elements (110 sec)	conformal mapping [13]
0.25	2.43 (10.3%)	2.23 (1.2%)	2.21
0.5	2.45 (7.9%)	2.28 (0.4%)	2.27
1	2.52 (6.2%)	2.38 (0.2%)	2.38
2	2.67 (5.5%)	2.53 (0%)	2.53
4	2.90 (5.7%)	2.71 (-1.4%)	2.75
8	3.24 (7.0%)	2.91 (-3.9%)	3.03

Electrostatic comb-finger

Another common structure requiring accurate flux calculation is the electrostatic comb-finger used for forcing and transduction in resonators. The BEM model comprises a rotor finger, a stator finger and a groundplane. Symmetry planes are used such that the model represents a section of an infinite comb-finger array. From the flux results shown in Figure 6, the capacitance and electrostatic forces are directly calculated. Note that no deflection of the rotor is required to calculate the force as is the case with a FEM analysis. The capacitance (Figure 7), lateral electrostatic force (Figure 8) and levitation electrostatic force (Figure 9) are shown for different comb-finger rotor positions. The capacitance and force calculations use the permittivity $\varepsilon = 1$.

Note that the drive force does not increase monotonically as the comb-fingers engage due to charge cancellation on the finger sidewalls, an effect which has not previously been reported. The maximum lateral force is obtained when the fingers do not engage: all the flux lines from the sides and tip of the rotor finger then have a component in the x-direction. As the fingers engage, most of the flux on the sides of the rotor no longer has an x-component which reduces the total lateral force. As the rotor tip approaches the sta-

tor anchor, the parallel-plate effect becomes dominant and the attractive force increases.

Also note that the levitation force is 4x larger than the lateral force. If this is undesired, a different shaped groundplane may be used to minimize the levitation force. The capacitance due to the fringing field is about 2x larger than the ideal overlapping plates. A large part of this capacitance can be treated as a constant parasitic capacitance since the lateral drive force is approximately the same as expected for ideal overlapping plates. The levitation force is about 3x larger than expected however due to fringing field components.

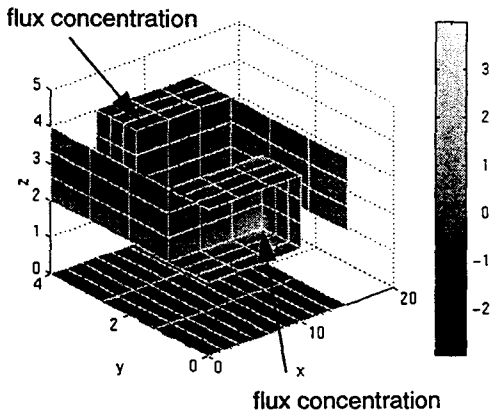


Figure 6. Electrostatic flux distribution on a section of a infinite comb-finger assembly.

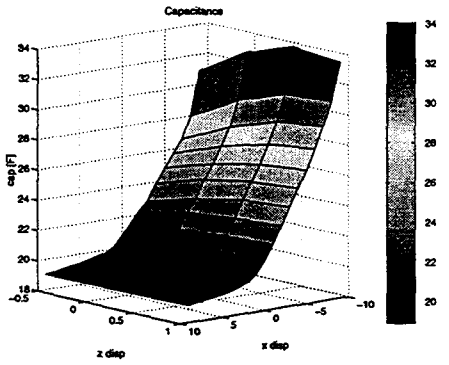


Figure 7. Comb-finger capacitance as a function of rotor position.

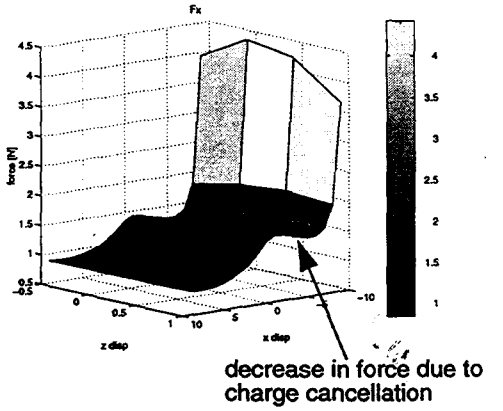


Figure 8. Lateral electrostatic force on comb-finger rotor as a function of rotor position.

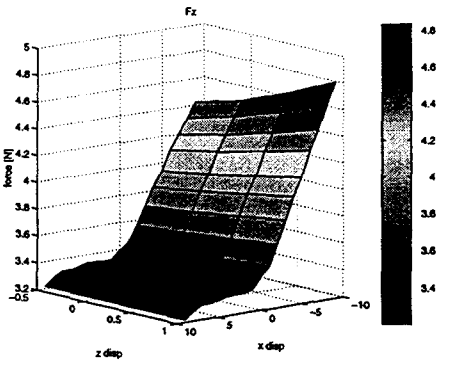


Figure 9. Levitation electrostatic force on comb-finger rotor as a function of rotor position.

Electrostatic etch-holes

A 3D BEM model of a plate with etch-holes is shown in Figure 10. Symmetry planes on the vertical surfaces are defined such that the model represents a section of an infinite plate.

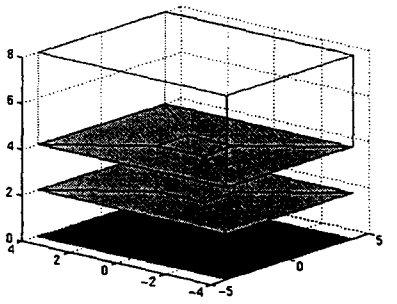


Figure 10. BEM model of a section of an infinite parallel plate with etch-holes.

The capacitance matrix consisting of capacitance contributions from the bottom, side and top surfaces is calculated for different size etch-holes. Large etch-holes do not significantly reduce the total capacitance since the flux associated with the edges and side-wall fringing increases. For the same reason, a parallel plate with small etch holes has a larger capacitance than an ideal parallel plate.

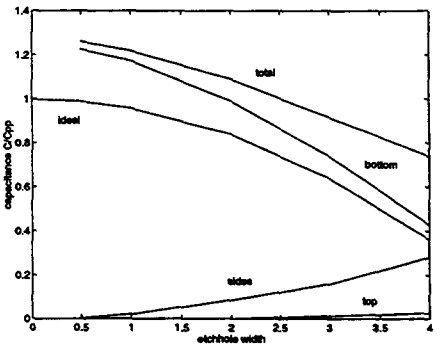


Figure 11. Capacitance vs. etch-hole width. The fringing field contributions from the bottom, side and top surfaces are shown.

The flux concentration along the edges of the bottom plate due to the etch-holes result in a bottom plate capacitance contribution

larger than that expected compared to an ideal parallel plate capacitor of the same area. With larger etch-holes, the capacitance contribution from the side walls becomes significant. The capacitance contribution from the top surface is negligible. As Figure 11 shows, an effective capacitance approximation is to ignore the effect of any etch-holes on a parallel-plate.

DIAPHRAGM WITH COUPLED FIELDS

Consider the diaphragm pressure sensor as an example of a common micromachined structure under the influence of several different coupled fields. During the fabrication process, a residual stress and residual stress gradient results in a warping of the diaphragm. During operation, pressure is applied resulting in an elastostatic deformation of the diaphragm. This deformation is capacitively measured by applying a voltage differential between the bottom of the diaphragm and the groundplane. If the resulting electrostatic field is large, the diaphragm will deflect due to the large attractive electrostatic forces. Finally, the natural frequencies of the structure are to be calculated. The effect of each of these applied fields are simulated below using a simple 2D model for clarity.

The diaphragm 2D BEM model consists of two closed regions: the diaphragm and the space between the diaphragm bottom surface and groundplane. The model consists of 30 nodes and 40 linear elements as shown in Figure 12. The appropriate boundary conditions and material properties are defined for each region.

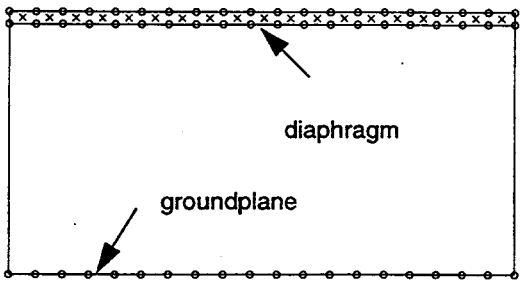


Figure 12. 2D BEM model of diaphragm pressure sensor consisting of 30 linear elements and 19 internal nodes.

Stress gradient deflection

Surface micromachined structures typically have residual strains σ and residual strain gradients Γ from the fabrication process. The strain gradient Γ results in a warping of the structure which must be modeled. Rather than apply stresses directly to the model, it is possible to calculate an equivalent temperature and temperature gradient which will result in the same stress state. Assuming a stress field varying in the z direction, a temperature field causing the same average stress and deflection is

$$\sigma(z) = E\alpha T(z) = E\alpha\left[T(0) + T'(0)z + T''(0)\frac{z^2}{2!} + \ldots\right] \quad (23)$$

By identifying terms one obtains expressions for the average temperature and the linear temperature gradient:

$$\sigma|_{z=0} = E\alpha T|_{z=0} \quad (24)$$

$$\Gamma = \sigma'|_{z=0} = E\alpha T'|_{z=0}. \quad (25)$$

The residual stress and stress gradient are modeled by applying a temperature T_1 to the top surface and T_2 to the bottom surface of the diaphragm. The resulting heat problem is readily solved for any unknown node temperatures and element temperature flux. These thermal state variables are then used to calculate the thermal elastostatic stress state. Any externally applied elastostatic deflections and tractions can also be linearly added. In this case, the ends of the diaphragm were specified as fixed $d = 0$ and no external tractions t were applied. The resulting thermoelastic deformation is shown in Figure 13.

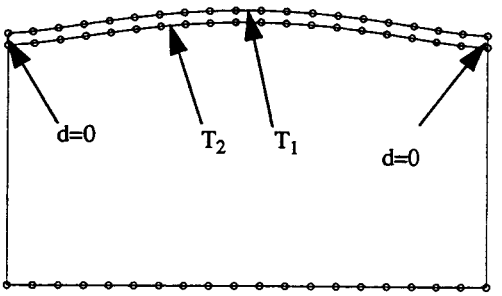


Figure 13. Deflection of diaphragm due to residual stress gradient.

Pressure deflection and capacitance

A pressure over the diaphragm is applied by specifying t at each element node on the top surface. Note that because the elements use linear interpolation, a constant t corresponds to a constant evenly applied pressure across the diaphragm. The resulting deflection due to the residual strain gradient and applied pressure is shown in Figure 14. The resulting capacitance can be calculated by applying arbitrary voltages u_1, u_0 to the bottom of the diaphragm and the groundplane. The resulting electrostatic problem is in the form of Laplace's equation which is readily solved for any unknown node voltage and element electrostatic fluxes.

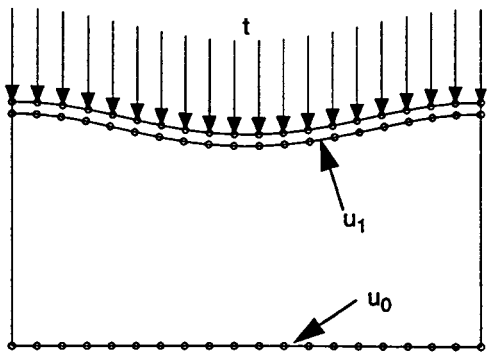


Figure 14. Deflection of diaphragm due to applied external pressure and residual stress gradient. (The tractions t are defined normal to the diaphragm surface.)

Electrostatic force and deflection

Depending on the diaphragm configuration, the attractive electrostatic forces due to the applied sense voltage can cause significant deflections. This elastostatic deflection in turn changes the electrostatic forces. This coupled problem can be solved by iterating between the elastostatic and electrostatic solvers until the change in deflection is not significant. Figure 15 shows the deflection of the diaphragm due to electrostatic forces, applied pressure and residual stress gradient. Equipotential lines and flux arrows are also shown.

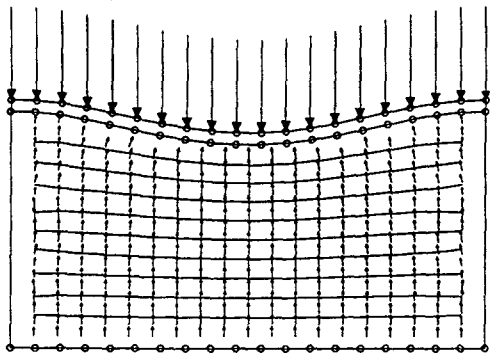


Figure 15. Deflection of diaphragm due to differential sense voltage, applied pressure, and residual stress gradient.

Modal analysis

The DR-BEM can be used to calculate the eigenvalues and mode shapes of the diaphragm. The additional stiffness caused by the applied electrostatic forces, pressure and residual stress gradients has not been included. An additional 19 internal nodes were arranged along the centerline of the diaphragm to form the right-hand side of Eq. 21. Rearranged into effective spring and mass matrices of Eq. 22 allows Matlab to directly calculate the eigenvalues and eigenmodes. Due to discretization approximations, some of these eigenvalues are complex. The first eigenvalues agree closely with those expected for a fixed-fixed diaphragm.

ACCELEROMETER WITH COUPLED FIELDS

Complex 3D structures such as lateral surface micromachined accelerometers can be modelled and simulated in the same manner. Figure 16 shows the exaggerated deflection of the accelerometer due to the same stress gradient. The BE mesh is used to solve for the elastostatic field *inside* the polysilicon structure. The same BE mesh is used to calculate the electrostatic field *outside* the polysilicon structure. In order for the capacitance calculations to converge, many elements are required along the edges of the comb-fingers. Since both FEM and BEM problems scale as $O\left(N^2\right)$, this leads to lengthy simulations. Each electrostatic and elastostatic iteration required 0.4 and 6 hours on a Macintosh PowerPC portable computer. The natural frequencies and eigenmodes agree closely with traditional FEM analysis.

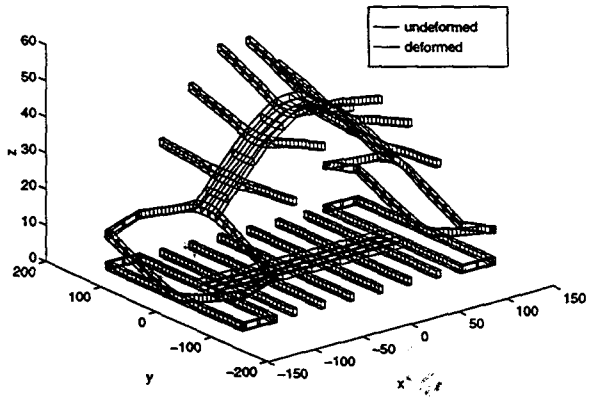


Figure 16. Deflection of surface micromachined accelerometer due to residual stress gradients. The ground-plane and stator fingers are not shown for clarity. (Z-axis scaled by 100)

CONCLUSIONS

A new portable BEM tool able to solve arbitrary 2D and 3D Poisson and Navier field equations has been introduced. The modeling effort is minimized since only the boundaries of the field must be described. Further, a single model is sufficient to allow sequences of different field equations to be solved. The compact model and DR-BEM formulation allows the software to be implemented in portable vectorized Matlab code on standard personal computers. Several electrostatic, elastostatic and thermoelastic fields in common micromachined structures are solved using this portable CAD tool.

REFERENCES

1. S. D. Senturia, "CAD for Microelectromechanical Systems", Transducers 95, Stockholm, SE, June 1995, pp. 5-8
2. J. R. Gilbert, R. Legtenberg, S. Senturia, "3D Coupled Electro-mechanics for MEMS: Applications of CoSolve-EM", Proc. MEMS 95, Amsterdam, NL, Feb. 1995, pp. 122-127
3. S. Senturia, R. Harris, B. Johnson, S. Kom, K. Nabors, M. Shulman, J. White, "A Computer-aided design system for microelectromechanical systems (MEMCAD)", J. MEMS, v.1, n.1, Mar. 1992, pp. 3-13
4. P. Osterberg, H. Yie, X. Cai, J. White, S. Senturia, "Self-Consistent simulation and modeling of electrostatically deformed diaphragms", Proc. IEEE MEMS, Oiso, Japan, Jan 25-28 1994, pp. 28-32
5. C. Xuejun, P. Osterberg, Y. Ha, J. Gilbert, "Self-consistent electromechanical analysis of complex 3-D microelectromechanical structures using relaxation/multipole-accelerated method", Sensors and Materials, 1994, v.6, n.2, pp. 85-99
6. K. Nabors, J. White, "FastCap: a multipole accelerated 3-D capacitance extraction program", IEEE Trans. CADICS, Nov. 1991, v.10, pp. 1447-59
7. M. Kamon, J. Ttsuk, J. White, "FastHenry: a multipole-accelerated 3-D inductance extraction program", IEEE Trans. MTT, Sept. 1994, v.42, n.9, pp. 1750-8
8. J. G. Korvink, J. Funk, M. Roos, G. Wachutka, H. Baltes, "SESES: A comprehensive MEMS Modelling System", Proc. IEEE MEMS, Oiso, Japan, Jan. 25-28 1994, pp. 22-27
9. P. W. Partridge, The Dual Reciprocity Boundary Element Method, Elsevier Applied Science, Southhampton, UK, 1992
10. Matlab Reference Guide, The Mathworks, Natick, MA, 1995
11. A. A. Becker, The Boundary Element Method in Engineering, McGraw Hill, 1992
12. D. P. N. Kontoni, P. W. Partridge, C. A. Brebbia, "The dual reciprocity boundary element method for the eigenvalue analysis of Helmholtz problems", Adv. Eng. Software, v.13 n.1, pp-2-16, 1991
13. W. H. Chang, "Analytical IC metal-line capacitance formulas", IEEE Trans. MTT, Sept 1976, pp. 608-612

ACKNOWLEDGMENT

I would like to thank the Hertz Foundation for their support.

APPLICATIONS OF X-RAY MICRO-IMAGING, VISUALIZATION AND MOTION ANALYSIS TECHNIQUES TO FLUIDIC MICROSYSTEMS

Ann-Marie Lanzillotto, Tzong-Shyng Leu, Mike Amabile and Rick Wildes
David Sarnoff Research Center, Princeton, NJ 08543-5300

John Dunsmuir
Exxon Research and Engineering Co. Annandale, NJ 08801

ABSTRACT

This paper describes x-ray imaging, visualization and analysis technologies which are utilized in an investigation of the micro-structure and micro-motion in various fluidic microsystems. More specifically, we have imaged and visualized the 3-D internal microstructure of Redwood Microsystems gas and liquid microvalves using x-ray microtomography and a massively parallel supercomputer, the Princeton Engine. In other experiments, x-ray microradiography was used to image fluid transport through the valves as well as valve actuation. In addition, we have begun to develop a method to visualize and measure steady-state flow in microchannels through the use of emulsions. The motion of the fluid and detailed flow pattern are directly viewable by motion of the liquid droplets. Finally, we have employed image motion analysis, bolstered by physical and geometric constraints derived from fluid mechanics, to recover meaningful flow fields and velocity profiles in microchannel flow. We found that our measured velocity profiles in 1000, 840, 750 and 640 micron inner diameter capillary tubes are in reasonably good agreement with the theoretical average velocity profiles of fully developed circular pipe flow.

INTRODUCTION

Numerous fluidic microdevices, such as valves, pumps and flow sensors, are rapidly emerging for chemical analysis and medical applications. To better design and fabricate these microdevices with increased reliability, a knowledge of the fluid flow behavior in this micron sized regime is crucial. However, an accurate model of flow behavior in the micron sized channels of these devices does not exist.

Most experimental studies of fluid transport have been limited to measurements of flow rates and pressure variations in flow systems or 2-D examination techniques such as optical microscopy. For example, J.N. Zemel and coworkers [1-3] have investigated flow in channels etched in Si ranging in area from 80 to 7200 square microns. In these experiments, the flow rate and the pressure drop were measured while varying the inlet pressure with a regulating valve. In addition, an optical microscope was used for direct observation of the fluid through the upper face of the channel i.e., a Pyrex glass cover bonded to the Si.

Our initial investigations have focused on the direct examination of the 3-D internal microstructure and microdynamics in microsystems using a combination of Exxon's x-ray microtomography capabilities, and Sarnoff's computer vision, visualization and supercomputing technologies. We have imaged and visualized the fluid streaming microscopically through various fluidic microvalves at high resolution. In addition, we have begun to develop a technique to visualize and measure steady-state flow in various micron-sized capillary tubes. Finally, we have developed vision-based fluid flow recovery algorithms by incorporating physical and geometrical constraints to retrieve information about the physical flow of fluids from corresponding video imagery.

METHODS

X-Ray Imaging of Microsystems

Exxon has established an x-ray microtomography/ microradiography facility at the National Synchrotron Light Source (NSLS), Brookhaven National Laboratories. This facility produces 3-D images of the internal structures of millimeter sized samples with resolution approaching 1 micron [4,5]. The synchrotron provides an intense, monochromatic, highly collimated x-ray beam with energies ranging from 5-30 keV. High quality microradiographs are generated with short exposure times of < 1 sec. The samples are uniformly illuminated by the x-ray beam. The wavelength is chosen so that the samples absorb approximately 80% of the incident x-rays. Differences in the x-ray absorption of materials form the basis of the radiography experiments. The phosphor screen converts x-ray light to visible light output for the CCD imager. A high power optical microscope objective is used to magnify the radiographic image formed in the phosphor onto the CCD camera. The digital image data is transferred to and stored in memory of a Macintosh computer.

Microtomography- Imaging Micro-Structure: This technique provides non-destructive, 3-D static characterization of internal microstructure of microsystems at high resolution. A typical experiment involves the collection of over one thousand radiographs or projections as a function of angle. These projections are reconstructed by a Direct Fourier Inversion algorithm to form 3-D opacity maps. Exxon's current reconstruction algorithm depends upon the sample being entirely covered by the x-ray beam throughout a 360 degree rotation.

Microradiography- Imaging Micro-Motion: This technique complements microtomography in that it provides information about the dynamics of the displacement process. Digital subtraction microradiography produces 2-D projections of the 3-D fluid flow through the channels where the background structure of the sample has been digitally removed from the image. Fluids are typically injected into a micro device by a syringe pump operating at a constant flow rate. The experiment proceeds by taking "pictures" (radiographs) approximately every second until enough frames are collected to visualize the displacement. Several hundred, 512 x 512 frames are accumulated. The sequence of radiographs are processed and then animated to create a "flow video" of the fluid moving through the device.

Visualization

Since the above instrumentation generates megabytes of data every few seconds, new computational methods for displaying, manipulating and extracting information from volumetric data sets and radiographs are

0-9640024-1-8 $3.00 © 1996 TRF

Solid-State Sensor and Actuator Workshop
Hilton Head, South Carolina, June 2-6, 1996

needed. Sarnoff has developed the Princeton Engine (PE) [6], a massively parallel supercomputer. Volume rendering algorithms have been implemented on the PE which enable the truly interactive volume visualization of the large, 3-D data sets collected at Brookhaven. These 3-D images, are computed and displayed in real-time, i.e., at multiple frames/sec [7]. We have also employed Silicon Graphics (SGI) hardware and software to display, clarify and enhance the flow data. Sample applications include simple viewers to display image sequences and volume data sets as well as various image processing and physical modeling tools.

Motion Analysis

A large body of work in image processing and computer vision deals with the extraction of motion information from image sequences [8]. Motion algorithms allow for the automatic detection and tracking of specific features of interest as well as for the estimation of a flow field in a given image sequence. Interestingly, relatively little has been done to make use of relevant physical and geometric constraints in applying image motion algorithms to specific domains of interest (e.g., the recovery of physical fluid flow) [9]. In the following paragraphs, we discuss image motion analysis, and how it is utilized to recover velocity flow fields.

As a point of departure, we are concentrating on the recovery of 2-D flow from pairs of temporally related 2-D images. The recovery of flow is a non-trivial matter due to local ambiguities in the apparent displacement of image intensities and the effects of noise [10]. Our approach to overcoming these matters is to constrain the flow recovery process by making use of physical and geometric constraints that are derived from fluid mechanics.

Currently, we are making use of three physically and/or geometrically motivated constraints on the flow recovery process. The first constraint enforces local rigidity of the recovered flow (i.e., translation and rotation of infinitesimal material particles). This is done by minimizing the total temporal derivative of imaged intensity, with intensity taken as a measurement of density. The second constraint enforces local smoothness of the recovered flow. This is done by minimizing spatial gradients in the recovered flow. The smoothness constraint is particularly important in ameliorating the effects of imaging noise. The third constraint concerns boundary conditions on the recovered flow. In particular, we enforce the boundary condition that at the interface of the fluid and its containing channel there can be no flow component orthogonal to the channel wall. Following [11] we combine these constraints and appeal to the calculus of variations [12] for a solution.

Algorithmically, we have embodied this iterative solution in a Jacobi relaxation scheme [13] and embedded it in a coarse-to-fine pyramid estimation framework [14]. The pyramid component allows initial estimates to be computed at low spatial resolution and used as initial conditions for further refinement at successively increased levels of resolution. Making use of coarse-to-fine refinement increases the range of motions that can be recovered even while leading to greater computational efficiency.

RESULTS

We have applied the above mentioned x-ray imaging technologies in an investigation of micro-structure and micro-motion in representative microsystems e.g. fluidic microvalves and micron sized capillary tubes. This technology was demonstrated using three example microsystems described below. Note: The volume data sets and dynamic image sequences are best illustrated on video tape.

Redwood's Normally Open (NO) Microvalve

The Fluistor microvalve, designed and fabricated by Redwood Microsystems [15], has been imaged using x-ray microtomography. A cavity, etched in a Si substrate, is filled with a control liquid. When the liquid is heated, the Si membrane flexes outward toward the valve seat. The valve contains channels and holes designed to regulate fluid transport. We have worked with several such valves including a Normally Open (NO) gas microvalve and a NO liquid microvalve. Figure 1 shows a snapshot of the 3-D microstructure of the NO gas microvalve. After rescaling the volume in terms of color and opacity, features of interest or selected materials can be identified inside the microdevice. For example, the silicon substrate and pyrex glass in Figure 1 were set to be transparent for easy presentation of the inner structures.

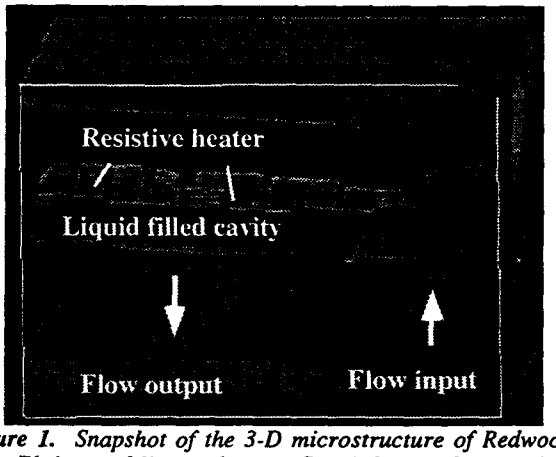


Figure 1. Snapshot of the 3-D microstructure of Redwood's NO Fluistor Microvalve. Spatial resolution is 6 microns/voxel.

The above microtomography experiments were coupled with a measurement of mechanical motion in this valve by using microradiography. For example, an image sequence of the NO gas microvalve opening and closing was collected. We have observed that when the power is switched on, the membrane quickly snaps outward towards the valve seat or orifice. When the power is switched off, the membrane slowly relaxes back to its original position. The boiling of the control liquid in the cavity is also apparent in the radiographs. A microradiograph of the valve in the closed position is shown in Figure 2.

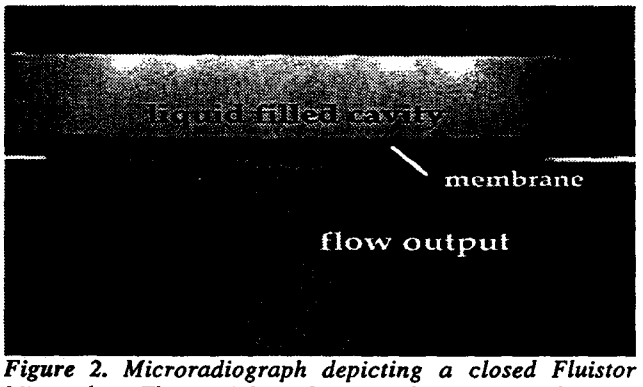


Figure 2. Microradiograph depicting a closed Fluistor Microvalve. The spatial resolution is 6 microns/pixel.

Digital subtraction microradiography, as described above, was also utilized to directly examine fluid transport. For example, we have investigated the initial filling of

Redwood's Normally Open liquid microvalve with iododecane. Iododecane was injected into the valve at a constant flow rate of 0.12μl/min. We have observed the liquid initially wetting the space (air gap) between the silicon membrane and the Pyrex. Only after complete filling of the air gap, does the fluid proceed through the output port. We can clearly see the onset of an air bubble forming. Three microradiographs selected from this image sequence are shown in Figure 3. The arrows indicate the direction of flow through the valve. Bubble formation and dead space were prevalent during these experiments.

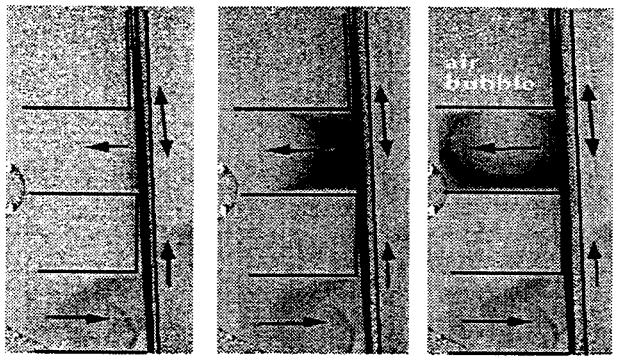


Figure 3. Microradiographs depicting the initial filling of Microvalve (after digital subtraction of background structure). The arrows indicate the direction of flow into the input port, through the air gap and finally leaving the output port. The spatial resolution is 6 microns/pixel.

Redwood's Normally Closed (NC) Microvalve

Microradiographs of an open (hot) and a closed (cold) NC valve are shown in Figure 4a and 4b, respectively. In this valve, the orifice is off-center with respect to the cavity and the membrane includes a fulcrum point. As the liquid is heated, the membrane pushes outward, causing the upper pyrex to rotate away from the orifice, allowing flow to occur. We have collected microradiographs of the NC valve during actuation and under varying inlet pressures. We have observed that increasing inlet pressure causes the upper pyrex to rotate towards the closed position, even though the valve is completely open (hot).

Micro Pipe Flow

We have begun to develop a technique to provide the simultaneous visualization and measurement of velocity flow field in microsystems. This technique can be utilized for steady state and unsteady state flows at the micron level. To accomplish this, an x-ray detectable contrast medium is first mixed with a fluid, e.g., oil mixed with water. After emulsification, the contrast medium is dispersed in the fluid in the form of fine droplets (diameter 1~20μm). Since the density of the contrast medium is very close to the density of water, millions of the fine droplets are neutrally buoyant and suspended in the flow field. The droplets follow the motion of the fluid in a micro channel and are directly viewable with x-ray imaging methods. Several image sequences were collected using this technique to measure steady state flow of a liquid in quartz capillary tubes of various sizes. In these particular experiments, fluid was injected at a constant flow rate from 0.004 to 0.008 microliters/sec. The qualitative flow field can be visualized by the motion of the droplets. In addition, the fluid velocity at any point can also be derived from the droplet movement between successive images and the known time interval between the images.

One image from a microradiograph pair is shown in the left panel of Figure 5. This image corresponds to an 840 micron quartz capillary tube with an emulsion solution being driven through at 0.004 microliters/sec captured at 500 msec exposure time. Examination of the image pair by eye suggests that in the (projected) center of the tube the emulsion droplets are displacing by approximately 10 pixels with a smooth transition to almost zero displacement at the (projected) left and right boundaries of the tube. The overall direction of motion is approximately vertical, with somewhat more directional variation in the center of the tube than at the sides. The right panel of Figure 5 shows the flow field recovered with our variational flow recovery algorithm.

The experimental velocity profiles across the channel at different flow rates, 0.004 and 0.008 μl/sec, are plotted in Figure 6. These profiles were compared with the theoretical average velocity profile of fully developed circular pipe flow at the same flow rates and show reasonable agreement[16]. The recovered velocity profiles in 1000, 840, 750 and 640 micron diameter capillary tubes were collapsed into a non dimensional velocity profile, as shown in Figure 7. The radius of the capillary tube, R, and the maximum velocity at centerline, Umax are used to non-dimensionalize the plot. A fully developed pipe flow velocity profile is also superimposed as a solid line in the same plot. We find that this theoretical profile fits the recovered velocity profiles very well. This means that the flow within our test range is still dominated by viscous and pressure forces.

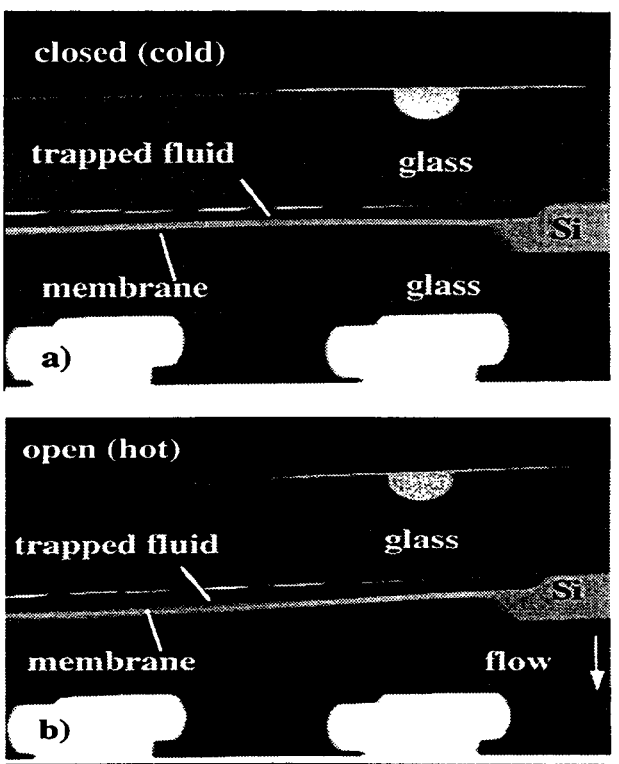


Figure 4. Microradiographs depicting Normally Closed Gas Microvalve a) closed (cold) and b) open (hot). The spatial resolution is 6 microns/pixel.

CONCLUSIONS

We have developed technologies which have allowed us to image, visualize and analyze micro-structure and micro-motion in fluidic microsystems. We have imaged and visualized the 3-D micro-structure and 2-D micro-fluidic motion in representative microdevices such as Redwood Microsystems' Normally Opened and Normally Closed Gas

Microvalves as well as their new Liquid Microvalve using x-ray microtomography and a massively parallel supercomputer, the Princeton Engine. We found that this technology can be used to verify the functionality of a particular valve as well as detect problem areas such as air bubbles and dead space. Finally, we have developed a method to image and visualize steady state fluid transport in various micron-sized quartz capillary tubes using emulsions. We have measured velocity profiles in 1000, 840, 750 and 640 micron inner diameter capillary tubes and found that they are in reasonably good agreement with the theoretical average velocity profiles of fully developed circular pipe flow.

ACKNOWLEDGMENTS

The authors wish to thank Dr. Mark Zdeblick and Ms. Robin Miles at Redwood Microsystems for providing the microvalves used in this work as well as technical assistance and helpful discussions. We also acknowledge with thanks the support of this work by DARPA/ETO under Contract No. DABT63-95-C-0057.

REFERENCES

1. J. Pfahler, J.Harley, H. Bau, and J. Zemel, *Sensors and Actuators*, A21-A23 (1990) 431-434.
2. L. Kricka, P. Wilding, J. Pfahler, J.Harley, H. Bau, and J. Zemel, *SPIE* 1167 Precision Engineering and Optomechanics (1989) 159.
3. J.Harley, H. Bau, J. Zemel, and V. Dominko, *IEEE* (1989) 25.
4. B.P. Flannery, H.W. Deckman, W.G. Roberge, and K.L. D'Amico, *Science*, 237 (1987) 1389-1544.
5. J.H. Dunsmuir, S.R. Ferguson, K.L. D'Amico, B.P. Flannery, and H.W. Deckman, *Review of Progress in Quantitative Nondestructive Evaluation*, 10A (1991) 443.
6. D. Chin, J. Passe, F. Bernard, H. Taylor, and S. Knight, "The Princeton Engine: A Real-time Video System Simulator", *IEEE Transactions on Consumer Electronics, 342*, 1988, pp. 285-297.
7. J. Kaba, J. Matey, G. Stoll, H. Taylor, and P. Hanrahan, "Interactive Terrain and Volume Rendering on the Princeton Engine", *Proceedings of Scientific Visualization*, September 1992, pp. 349-355.
8. J. K. Aggarwal and N. Nandhakumar, "On the computation of motion from sequences of images -- A review", Proc. IEEE 76 (8), (1988), 917--935.
9. A. Del Bimbo, P. Nesi, and J. L. C. Sanz, "Analysis of optical flow constraints", *IEEE Transactions on Image Processing (4) 4*, (1995), 460--469.
10. B. K. P. Horn, *Robot Vision*, MIT Press, Cambridge, MA (1986).
11. B. K. P. Horn and B. G. Schunk, "Determining optical flow", *Artificial Intelligence 17 (1)*, (1981), 185--203.
12. R. Courant and D. Hilbert, *Methods of Mathematical Physics, Vol. I*, John Wiley and Sons, New York, (1953).
13. W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, *Numerical Recipes*, Cambridge University Press, New York, New York, (1988).
14. J. R. Bergen, P. Anandan, K. Hanna and R. Hingorani, "Hierarchical model-based motion estimation" *Proc. ECCV*, (1991), 5--10.
15. M.J. Zdeblick, R. Anderson, J. Jankowski, B. Kline-Schoder, L.Christel, R. Miles, W. Weber, " Thermpneumatically activated microvalves and integratedelectro-fluidic circuits," *Solid State Sensor and Actuator Workshop, Hilton Head, South Carolina, June 13-16, 1994*, p 251.
16. H. Schlichting, "Boundary Layer Theory" *Mcgraw Hill Book Company, seventh edition*, (1979), p85.

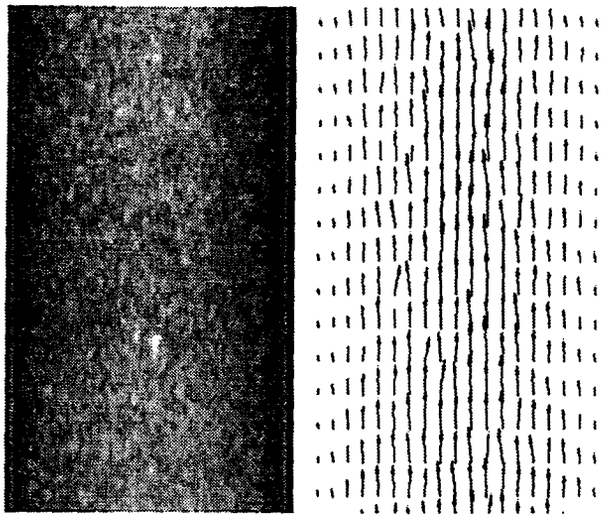


Figure 5. The left panel shows one image from an input image pair. The right panel shows the flow field recovered by our variational flow recovery algorithm. The channel diameter is 840 microns. The spatial resolution is 3 microns/pixel.

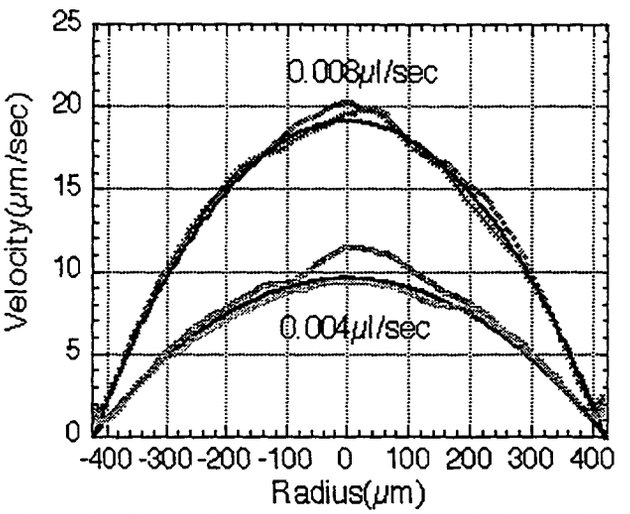


Figure 6. Comparison of experimental velocity profiles with theoretical (solid line) fully developed pipe flow average velocity profile

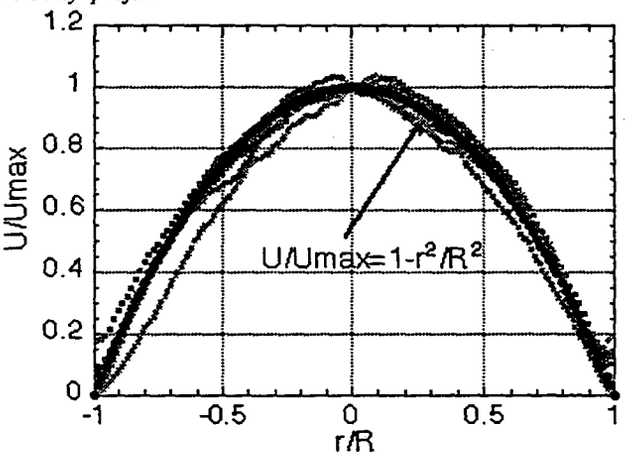


Figure 7. Nondimensional average velocity profiles experimental versus theoretical (solid line)

MICROMACHINED THERMALLY ISOLATED CIRCUITS

Erno H. Klaassen[*], Richard J. Reay[**], Christopher Storment[*], Jonathan Audy[***], Peter Henry[***], A. Paul Brokaw[***], and Gregory T.A. Kovacs[*]

[*] Center for Integrated Systems, CIS-X 202, Stanford University, Stanford, CA 94305-4075
[**] IBM Almaden Research Center, San Jose, CA 95120
[***] Analog Devices Inc., Santa Clara, CA 95052

ABSTRACT

This paper details a post-process technique by which circuitry in an unmodified IC technology is thermally and electrically isolated from the silicon substrate. This method enables new applications for micromachining, including temperature regulation of analog integrated circuits. The process will be discussed in detail, along with improved tetramethyl ammonium hydroxide (TMAH) etching chemistries that use strong oxidizers to eliminate hillock formation. Also presented is an electrochemical biasing method that uses circuitry on the silicon being etched during the micromachining step. Work is currently underway to evaluate the use of these micromachining techniques for commercial analog circuit applications, several of which will be highlighted in this presentation.

INTRODUCTION

Silicon has a thermal conductivity close to that of aluminum (1.57 W/cm·K and 2.36 W/cm·K, respectively). This material property is highly beneficial for most circuits, since it allows heat generated in devices to be efficiently conducted into the package. There are applications, however, in which it is necessary to generate or detect precise, localized changes in temperature on the surface of a silicon chip. In order to do this without dissipating large amounts of power, the thermally conductive path through the silicon must be selectively removed so that peripheral circuitry still benefits from efficient cooling.

Techniques referred to as CMOS post-processing have been used to isolate regions of dielectric, inside which the interconnect layers in the process, usually aluminum and polysilicon, can be thermally isolated to form heating resistors and thermopiles. These devices can be processed using a front-side anisotropic undercut etch to create cantilevers [1] or a backside etch to form diaphragms [2].

A much wider range of applications is made possible with a new method that allows active devices to be thermally isolated. This is achieved by suspending silicon at the end of dielectric support beams. Entire circuits, or critical portions of circuits, can now be thermally isolated. These circuits can contain, depending on the IC process used, MOS and bipolar transistors, diffused resistors, pn diodes, and polysilicon resistors and capacitors. Thermally isolated circuits can be used for a wide range of applications involving temperature measurement of these suspended devices and low-power temperature regulation of thermally isolated circuitry.

CIRCUIT ISOLATION PROCESS

Figure 1 shows the post-processing fabrication method. Portions of the silicon substrate are exposed by superimposing the oxide cuts in the process [3]. Each oxide layer that is grown or deposited is removed in a subsequent etching step in these "open" regions. One or more n-wells are located at the ends of support beams defined by the exposed regions, as shown in Figure 1(a).

The wells are biased in tetramethyl ammonium hydroxide, an anisotropic silicon etchant, at -0.8V relative to an Ag / AgCl electrode using a potentiostat. The p-type substrate is maintained near the open circuit potential at -1.5V. A platinum counter electrode is used in this four-electrode potentiostat configuration [4]. The n-wells are passivated, while the p-substrate etches away in the typical anisotropic fashion. The substrate underneath the well is etched primarily along (110) crystal planes.

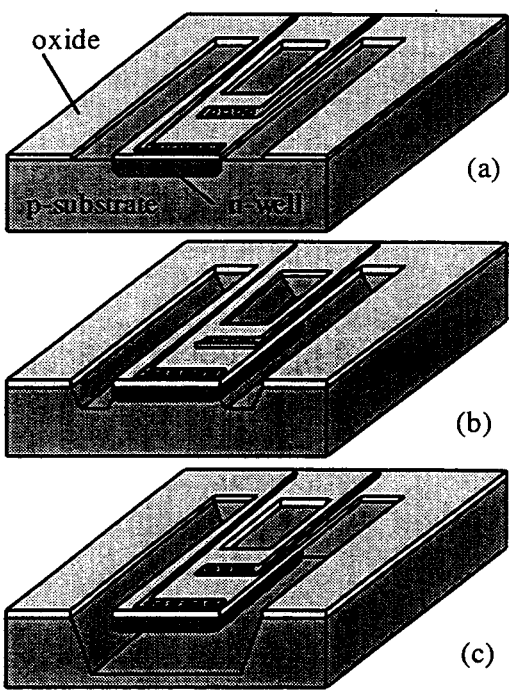


Figure 1. Circuit undercutting process, upon completion of CMOS processing (a), beginning of anisotropic etch (b), and after release of the micromachined circuit (c).

The resulting structure, shown in Figure 1(c), contains a region of single crystal silicon suspended by dielectric support beams. Aluminum interconnects are embedded inside these oxide layers. Since the thermal conductivity of silicon dioxide and

0-9640024-1-8 $3.00

itride is much lower than that of aluminum, the only significant thermal path through the supports is formed by these metal traces. This type of micromachined element provides circuitry with high thermal resistance to the substrate.

IMPROVED TMAH ETCH CHEMISTRY

The masking layer for the post-processing etch is the oxide or nitride passivation layer of the circuit process. The aluminum bond-pads are exposed during the etch. It is therefore necessary to use a silicon etchant with high oxide, nitride, and aluminum selectivity. Ethylene-diamene pyrocatechol (EDP) could be used as a silicon etchant, as certain formulations of this chemical have been shown to offer the required selectivity [5]. EDP however, is toxic and mutagenic. Similar safety concerns, in addition to explosion hazards, apply to hydrazine silicon etchants. Tetramethyl ammonium hydroxide (TMAH), another anisotropic etchant, also has very high oxide and nitride selectivity, and exhibits selectivity to aluminum if the solution is doped with silicon [6] or silicic acid ($Si(OH)_4$) [7]. TMAH is relatively safe to use, and presents no special disposal issues. The addition of silicates to the etchant acts in two ways to passivate the aluminum from attack by the alkaline etchant. The pH of the solution is reduced, but more importantly, the presence of silicates in solution will form aluminosilicates that are less soluble at moderate pH (between 8 and 12) than the hydrated aluminum oxides that would otherwise form at the metal surface. In the packaging industry, silicates are commonly used when aluminum is put into contact with alkali solutions of soaps or other cleaning agents [8]. The required amount of dissolved silicates is proportional to the concentration of the TMAH etching solution. In order to limit the amounts of silicon or silicic acid required in solution, a five weight percent TMAH solution was chosen. The passivation of aluminum requires at least 16 g/l of dissolved silicon or 40 g/l of dissolved silicic acid for this concentration.

One troublesome property of TMAH is the formation of hillocks on (100) silicon for pH levels lower than 13 [9]. When these small pyramids form, the silicon etch rate drops dramatically. Figure 2 is a SEM photograph of a hillocked (100) p-type silicon surface. The mask was 100. nm of thermal SiO_2, and the etchant was 5 wt. % TMAH with 16 g/l of dissolved silicon at 80 °C.

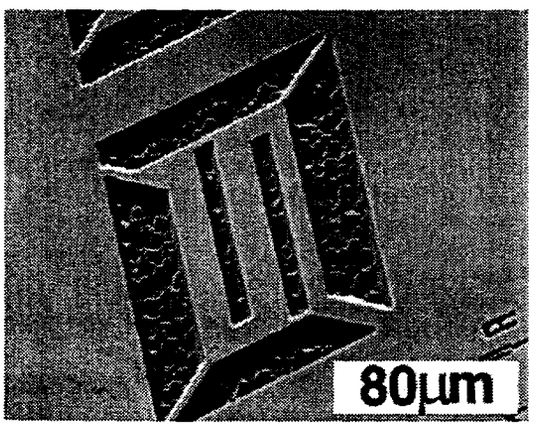


Figure 2. SEM photo of hillock formation on (100) silicon after etching in 5 wt. % TMAH and 16 g/l of dissolved silicon at 80 °C.

The alkaline silicon etching can be broken down into two separate steps and written as

$$Si + 2OH^- \rightarrow Si(OH)_2^{2+} + 4e^- \quad [1]$$

$$Si(OH)_2^{2+} + 4e^- + 4H_2O \rightarrow Si(OH)_6^{2-} + 2H_2 \quad [2]$$

In the first step, in Eqn. 1, electrons are injected into the conduction band of the silicon. These are used up in the reduction step in Eqn. 2. Hydroxyl ions from the water bind to the silicate formed in the first step, and produce soluble silicic acid, with hydrogen gas as a byproduct. Hillock formation is likely due to local micro-masking of the silicon. If small hydrogen gas bubbles cling to the surface before they release, they temporarily mask the surface underneath from the etching solution. Hillocks form beneath these adhered bubbles. The bubbles block the free diffusion of reactants and products to and from the surface. A precipitate will form at the point of contact between the bubble and the silicon that creates a temporary masking layer as it is slowly dissolved upon release of the hydrogen bubble. This localized mask is composed of partially hydrated silicate species with variable stoichiometries. Initially the number of pyramids at the surface is small, but as etching continues new pyramids are formed and begin to superimpose over previous hillocks. Once the surface is fully covered with hillocks, the (100) etch rate will drop precipitously.

For increasing TMAH concentrations, the silicon etch rate decreases, and hillocking is reduced. This dependence on concentration, which occurs above two weight percent, is based on two key factors. First, as the concentration of water drops with increasing concentration, the formation of progressively less hydrated and less soluble silicate species reduces the rate of product removal from the surface. This effectively blocks the surface reduction reaction, until the products being solvated are removed. With further increases in concentration, the surface becomes more uniformly covered with these low solubility silicate species. The slightly soluble silicates that were located beneath bubbles for low concentration solutions are now spread uniformly across the entire surface. The localized blocking effects are no longer important, and hillocking is reduced. The second factor is that the amount of water available for reduction by the silicon (Eqn. 2) is lower. This reduces the silicon etch rate, which in turn drops the formation rate and size of the bubbles clinging to the surface.

Palik, *et al.* [10] observed that when silicon is etched in KOH, the average size of hydrogen bubbles diminishes with increasing KOH concentration, while the areal density of these bubbles goes up. They theorize that since smaller bubbles have a smaller "footprint" and release more quickly, they leave the surface less rough than larger bubbles with a longer "dwell time." The wetting properties of the solution may therefore in part determine the appearance of an etched surface. Surfactants could be added to TMAH to change the wetting angle of the solution on the silicon interface. This could keep bubbles from creating hillocks by reducing the dwell time. TMAH, however, is very sensitive to contamination from organic sources. The addition of a fluorochemical surfactant, FC-129 [11], to TMAH solutions caused silicon etching to stop. Ultrasonic agitation of the solution could also shorten the dwell time of adhered hydrogen. The agitation could, however, destroy micromachined devices. Another solution to the hillocking problem is therefore called for.

The addition of pyrazine to TMAH solutions has been shown to eliminate hillock formation [2]. This benefit did not apply to silicon-doped etching solutions, however. Certain potassium hydroxide (KOH) - alcohol etchants also cause hillock formation. It has been reported that etching silicon in a KOH-isopropanol solution saturated with oxygen eliminates hillock formation [12]. This method may yield similar benefits for TMAH

solutions, but working with oxygen under high pressure presents an explosion hazard.

Alkaline silicon etches can produce smooth (100) surfaces if oxidizers are added to the etchant [6,13]. When an oxidizer with a higher redox potential than water is added to the etching solution, hydrogen gas evolution can be avoided. The oxidizer takes the place of water in Eqn. 2. If the byproducts of the solution are aqueous, no micro-masking by gas bubbles occurs and no hillocks form. This has been demonstrated with hydrogen peroxide in ammonium hydroxide, but the oxidizer concentration needs to be very tightly controlled [14]. If insufficient peroxide is added, hillock formation persists, but if too much is introduced, silicon etching ceases since the hydrogen peroxide rapidly reduces the pH of the etching solution. The required range was reported to be between 0.65 x 10^{-2} M and 1.84 x 10^{-2} M.

Peroxydisulfates, another group of inorganic peroxides, are some of the strongest known oxidizers in aqueous solution [15]. This group includes potassium peroxydisulfate ($K_2S_2O_8$) and ammonium peroxydisulfate ($(NH_4)_2S_2O_8$). Peroxydisulfate is an even stronger oxidizer than peroxide. The half-cell reaction for the reduction of hydrogen peroxide can be written (relative to a standard hydrogen electrode, with 1 M concentrations of each species) as

$$H_2O_2 + 2H^+ + 2e^- \leftrightarrow 2H_2O \qquad [3]$$

and has a redox potential of E^0=1.776 V [16]. The redox reaction for the peroxydisulfate ion is

$$S_2O_8^{2-} + 2e^- \leftrightarrow 2SO_4^{2-} \qquad [4]$$

which has a redox potential of E^0=2.01 V.

Adding 5 to 10 grams per liter of either potassium or ammonium peroxydisulfate to TMAH solutions, even ones doped with silicon or silicic acid, eliminates hillock formation. In contrast to hydrogen peroxide, the exact peroxydisulfate concentration is not critical. The addition of excess oxidizer does not stop, or even slow, the silicon etch rate.

Figure 3 shows the (100) silicon etch rate for a 5 wt. % TMAH solution for several different temperatures. One set of data points represents the silicon etch rate with a pure solution, while the second shows the etch rate with the addition of 5 grams per liter of ammonium peroxydisulfate. The surface of the silicon etched in the pure TMAH-water solution was hillocked, while that for the solution with added oxidizer was smooth. The data was taken for 5-10 Ω·cm p-type silicon, using a 100 nm thermal oxide mask.

Figure 4 shows SEM photos of two sets of elements used in a thermal RMS converter [17]. The devices in Figure 4(a) were etched in a 5 wt. % TMAH solution containing 16 g/l of dissolved silicon. While the aluminum bond-pads were not attacked by the etchant, there is significant hillocking at the bottom of the cavity. The devices are thermally isolated from the substrate, but the cavity is only 45-50 μm deep. The photo in Figure 4(b), however, shows an identical set of thermo-elements etched in a solution containing 5 g/l of ammonium peroxydisulfate. The cavities under these devices have smooth bottoms (not visible) and are approximately 125 μm deep. The etch time for each of these sets of devices was approximately 150 minutes at 80 °C.

Silicon doped TMAH solutions, with or without added oxidizers, have been found to effectively etch silicon for about one day. The pH of the solution drops to unacceptably low levels after 6 to 8 hours. The time will be shorter if large volumes of silicon are etched during use. This effect could be mitigated by appropriately buffering the pH level of the solution. Even without this added measure, however, the low cost of the solution and ease of disposal make it feasible to use a fresh solution for each day of use.

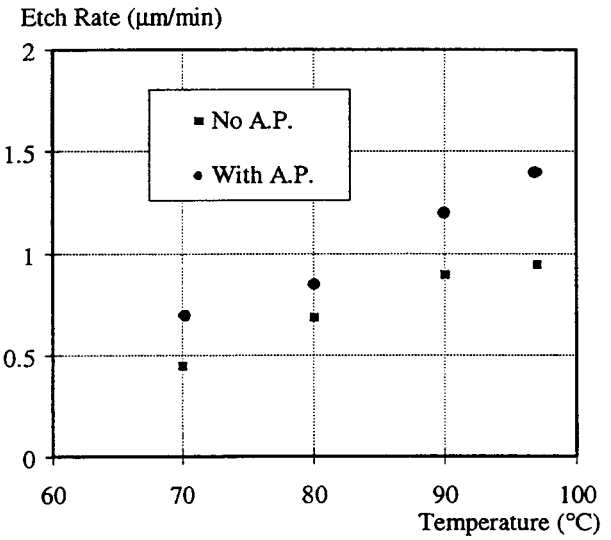
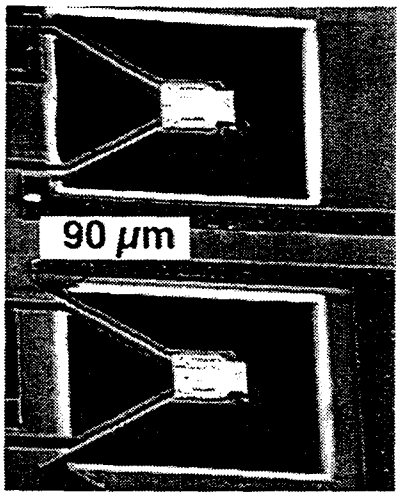


Figure 3. (100) silicon etch rate at several temperatures for 5 wt. % TMAH with and without added ammonium peroxydisulfate. No dissolved silicates were added to these solutions.

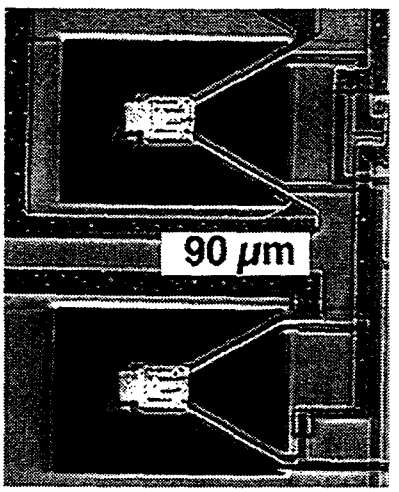
AUTO-BIASED ETCHING

Since the etch is a post-process step, the circuitry built during the CMOS process is functional during etching. Portions of these circuits can play an active role during the micromachining etch step. One application for this is the distribution of the passivation potential to multiple n-wells. Figure 5 shows a device configuration that achieves this. In this case, each element contains a suspended diode, which are series connected, as shown. Many other configuration are of course also possible. Each of the n-wells must be maintained at the passivation potential during the undercutting etch. To prevent having to short-circuit the wells together permanently, a series of transmission gates are used. The passivation potential is applied to the lower signal line, labeled V_{pp}. In solution this represents a voltage of -0.8 V relative to the reference electrode. A control signal, V_{etch} is then applied to turn the NMOS transmission gates on. This voltage must be more than one threshold voltage greater than the passivation potential. With the transmission gates turned on, each n-well shown is electrically connected to the passivation potential. Following the etch, the transmission gates can simply be turned off by keeping the two lines at the lowest potential on the chip (V_{ss}). The n-wells are then electrically disconnected from each other.

This concept is demonstrated by the 70 element array of test devices shown in Figure 6. The transmission gates are located between each row of elements. The passivation potential was introduced at a single point, and distributed by the transmission gates to each individual element.



(a)



(b)

Figure 4. Two SEM photos of thermal RMS converter elements. The cavities under the devices in (a) are covered with hillocks, while those in (b) are smooth, due to the addition of peroxydisulfates to the silicon-doped TMAH solution.

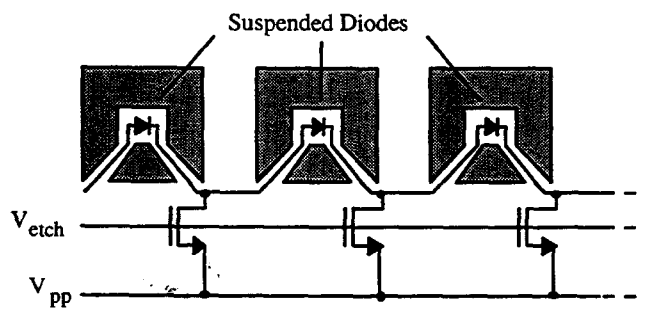


Figure 5. Schematic of a configuration that employs on-chip circuitry during the micromachining etch step.

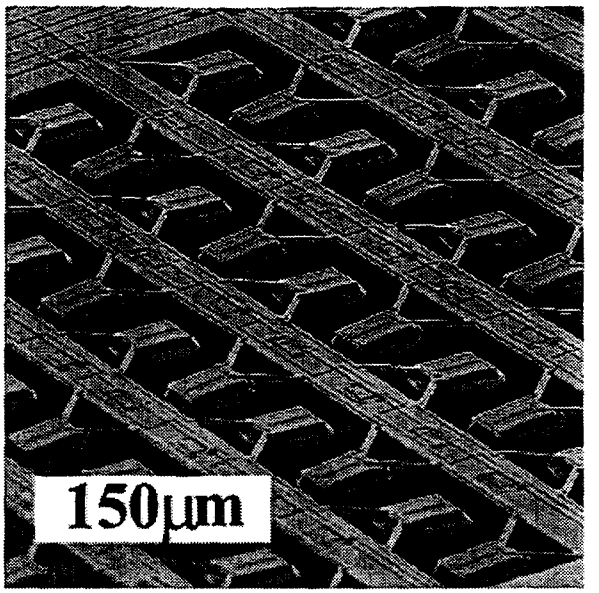


Figure 6. An array of 70 suspended n-wells fabricated using auto-biased etching.

APPLICATIONS AND RESULTS

The silicon undercutting technique has been employed on the Mosis 2.0 μm Analog CMOS process, as well as on an Analog Devices BiCMOS process. Thermal isolation up to 50,000 K/W has been achieved.

The technology presented here has been used to fabricate a thermal AC to RMS converter with a 60 dB dynamic range. Figure 4 is an SEM of a pair of thermoelements used in this RMS converter. Each of the matched elements contain a polysilicon resistor and a diode [17]. The resistor is used as a heater, while the diode is used under forward-bias as a temperature sensor. One of the two elements is heated by the AC signal. The second element is heated with a DC signal by feedback circuitry, such that the two elements are maintained at equal temperatures. The DC signal is the RMS value of the AC signal. The control circuits for the RMS converter system were integrated on-chip, alongside the micromachined elements. The converter has a bandwidth in excess of 400 MHz and a nonlinearity better than 1%.

Another application for this thermal isolation method is temperature-regulation of analog circuits. This technique can be used to mitigate the effects of ambient temperature changes on precision circuitry, such as voltage references. A bandgap voltage reference that employs such temperature regulation was built [18]. Figure 7 is an SEM of such a suspended circuit. The device contains PMOS heating transistors, which, together with a feedback network, maintain the silicon underneath the rectangular element at a constant temperature. The output voltage drift of the device was reduced by a factor of 40, to 9 ppm/°C, by using temperature regulation. The high thermal isolation of the circuits enabled a power dissipation of only 2mW. Work is currently underway to further improve the performance of a micromachined voltage reference.

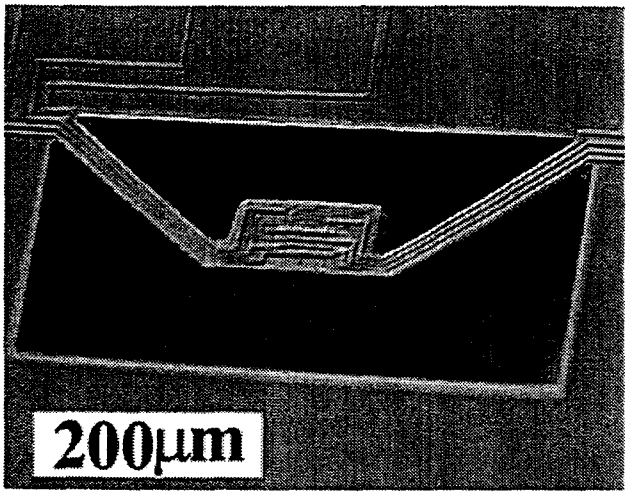


Figure 7. SEM of a micromachined bandgap voltage reference.

CONCLUSIONS AND OUTLOOK

We have presented a new IC post-processing method that requires no additional masking steps beyond those used for the circuit fabrication. The technique is capable of providing circuitry with thermal isolation as high as 50,000 K/W at atmospheric pressure. An improved TMAH etching chemistry uses strong oxidizers to eliminate hillock formation. The resulting etchant provides good silicon etch rates, and exhibits high selectivity to oxide, nitride, and aluminum. Since the undercutting etch is performed as a post-process step, finished circuitry from the CMOS process can play an active role during the micromachining operation. A method to use transmission gates for the distribution of the electrochemical bias was described. In this fashion, many separate islands of suspended silicon can be realized. Two applications of the technology, an RMS converter and a temperature-regulated circuit, were discussed as examples. The micromachined bandgap voltage reference demonstrates a use for micromachining in improving the performance of integrated circuits.

To date, the process has been sucessfully implemented using two different IC technologies: a 2.0 µm n-well analog CMOS process offered through MOSIS, and an Analog Devices BiCMOS line.

REFERENCES

1. R. Lenggenhager, H. Baltes, J. Peer, and M. Forster, "Thermoelectric infrared sensors by CMOS technology," IEEE Electron Device Letters, No. 13, pp. 454-456, 1992.

2. R. Lenggenhager, D. Jaeggi, P. Malcovati, H. Duran, H. Baltes, E. Doering, "CMOS membrane infrared sensors and improved TMAHW etchant," In Tech. Dig., IEEE Intl Elect. Dev. Meeting, San Francisco, CA, Dec. 11-14, 1994, pp. 531-534.

3. M. Parameswaran, H. Baltes, L. Ristic, A. Dhaded, and A. Robinson, "A new approach for the fabrication of micromechanical structures," Sensors and Actuators, vol. 19, pp. 289-307, 1989.

4. B. Kloeck, S. D. Collins, N. F. de Rooij, and R. L. Smith, "Study of electrochemical etch-stop for high-precision thickness control of silicon membranes," IEEE Trans. Electron Devices, vol. 36, no. 4, pp. 663-669, 1989.

5. D. Moser, "CMOS flow sensors," Ph.D. Dissertation, Physical Electronics Laboratory, ETH Zurich, p. 96.

6. U. Schnakenberg, W. Benecke, P. Lange, "TMAHW etchants for silicon micromachining," In Proceedings, Transducers '91, San Francisco, CA, pp. 815-818, 1991.

7. G. Lin, B. Eyre, K. Pister, "Bulk micromachining of standard CMOS using silicic acid added to TMAH for aluminum passivation," UCLA MEMS Laboratory internal report, TRN 95-13, June 1995.

8. The Corrosion Handbook, John Wiley, New York, p. 915, 1948.

9. O. Tabata, R. Asahi, H. Funabashi, S. Sugiyama, "Anisotropic etching of silicon in $(CH_3)_4NOH$ solutions", In Proceedings, Transducers '91, San Francisco, CA, pp. 811-814, 1991.

10. E.D. Palik, O.J. Glembocki, I. Heard, P.S. Burno, L. Tenerz, "Etching roughness for (100) silicon surfaces in aqueous KOH," J. Appl. Phys., vol. 70, no. 6, pp. 3291-3300, 1991.

11. 3M Fluorad Fluorochemical Surfactants, 3M Specialty Chemicals Division, 908 North Elm St., Hinsdale, IL 60521.

12. S.A. Campbell, K. Cooper, L. Dixon, R. Earwaker, S.N. Port, and D.J. Schiffrin, "Inhibition of pyramid formation in the etching of Si p<100> in aqueous potassium hydroxide-isopropanol" J. Micromech. Microeng. nol. 5, pp. 209-218, 1995.

13. W. C. Erdman, P. F. Schmidt, "Anisotropic etching of silicon and germanium," US Pat No. 3,738,881, June 12, 1973.

14. U. Schnakenberg, W. Benecke, B. Loechel, S. Ullerich, and P. Lange, "NH_4OH based etchants for silicon micromachining: influence of additives and stability of passivation layers," Sensors and Actuators, vol. A25-27, pp. 1-7, 1991.

15. Kirk-Othmer Encyclopedia of Chemical Technology, Third Edition, John Wiley & Sons, New York, volume 17, pp. 15-19.

16 Handbook of Chemistry and Physics, 66th edition, CRC Press, 1986, p. D157.

17. E. H. Klaassen, R. J. Reay, G.T.A. Kovacs, "Diode-based thermal RMS converter with on-chip circuitry fabricated in CMOS," In Transducers '95 Tech. Dig., Stockholm, pp. 154-157, 1995.

18. R.J. Reay, E.H. Klaassen, G.T.A. Kovacs, "A low-power temperature-regulated bandgap voltage reference," In Tech. Dig., ISSCC '95, San Francisco, pp. 166-167, 1995.

Study of Screen Printed Wells in Solid-State Ion Selective Electrodes

R. W. Hower and R. B. Brown
Center for integrated circuits and sensors, University of Michigan, Ann Arbor, MI 48109

E. Malinowska
Department of Analytical Chemistry, Warsaw University of Technology, Warsaw, Poland

R. K. Meruva and M. E. Meyerhoff
Department of Chemistry, University of Michigan, Ann Arbor, MI 48109

ABSTRACT

The size of integrated ion sensors is dictated by the size and spacing of their polymeric membranes, which provide the chemical selectivity. The area occupied by an array of these membranes can be significantly reduced through the use of wells (areas separated by barrier walls, into which the membrane solutions are deposited). Screen-printable epoxies are obvious candidates for forming the wells, but their physical and chemical compatibility with polymer-based membranes has not been known. This paper describes a methodology for selecting a screen-printed epoxy well material, and compares 5 perspective materials.

INTRODUCTION

The size (and therefore cost) of solid-state ion-selective sensors is usually determined by membrane dimensions and spacing, rather than by the size of circuitry, interconnects, or bonding pads. Membrane design rules are typically dictated by the requirement to keep membranes which are selective to different chemicals from touching. We have developed membranes for automated deposition by both screen printing [1] and dispensing. In both cases, membrane components are dissolved in solvents which are evaporated subsequent to deposition. The design rules must allow for flow-out of the membrane/solvent solution after it is applied to the sensor surface, making the sensors much larger than they would otherwise need to be.

To reduce the size of sensor arrays, we pattern wells (see Figure 1) which limit the flow-out of the membrane components, allowing membranes to be smaller and closer together. Wells provide the additional advantage of making final membrane thickness more uniform, and the deposition process more tolerant of variations in the viscosity of membrane solutions. We and others have previously used epoxies [2,3], acrylic photopolymers [2], polyimide [2,4], and silicon [5], to form wells or cavities. This work focuses on screen printing of epoxy wells, which can be a reliable and economical approach for forming wells if the characteristics of the epoxy are acceptable.

In this paper we report details of our screen-printed epoxy well process, and present the methodology for selecting the best material. Data will be presented which compares prospective materials based on: their chemical compatibility with the ion-selective membranes; adhesion of the well materials to the sensor surface (which is typically excellent); adhesion of the membranes to the well materials; and screen printing characteristics (viscosity and thixotropy). The most important material property is chemical compatibility with the membrane, which means that the well material will not exude contaminants that interfere with the selectivity of the membrane, nor absorb active components from the membrane, reducing over time the magnitude of the sensor's response. Good adhesion of the membrane to the well is necessary to prevent electrolytic shunts around the membrane. When sensors have internal electrolyte layers, adhesion of the membrane to the well becomes even more important because osmotic pressure works to dislodge the membrane, and the area of adhesive contact between the membrane and sensor surface is reduced by the area of the internal layer.

Well materials for this study were selected for their screen printability and adhesion to silicon nitride. Screen masks were designed to help optimize the well dimensions.

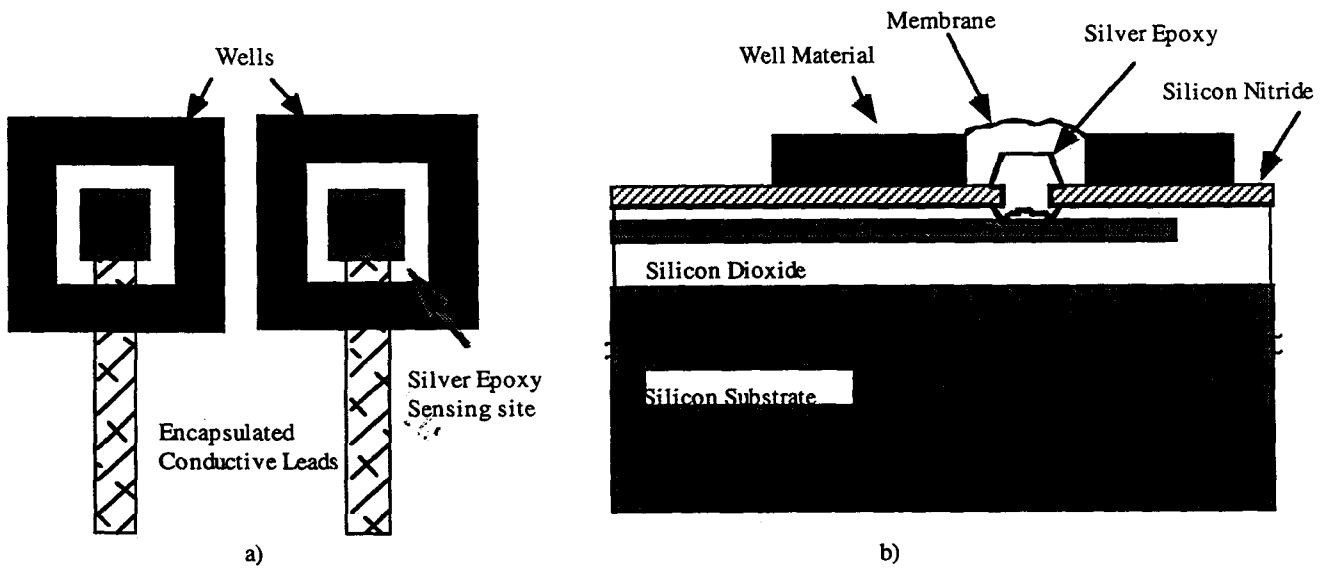


Figure 1. The well surrounding the ion-selective electrode sensing site: a) top view, b) side view.

Solid-State Sensor and Actuator Workshop
Hilton Head, South Carolina, June 2-6, 1996

EXPERIMENTAL

Reagents: Tecoflex polyurethane SG-80A (PU) was purchased from Thermedics (Woburn, MA). A terpolymer denoted as PVA (MW 40,000), which is a mixture of poly(vinyl-chloride)(80%), poly(vinyl-acetate) (15%) and poly(hydroxy propyl acrylate) (5%), was purchased from Scientific Polymer Products, Inc. (Ontario, NY). The plasticizer dioctyladipate (DOA) was purchased from Scientific Polymer Products, Inc. (Ontario, NY). The well materials used for this study were 688-PFC and 905 from Epoxy Technologies, Inc. (Billerica, MA), 3140 Silicon Rubber from Dow Corning (Midland, MI), EP3HT from Master Bond, Inc. (Hackensack, NJ), and BA-2116 from Tra-Con (Medford, MA).

Adhesion: The adhesion study was conducted by screen printing an unpatterned layer of each prospective well material onto silicon wafers, the surface of which had been coated with a silicon nitride layer. The well materials were cured, as appropriate for each material, and a scribe mark was made by sawing the wafer part way through. Membranes for the adhesion tests were formed by combining 1.06 g of DOA with 0.54 g of PU/PVA (4:1) to which was added 10 ml freshly distilled Tetrahydrofuran (THF) to dissolve the components. After 5 hours of mixing, approximately 7 ml of THF was evaporated from the solution under nitrogen, resulting in a slightly viscous solution. 500 ml of this membrane cocktail was dispensed into a 2-cm square Teflon ring, centered on the wafer's scribe lane, on the surface of the cured well material. The membranes were allowed to cure for 24 hours before the ring was removed to allow the remaining THF to evaporate. The samples were cleaved on the scribe lane, and the membranes were pulled on a Q-Test II custom adhesion tester at 1 cm/min, at approximately 90 degrees angle from one side of the wafer. Five or six samples were pulled for each well material, and averages and standard deviations were compared with those of control samples, which had membranes in direct contact with the silicon nitride.

Membrane contamination: To study the chemical compatibility of the various materials to polymeric membranes, samples of each prospective well material were cast on Teflon and cured according to their normal cure schedules. 1.0 g of the cured material was removed from the Teflon and placed in a container with 4 ml freshly distilled THF. The well material and THF were agitated for 24 hours, followed by the addition of 4 ml THF and agitation for another 24 hours. The remaining solids were removed from the THF, and it was used in the preparation of test membranes, which were mounted in conventional Phillips electrode bodies. The THF contained the extractable contaminants from the well material, which were concentrated in the resulting membranes by evaporating the solvent. This is an extreme test of chemical compatibility.

Blank standard membranes contained 33 wt.% PU/PVA (4:1) and 67 wt.% DOS. The first set of sodium membranes contain 1 wt.% Monensin Methyl Ester (MME), 30 mol-% KTpClPB, 32-33 wt.% PU/PVA (4:1) and 65-67 wt.% DOS. In later membranes, the MME was increased to 4 wt.%, with the other components decreased proportionally. Borates were used because they have been shown to decrease the anionic response in cationic selective membranes [6,7]. 3140 RTV Silicone Rubber was not tested because it is known to be chemically pure enough to form ion-selective membranes [8].

Screen-printing: The wells were printed using a 50-µm-thick, 25-µm-mesh, screen mask. Wells were formed in a variety of sizes ranging from 0.75 mm to 1.15 mm. The wells were screen printed on wafers having 50-µm-high, 300-µm-square, silver epoxy contacts centered in the wells.

Design rules: The optimal well design rules (spacing of well from silver epoxy contact and minimum wall thickness) were determined by examining screen-printed membranes in wells of various sizes. Ion-selective membranes were screen printed into the wells through 0.75 mm square openings in a 150-µm-thick stencil mask. The membranes were examined under a microscope to determine the integrity of the well walls, and whether the wells were completely filled. They were cleaved and examined with a scanning electron microscope (SEM) to evaluate the shape of the printed membranes.

RESULTS

Results of the adhesion, membrane compatibility, and screen printing experiments are shown in Table 1. Our PU/PVA membranes adhered satisfactorily to all of the well materials tested except 3140 RTV Silicone Rubber. The best adhesion was provided by the EP3HT and the 688-PFC epoxies.

Some of the blank (no ionophore) membranes and sodium membranes having little ionophore, showed anionic response; results for the blank membranes are shown in Figure 2. The top curve in this figure is for the control

	Adhesion		Chemical		
Well Material (number of adhesion samples)	Average (N)	Std. deviation	Membrane Contamination	Detection Limit	Screen printability
Control (5)	1.99	0.31	none	10^{-4}	
Tra-Con BA-2116 (6)	0.56	0.12	some anionic	10^{-3}	1
Master Bond EP3HT (5)	2.01	0.17	none	$3x10^{-4}$	1
Epo-Tek 688-PFC (5)	1.89	0.16	some anionic	10^{-3}	1
Epo-Tek 905 (6)	1.29	0.30	some anionic	no response	2
3140 RTV Silicon Rubber (6)	0.15	0.11	none	10^{-4}	3

Table 1. Experimental results for studies of adhesion, chemical contamination, and screen printability. The detection limit is for sodium selective membrane with increased ionophore. The control for the adhesion test was silicon nitride, and for the chemical tests was a membrane containing distilled THF.

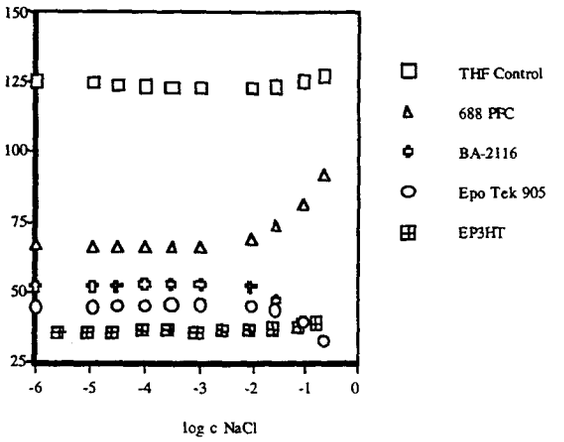


Figure 2. Anionic contamination of blank membranes in a calibration solution of 0.1 M CaCl2. The top curve is for a standard blank (no ionophore) membrane; the lower curves are for blank membranes with extracted contaminants from the materials noted in the legends.

membrane, which was made with pure (uncontaminated) THF. The contaminated membranes (bottom three curves) have significantly lower baseline potentials, and respond to anions at high concentrations. This effect can be overcome in most membranes by adding more ionophore (e.g., 4 wt.% instead of 1 wt.%). Figure 3 demonstrates this for sodium membranes; here the contamination has only increased the detection limit.

Screen printability was good to excellent, as expected, for all the well materials, as they were selected for this property. Table 1 lists a subjective evaluation of the printability of each material (1 is excellent, 5 is unacceptable). While silicone rubber has a number of very good properties from a screen printing point of view, it was given a fair rating because of the difficulty in cleaning it from the screen. Special solvents and mechanical action were needed to remove it from the screen, while the other well materials could be cleaned in acetone and isopropyl alcohol.

The relationship between well dimensions and mask opening for a particular membrane material is important to

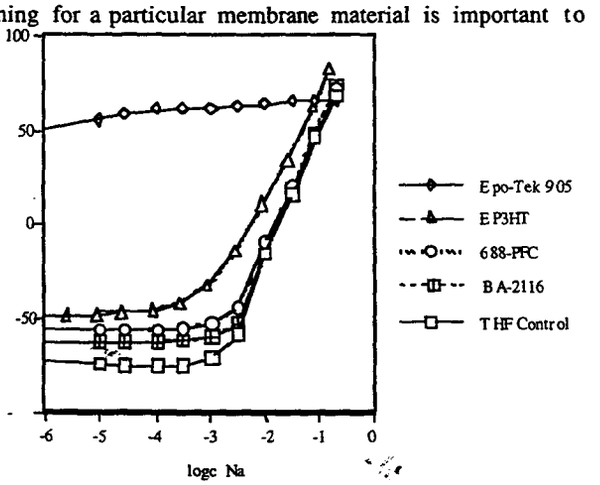


Figure 3. Improved sodium response by increasing the ionophore and borates to 4 wt. %.

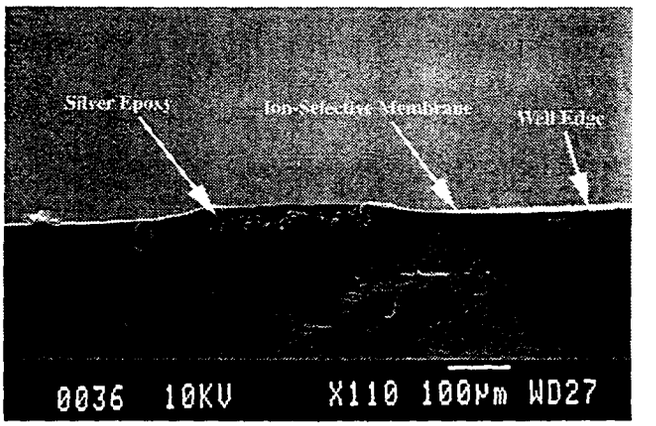


Figure 4. SEM of Ion selective membrane printed into a EP3Ht well.

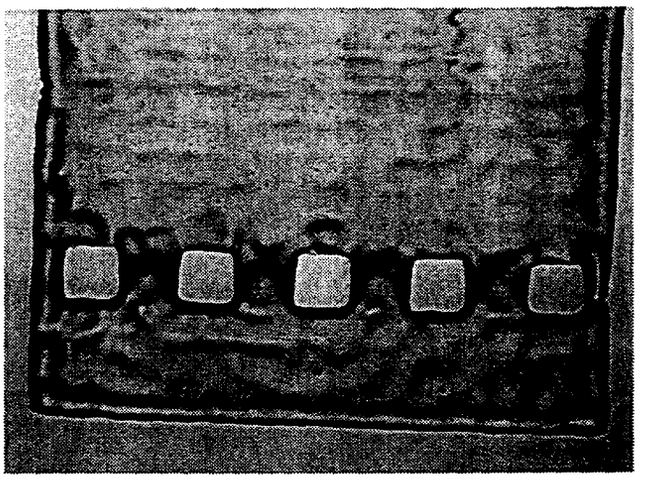


Figure 5. Picture of screen printed 3140 RTV silicone rubber wells.

successful use of the well concept. If the well is too small, membrane solution will overflow it, and may contaminate another membrane. If the well is too large, membrane thickness and dimensions will be more sensitive to viscosity of the membrane solution, and membranes will not be as uniform (as in the case of using no wells). Figure 4 is a SEM micrograph of a properly printed well. Figures 5 and 6 are pictures of screen-printed wells. We found that for our PU/PVA membranes, the well should be 200 - 400 µm larger on a side than the opening in the membrane mask. When silver epoxy contacts partially fill the well, the membrane mask opening must be reduced in size. When printing over the silver epoxy bumps, it was observed that the well size was 100 µm smaller than the drawn width. This needs to be taken into concideration when designing the wells.

SUMMARY

Thick-film wells on solid-state liquid chemical sensor surfaces facilitate the manufacturing process by making it less sensitive to the viscosity and surface-tension properties of the membrane solution. Properly-sized wells control the thickness and lateral dimensions of the

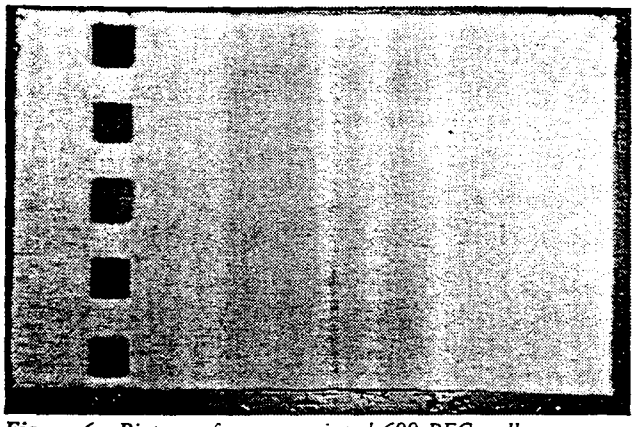


Figure 6. Picture of screen printed 688-PFC wells.

membranes. Wells allow closer membrane spacing, and result in more uniform sensors.

Materials must be evaluated in three general areas to determine their appropriateness for use as membrane wells. First, membrane compatibility must be determined to see if the material would contaminate the membrane. Second, adhesion tests must be done to determine if the material is chemically compatible with the membrane material. Finally, the screen-printing properties (primarily viscosity and thixotropy) must be examined by actually printing wells.

In this study, four epoxies and a silicone rubber were evaluated. The results showed that two of the epoxies have all of the properties needed for use as wells. The first is Epo-Tek 688-PFC, which has excellent screen-printability, and excellent adhesion to both the silicon nitride surface and to the ion-selective membrane. It showed some anionic response due to contaminants which might leach into the membrane, but even in the extreme case that was evaluated (extraction of contaminants with THF), this anionic sensitivity was overcome by increasing the amount of ionophore in the membrane. The second epoxy is the Master Bond EP3HT, which also has excellent screen printability and has excellent adhesion to the silicon nitride and the membrane. This material showed no anionic response in the contaminated membrane, and had very little change in the detection limit.

REFERENCES

1. R. W. Hower, J. H. Shin, G. S. Cha, R. K. Meruva, M. E. Meyerhoff, and R. B. Brown, "New Solvent System for the Improved Electrochemical Performance of Screen-Printed Polyurethane Membrane-based Solid-State Sensors.", *Proceedings of Transducers 95 / Eurosensors IX*, Stockholm, Sweden, June 1995, pp. 858-862.

2. R. B. Brown, *An Integrated Multiple-Sensor Chemical Transducer*, Ph. D. Dissertation, University of Utah, 1985.

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

4. S. N. Cozzette, G. Davis, J. A. Itak, I. R. Lauks, R. M. Mier, S. Piznik, N. Smit, S. J. Steiner, P. Van Der Werf, and H. J. Wieck, "Wholly Microfabricated Biosensors and Process for the Manufacture and Use Thereof.", U. S. Patent, 5,200,051, May 18, 1993.

5. L. J. Bousse, J. W. Parce, J. C. Owicki, K. M. Kercso, "Silicon micromachaning in the fabrication of biosensors using living cells.", *Technical Digest. IEEE Solid-State Sensor and Actuator Workshop* 173-6; Hilton Head Isl., SC, 6/90, Transducer Research Foundation, Cleveland (1990).

6. R. Eugster, P.M. Gehrig, W. E. Morf, U. E. Spichiger, W. Simon. "Selectivity-Modifying Influence of Anionic Sites in Neutral-Carrier-Based Membrane Electrodes.", *Analytical Chemistry*, vol. 63, pp. 2285-2289 (1991).

7. U. Schaller, E. Bakker, U. E. Spichiger, E. Pretsch, "Ionic Additives for Ion-Selective Electrodes Based on Electrically Charged Carriers.", *Analytical Chemistry*, vol. 66, pp. 391-398 (1994).

8. E. Malinowska, V. Oklejas, R. W. Hower, R. B. Brown, and M. E. Meyerhoff, "Enhanced Electrochemical Performance of Solid-State Ion Sensors Based on Silicone Rubber Membranes.", *Proceedings of Transducers 95 / Eurosensors IX*, Stockholm, Sweden, June 1995.

USE OF THE PSEUDO-RIGID BODY MODEL TO SIMPLIFY THE DESCRIPTION OF COMPLIANT MICRO-MECHANISMS

L. G. Salmon, D. B. Gunyan
Integrated Microelectronics Laboratory, Department of Electrical and Computer Engineering
Brigham Young University
Provo, Utah 84602

J. M. Derderian, P. G. Opdahl, and L. L. Howell
Mechanical Engineering Department
Brigham Young University
Provo, Utah 84602

ABSTRACT

This paper reports the use of the pseudo-rigid body method to predict the motion of compliant micro-mechanisms (CMEMS). Physical measurement of the motion of CMEMS structures is compared with predictions of the pseudo-rigid body model and finite element analysis. It is shown that for some structures the computationally simple pseudo-rigid body model can accurately predict the motion of these structures with error of less than the measurement accuracy of 0.5μm for deflections as large as 30μm. This paper also discusses the use of the pseudo-rigid body model to provide a hierarchical approach to CMEMS design that sharply reduces the computation time required to predict the motion of complex CMEMS structures.

INTRODUCTION

The use of compliant mechanisms for the construction of MEMS provides numerous advantages over the use of conventional rigid-body structures. Compliant micro-mechanisms can be easier to fabricate, have a simpler topology, and require no assembly. [1] Compliant mechanism have no joints, can be fabricated using a single mask level, and its motion is dependent on materials parameters and dimensions rather than critical tolerances and friction.

The use of compliant mechanisms in MEMS technology is limited, however, because their motion is highly non-linear and thus difficult to simulate. Linear equations are often inaccurate and non-linear finite element analysis is often too computation intensive. The linear deflection equations such as the Bernoulli-Euler equation do not adequately describe the motion of compliant mechanisms because deflection in compliant micro-mechanisms typically exceed the angular deflection appropriate for these approximations. Non-linear finite element analysis can be used to accurately predict the non-linear motion of compliant micro-mechanisms, but the computation time required to describe complex structures is often prohibitive.

THE PSEUDO-RIGID BODY MODEL

Recently a new formalism has been developed to simplify description of compliant mechanisms. [2] The pseudo-rigid-body (PRB) method sharply reduces the complexity of compliant mechanism simulation by modeling the compliant mechanism as a parameterized rigid-body mechanism. The complex non-linear motion of CMEMS structure can then be predicted using simple, computationally efficient kinematic methods.

Compliant mechanisms gain some or all of their motion from the deflection of flexible members. In general, the deflections of flexible members in compliant mechanisms are large, and non-linear analysis methods must be employed to accurately predict their behavior. Early efforts in compliant mechanism analysis involved using elliptic integral solutions to model simple mechanisms. [3,4] More recently numerical methods, such as nonlinear large-deflection finite element methods, have been used to analyze the behavior of compliant mechanisms. [5]

Theory has recently been developed that simplifies the nonlinear analysis of compliant mechanisms. Before the advent of compliant mechanism analysis theory, the difficulty associated with non-linear analysis often limited the use of compliant mechanisms to very simple tasks. For example, to use non-linear finite element analysis, a designer must already have obtained a detailed initial design before the mechanism could be modeled. In contrast, the pseudo-rigid-body model in compliant mechanism theory serves as a fast and efficient method of evaluating many different trial designs to meet the specific design objectives. Once a design is obtained such that it meets the specified design objectives, it may be validated or further refined using methods such as non-linear finite element analysis.

One of the key concepts in compliant mechanism theory is the pseudo-rigid-body model. This model provides a means of simplifying the nonlinear analysis associated with large deflections. It also is valuable in predicting the behavior of complex systems of flexible members. Pseudo-rigid-body models have been developed for a number of types of flexible segments, two of which are briefly described here as prototypes for other more complex compliant structures. If a system has flexural segments with lengths small relative to those of the rigid segments, the flexural segments are modeled as pin joints with torsional springs to represent the member stiffness, as shown in Figure 1a. [6] These segments are called small-length flexural pivots. The spring constant is calculated as:

$$K=\frac{EI}{l}$$

where E is the modulus of elasticity, I the moment of inertia, and l the length of the flexible segment.

Figure 1b illustrates the pseudo-rigid-body model for a long flexible segment with an end-force load. Again, the model consists of two rigid links, connected by a "characteristic pivot" to represent the displacement, and a torsional spring to model the beam resistance to the

0-9640024-1-8 $3.00 © 1996 TRF

applied force [2]. This model predicts a deflection path of the beam that is within 0.5% of the closed-form elliptic integral solutions for quite large deflections. The location of the characteristic pivot is specified by γ, where the length of the pseudo-rigid link is γL. A nondimensionalized "stiffness coefficient", K_Θ, is used to specify the stiffness of the torsional spring as:

$$K=\gamma K_\Theta \frac{EI}{L}$$

where L is the length of the flexible segment (Figure 1b). Values of $\gamma = 0.85$, and $\gamma K_\Theta = 2.16$ are used for the constant-force mechanisms described herein. Figure 1c illustrates a similar model for fixed-guided flexible segments [7].

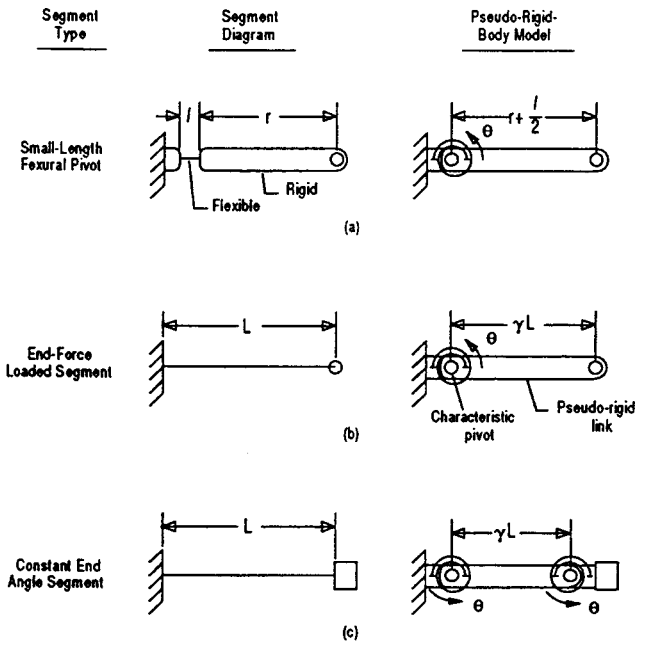


Figure 1. *Pseudo-Rigid-Body models for various flexible segments.*

The pseudo-rigid-body models for individual flexible segments may be used to model more complex systems that include such members. This provides the ability to develop a pseudo-rigid-body model of a compliant mechanism, and then use the large body of knowledge available in the field of rigid-body mechanism analysis and design. In this way, the pseudo-rigid-body model concept acts to unify compliant mechanism and rigid-body mechanism theories.

THE PARALLEL-GUIDING MECHANISM

In order to validate use of the PRB model to describe compliant micro-mechanisms, we fabricated several types of CMEMS structures, measured their deflection characteristics, and compared their motion with predictions made by the PRB model. In this paper we report on the results obtained from a compliant structure that is a combination of two of the connected constant end angle segments illustrated in Figure 1 c). This structure is called a parallel-guiding mechanism and was chosen because it is simple enough to be modeled using non-linear finite element analysis, complex enough to be poorly modeled by use of the Bernoulli-Euler equations, and robust enough to measure deflection into the non-linear regime.

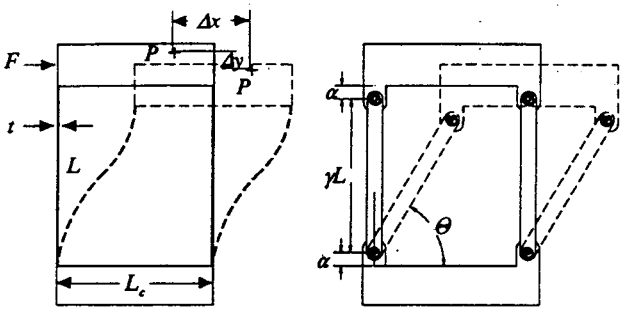


Figure 2. *A compliant parallel-guiding mechanism and its pseudo-rigid-body model.*

Figure 2 is a sketch of a prototype compliant parallel-guiding mechanism together with the equivalent PRB model representation. As can be seen from the figure, the motion of the compliant parallel-guiding mechanism can be described using the PRB model and standard kinematic position analysis. The path equations for a point "P" in the compliant parallel-guiding mechanism are:

$$\Delta x=\gamma L\cos(\Theta)$$

$$\Delta Y=\gamma L(\sin\Theta-1)$$

$$\alpha=\frac{L(1-\gamma)}{2}$$

where the characteristic radius factor, γ, is determined from the orientation of the input force.

Using the stiffness coefficient, K_Θ, and the PRB equivalent model for the compliant parallel-guiding mechanism we find:

$$F=\frac{8K_\Theta EI(\frac{\pi}{2}-\Theta)}{L^2\sin\Theta}$$

where the applied force, F, is the horizontal forcing acting on the coupler. Displacement values calculated from these equations were used to compare with displacement measurements from fabricated compliant parallel-guided micro-mechanisms and the results are discussed in the next section of the paper.

EXPERIMENTAL VERIFICATION OF THE PSEUDO-RIGID BODY MODEL

Two processes were used to fabricate the structures used in this study. The first was a single level poly-silicon process with 4µm of polysilicon and a 2µm SiO_2 release layer. The second process was the MUMPS process provided by MCNC with 2µm of poly-silicon and a 2µm SiO_2 release layer. An SEM photograph of a fabricated structure is shown in Figure 3.

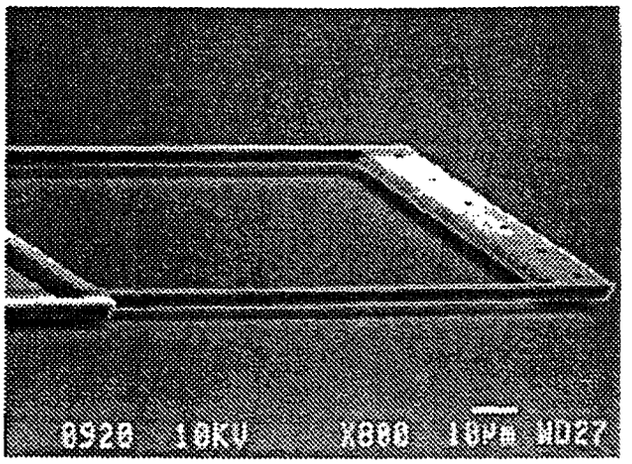


Figure 3. *SEM micrograph of fabricated parallel-guiding mechanism.*

Table 1 indicates the dimensions of the two mechanisms used in this work.

Table 1. *List of the physical parameters for the two parallel-guided mechanisms measured in this study.*

Structure	L	t	L_c	
A	100 µm	3 µm	103 µm	
B	100 µm	3 µm	53 µm	

Figures 4 and 5 compare the measured deflection of parallel mechanism A and B with predicted displacement extracted from the PRB model and non-linear finite element analysis using ANSYS. Data for several instances of each structure are represented in the graph. Data for

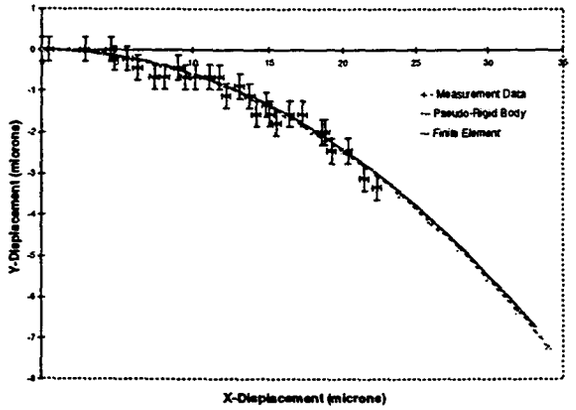


Figure 4. *Graph of measured and predicted x- and y-displacement for parallel-guided mechanism B.*

greater deflection was not measured because of mechanism failure. The error bars represent the uncertainty due to the pixel size of the digitized image used to determine position of the mechanism.

Both graphs show excellent correlation between measured data and predictions from the PRB model and the non-linear finite element analysis. The data from mechanism A with a longer coupler length is somewhat closer to the predicted values than mechanism B, but the data from both structures closely match predicted values.

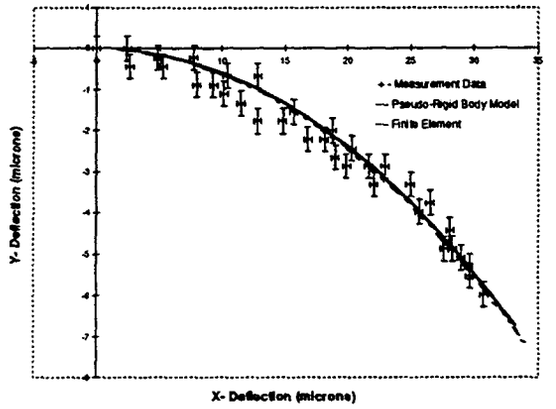


Figure 5. *Graph of measured and predicted x- and y-displacement for parallel-guided mechanism A.*

The Bernoulli-Euler equations would predict zero vertical deflection, and the predicted path is coincident with the x-axis. This is an accurate approximation for small, linear deflections, but the error increases as the deflection increases.

The close correlation between predicted and measured data indicates that little or no torsion is occurring in the measured structures. Torsion is a major problem with analysis of the path of CMEMS because all of the analysis assumes that the displacement of the structures is restricted to the plane of the wafer and measurement of out-of-plane displacement is difficult to determine. We measured other types of CMEMS structures and found that torsional movement severely compromised measurement data. We are currently fabricating small linewidth devices using thicker poly-silicon films in order to minimize torsional effects.

CONCLUSIONS

The close correlation between the measured data and the pseudo-rigid body predictions for parallel-guiding CMEMS mechanisms indicates the accuracy of the PRB method for moderately complex CMEMS structures. The comparatively short computation time required to simulate structures using the PRB method indicate great promise for simulation of complex CMEMS structures.

In future work we will verify the accuracy of the PRB approach to model more complex CMEMS mechanisms and to develop a formalism for complex CMEMS designs. These designs will include precise movement and bistable CMEMS mechanisms.

ACKNOWLEDGMENTS

The authors express their appreciation to the staff of the Integrated Circuit Laboratory for their assistance in fabricating the structures reported in this work. We also appreciate the fabrication efforts of MCNC in processing the wafers fabricated using the MUMPS process. This work is supported by the National Science Foundation under grant No. ECS-9528238.

REFERENCES

1. G. K. Ananthasuresh, S. Kota, and Y. Gianchandani, "A Methodical Approach to the Design of Compliant Micromechanisms," *Technical Digest of the 1994 IEEE Solid-State Sensor and Actuator Workshop*, Hilton Head Isl., SC, June 1994. pp. 189-192.

2. L. L. Howell and A. Midha, "Parametric Deflection Approximations for End-Loaded Large-Deflection Beams in Compliant Mechanisms," *ASME Journal of Mechanical Design*, Vol. 116, 1994. pp. 156-165.

3. Burns, R.H., and Crossley, F.R.E., "Kinetostatic Synthesis of Flexible Link Mechanisms," ASME Paper No. 68-Mech-36, 1968.

4. Shoup, T.E., and McLarnan, C.W.,"On the Use of the Undulating Elastica for the Analysis of Flexible Link Devices," *Journal of Engineering for Industry*, Trans. ASME, 1971, pp. 263-267.

5. Her, I., Midha, A., and Salamon, B.A., "A Methodology for Compliant Mechanisms Design: Part II-Shooting Method and Application," *Advances in Design Automation*, DE-Vol. 44-2, 18th ASME Design Automation Conference, 1992, pp. 39-45.

6. Howell, L.L. and Midha, A.,"A Method for the Design of Compliant Mechanisms with Small-Length Flexural Pivots," *ASME Journal of Mechanical Design*, Vol. 116, No. 1, 1994, pp. 280-290.

7. Howell, L.L., Midha, A., and Norton, T.W., "Evaluation of Equivalent Spring Stiffness for Use in a Pseudo-Rigid-Body Model of Large-Deflection Compliant Mechanisms," *ASME Journal of Mechanical Design*, Vol. 118, No. 1, 1996, pp. 126-131.

SURFACE-MICROMACHINED BEAMS WITHOUT SPRING EFFECT OF ANCHOR STEP-UP

Lan V. Ngo, Phyllis Nelson*, and Chang-Jin Kim
Mechanical and Aerospace Engineering Department
*Electrical Engineering Department
University of California, Los Angeles
Los Angeles, CA 90095

ABSTRACT

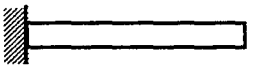
A new anchor design for a surface-micromachined beam is proposed for eliminating the extra spring effect at the step-up anchor. With a minor modification in the sacrificial-layer mask layout, the spring effect is eliminated by forming a reinforcement hump at the beam anchor. The design idea is verified by comparing the resonant frequencies of the conventional and proposed microbeams numerically in FEM analysis and experimentally with a Michaelson laser interferometer.

INTRODUCTION

Surface-micromachined beams are a main component in microelectromechanical systems (MEMS). When surface-micromachined beam is fabricated, it has a step-up (Fig. 1a) at the anchor, which contributes an extra spring effect. For many MEMS applications, the anchor spring effect is not accounted for in the design and analysis, leaving questions in the accuracy of their performance. Although this effect can be accounted for with additional beam theory equations, they often lead to a complex analysis. For simplicity, ideal fixed-end boundary condition (Fig. 1c) is commonly used for modeling the boundary condition of the surface-micromachined beams. However, results showed that care must be taken when surface-micromachined beams are modeled as a cantilever beam [1].

a. Typical surface-micromachined beam with step-up.

b. Limiting case.



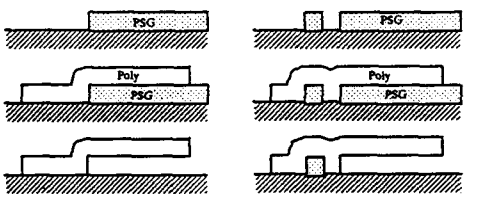
c. Ideal fixed end.

Fig. 1: Various Anchor Types

Work have been done to fabricate a surface-micromachined beam similar to a cantilever beam of the limiting case and to alleviate this problem, so that analytical results are more accurate. However, the design usually requires modifying the process and/or introducing new process steps [2]. In this work, a surface-micromachined beam is designed to eliminate the step-up spring effect. Without altering the fabrication process or adding new processing steps, it makes only a simple modification in the layout of the sacrificial layer. The idea of the mask modification is to form a reinforcement at the step-up anchor by taking advantage of the gap-filling characteristic of LPCVD, so that the spring effect is eliminated and the beam can be modeled more accurately as a simple cantilever beam. Since the anchor condition of surface-micromachined beam has a limit, the goal of this project is to design a micro-beam that has an anchor condition approaching the limit (Fig. 1b). The design idea is initially verified by FEM analysis, and microbeams are made and tested.

DESIGN AND FABRICATION

Fig. 2 illustrates two surface-micromachined beams with different anchors: typical surface-micromachined beam with conventional anchor (Fig. 2a) and the microbeam with the new purposed anchor (Fig 2b). As show in the figure, both microbeams have the same fabrication process; the only difference is the additional PSG block made by minor modification on the sacrificial layer mask pattern. When conventional surface-micromachined beam is fabricated, all the sacrificial layer (usually PSG) is removed in the anchoring region, Fig. 2a. However, when the beam with new reinforced anchor is made, a block of PSG is left in the anchoring region to create a narrow gap, Fig. 2b. If the gap is narrow enough, the subsequent polysilicon layer will fill the gap, forming a reinforced anchor and approaching the limit condition of Fig. 1b.



(a) *(b)*

Fig. 2: (a) Microbeam with conventional anchor. (b) Microbeam with proposed anchor

Fig. 3 is a demonstration of the polysilicon filling the narrow oxide gap. It is fabricated by MCNC/MUMPs processing foundry [3]. The polysilicon and the oxide layers are 2 μm thick, and the gaps are 3 μm and 2 μm in the mask level. The fabricated size of the gaps are about 0.5 μm wider because of the overetching during the process. The picture shows that the new design technique may make beams approaching the limit case (Fig. 1b).

The testing beams are also fabricated by using the standard MUMPs foundry. Since the fundamental frequency of the beam is measured by using laser interferometer, the beam is made relatively wide (50 μm). However, the wide beam is subject to squeeze film damping effect which may attenuate the vibration amplitudes of the beam and decrease the sensing signals. In an

0-9640024-1-8 $3.00 © 1996 TRF

attempt to minimize this effect, a post process is performed to remove the substrate underneath the beam.

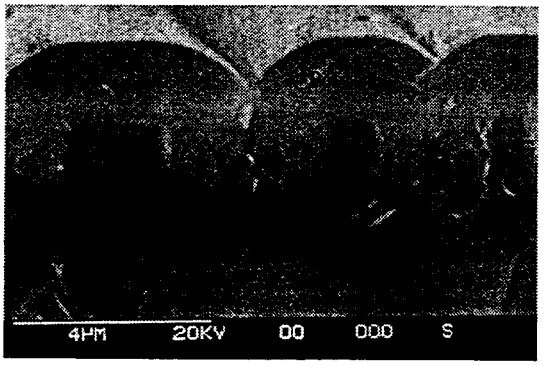


Fig. 3: Polysilicon filling the oxide gaps. The gaps are 3 μm (left) and 2 μm (right) width. The poly and the PSG are 2 μm thick.

This post process is a non-mask etching step by taking advantage of the overetching during the MUMPs fabrication. The overetch will define a poly protected mask by opening pit on the substrate while the oxide still covers the poly. Fig. 4 is the process flow of the post etch. After removing the protective photoresist in the MUMPs dice, xenon diflouride (XeF_2) is used to remove the silicon substrate underneath the beams. The nitride layer underneath the beam is etched by phosphoric acid (H_3PO_4). Finally, the structure is released with HF.

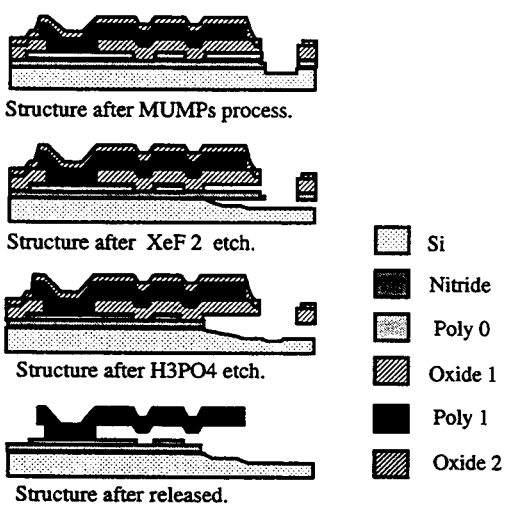


Fig. 4: Process flow for the post etch. Poly2 and metal layers of MUMPs are not shown in the figure.

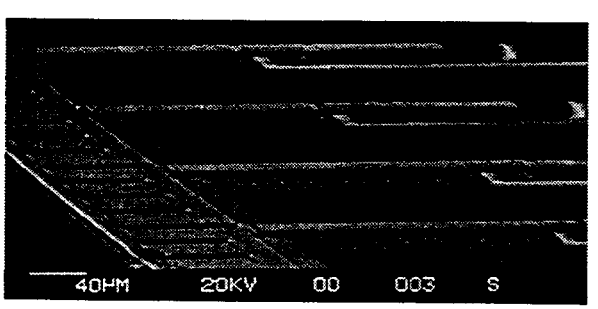


Fig. 5: Microbeams with conventional anchor and proposed anchors. Substrate underneath the beam ends are recessed.

Fig. 5 is the complete beams after the post processing. Beams with conventional anchor and identical beams with proposed anchor are placed next to each other as pairs for direct comparison. Fig. 6 shows the proposed anchor at the base of a microbeam.

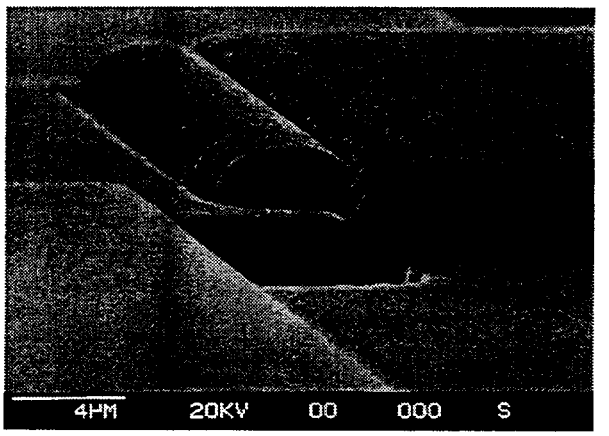
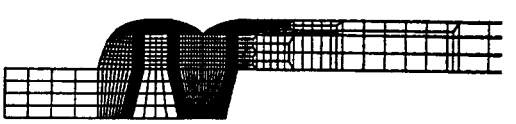


Fig. 6: Proposed anchor at the base of a microbeam.

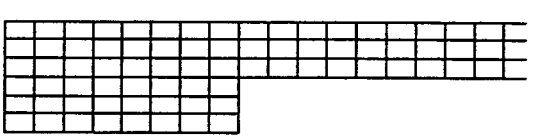
FINITE ELEMENT MODELING

FEM models are developed and coded in a commercial PATRAN package for predicting the resonant frequency. The beam is modeled as a two-dimensional problem. Beams with three different types of anchoring are analyzed and compared. Fig 7 shows the models of the beams with different types of anchors: conventional step-up anchor, proposed reinforced anchor, and limiting case. For the step-up anchor and the proposed reinforced anchor, the curve profile of the anchor is fed directly from the actual structure fabricated, i.e., from the SEM pictures of the cross section of the microstructures.

(a) Conventional anchor.



(b) Proposed anchor.



(c) Limiting case.

Fig. 7: FEM models of beams with various anchors. Curved profiles in (a) and (b) are based on the actual cross-section of the fabricated structures.

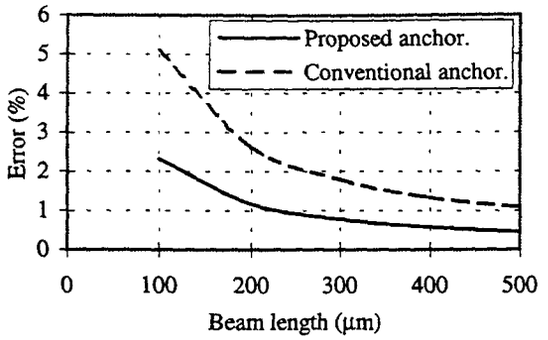


Fig. 8: Deviation of the resonant frequency of microbeams with two different anchors from the case of fixed end beams.

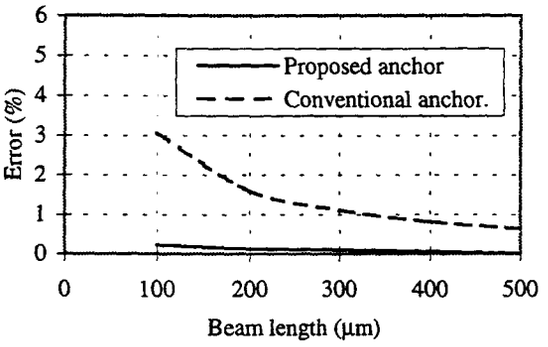


Fig. 9: Deviation of the resonant frequency of microbeams with two different anchors from that of the limiting case.

By analyzing the resonant frequency of beams with various anchors, the spring effect of different anchors are evaluated. Fig. 8 compares the percentage deviation of the resonant frequency of microbeam with conventional and proposed anchor from the case of an ideal fixed end support (Fig. 1c). The deviation is the error in predicting resonant frequency of the beam if an ideal fixed end is assumed for analysis. The curve for the proposed anchor almost coincides with that for the limiting case (Fig. 1b) as evidenced by Fig. 9, which compares the errors of the conventional anchor and proposed anchor when they are approximated as the limiting case. The proposed anchor has negligible error when approximated as the limiting case support. Same comparison has been made for microbridges in Fig. 10 and Fig. 11, as opposed to the cantilevers.

MEASUREMENT

To evaluate the anchor design, the motion of microbeams with the conventional anchor and the proposed is monitored by a Michaelson interferometer. Fig. 12 is a schematic drawing of the apparatus. Mirrors M1 and M2 are used to align a helium neon (HeNe) laser used as the light source for the interferometer. This input beam is divided into two orthogonal beams by a beam splitter (BS). One beam (B1), which is used as the reference, is reflected back to the beam splitter by a mirror (M3) mounted on a piezoelectric translator. The other beam leaving the beam splitter(B2) is focused on the microstructure by a lens. The reflection from the device under test returns through the lens to the beam splitter, where it is combined with the reference beam. Upon their return to the beam splitter, the transmitted portion of B1 and the reflected portion of B2 pass through a

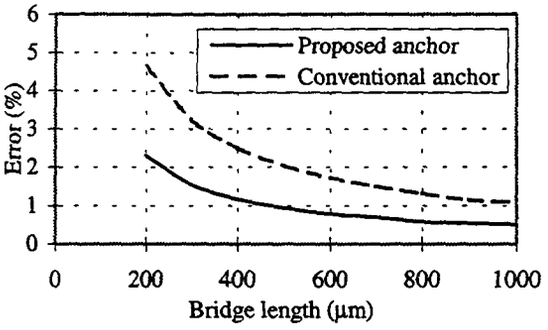


Fig. 10: Deviation of the resonant frequency of microbridges with two different anchors from the case of fixed end bridges.

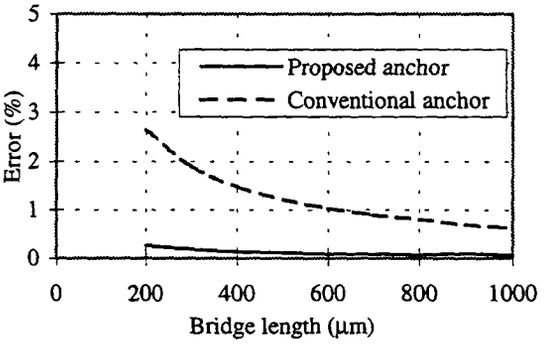


Fig. 11: Deviation of the resonant frequency of microbridges with two different anchors from the case of the limiting case.

diverging lens. The system is focused to give a projected image of the device under test at the plane of an aperture (A) with beam B2 alone. When the return reflection of B1, the reference, is added, an interference pattern is superimposed on the image due to the difference in the phase of the two beams.

The intensity of the light passing the aperture is detected by a silicon photodiode having a time constant of less than 1 μs. By adjusting the linear translator, the relative optical path lengths traveled by B1 and B2 can be adjusted for a linear small signal response to displacement of the device under test. A lock-in amplifier is used to detect the device motion by isolating the intensity variations at the appropriate frequency. Measurement of the fringe contrast allows small signal intensity changes to be converted to absolute displacement of the test structure. Therefore, the dynamics response of the system can be evaluated.

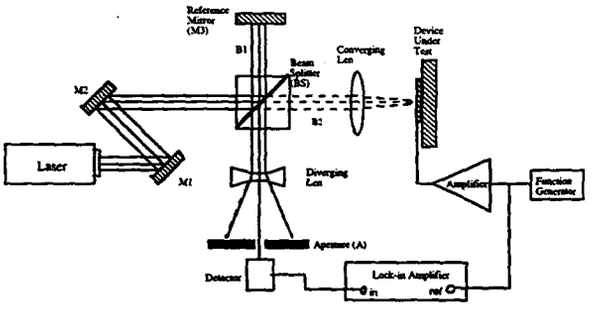


Fig. 12: The Michaelson interferometer used to measure the resonant frequency.

The microstructures are excited by an electrostatic actuator, which is built in the test structure. At a constant DC offset, sinusoidal signals are applied to the system and swept at a frequency range to determine the resonance.

We estimated the resonant frequency of 500 μm long, 50 μm wide, 2 μm thick microbeams with both types of anchors from the measured phase shifts of their dynamic responses. Fitting the phase data for the conventional microbeam to that expected for a second order damped spring mass system yields a resonant frequency of 8.92 kHz. The data points and fit are shown in Fig. 13. (This fit also includes a phase offset of -28 degrees to account for the phase response of the actuation and measurement electronics.) The measured phase data for the structure with the proposed anchor, which is also plotted in Fig. 13, cannot be fit to the simple second order form. However, this structure yields a smaller phase shift at any given frequency than the conventional one, which is consistent with the prediction that the proposed structure is closer to the fixed end condition (i.e., stiffer).

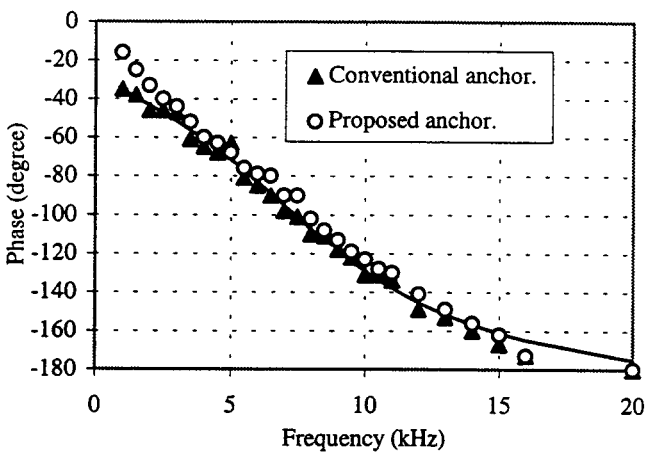


Fig. 13: Phase response of microbeams with conventional anchor and proposed anchor. (The fit of the conventional anchor to a second order damped spring mass system is shown as a solid line.)

Our measurements are limited to less than 20 kHz by the bandwidth and slew rate of the amplifier used to actuate the microbeams. Thus, we have only examined relatively long microbeams with low frequency resonance. However, these are structures in which the effect of the anchor is relatively small. We are currently improving the actuation system, and will then study shorter beams with a larger dependence on the anchor structure. In addition, we will investigate the possibility of using an acousto-optic modulator and optical heterodyne detection to improve the signal-to-noise ratio of the measurement system.

CONCLUSION

A new anchor that reduces the step up spring effect in surface-micromachined beams has been designed and fabricated. This new design requires no change in the existing fabrication process. The modified anchor performance has been verified by FEM analysis and measured experimentally. The results show that the anchor spring can be eliminated by minor modification in the sacrificial layer layout.

ACKNOWLEGEMENT

The authors would like to thank Prof. Oscar Stafsudd for valuable discussion and the loan of experimental equipment and Prof. Yang for the use of CAD lab.

REFERENCES

1. Q. Meng, M. Mehregany, and R. L. Mullen, "Theoretical Modeling of Microfabricated Beams Elastically Restrained Supports," *J. Microelectromechanical Systems*, Vol. 2, (1993), pp. 128-137.

2. Y.X. Li, P.J. French, and R.F. Wolffenbuttel, "Plasma Planarization for Sensor Applications," *J. Microelectromechanical Systems*, Vol. 4, (1995), pp. 32-138.

3. Multi-User MEMS Process (MUMPs) of the MCNC Center for Microelectronics Systems Technologies, Research Triangle Park, NC 27709-2889.

4. F. I. Chang, R. Yeh, G. Lin, P. B. Chu, E. Hoffman, E. J. J. Kruglick, K. S. J. Pister, and M. H. Hecht, "Gas-Phase Silicon Micromachining With Xenon Difluoride," *SPIE 1995 Symposium on Micromachining and Microfabrication*, Austin, Texas, (1995).

5. P. B. Chu, P. R. Nelson, M. L. Tachiki, and K. S. J. Pister, "Dynamics of Polysilicon Parallel-Plate Electrostatic Actuators," *Proc. 8th Int. Conf. on Solid-State Sensor and Actuator*, Stockholm, Sweden, (1995), pp. 356-359.

Microscale Pumping with Traversing Bubbles in Microchannels

Thomas K. Jun and Chang-Jin Kim
Mechanical and Aerospace Engineering Department
University of California, Los Angeles
Los Angeles, CA 90095

ABSTRACT

Pumping in microchannels using asymmetric heating and a single or multiple traversing bubble(s) is studied and demonstrated. This pumping mechanism requires no mechanical moving parts for actuation. A heat and mass transfer analysis was performed to understand the pumping mechanism and revealed that the two dominant forces are vapor pressure and surface tension. A pumping device with a channel hydraulic diameter of 3.4 μm was fabricated on a silicon wafer using surface micromachining and tested. Testing results showed pumping velocities as high as 160 μm/s (505 pl/min) with a pressure gain of approximately 800 Pa.

INTRODUCTION

The effect of phase change in a fluid was studied to determine if this process could be used to achieve net motion of a fluid in a microchannel. A heat and mass transfer analysis revealed that the two dominant forces would be variations in vapor pressure and surface tension due to heater induced temperature gradient.

In this study, we introduce two methods to achieve net fluid movement in microchannels. The first method (Fig.1a) consists of heating a single bubble and growing it in one direction, similar to the method used by Ozakis [1] in much larger metal channels (180μm vs. our 4μm hydraulic diameter). The second method (Fig.1b) utilizes multiple bubbles that are generated in sequence from left to right. The existing bubble(s) acts as an anchor during the subsequent nucleation, causing asymmetric liquid flow. The end result for both methods is net fluid displacement in the microchannel.

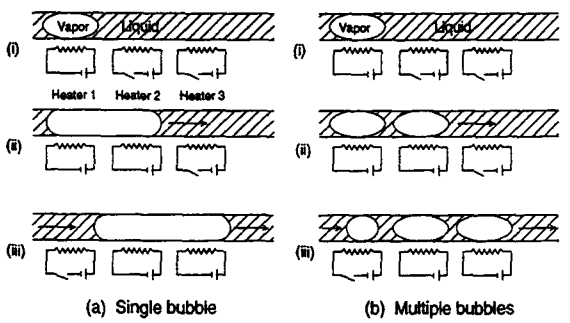


Figure 1. Two different modes of obtaining net fluid displacement

To verify the pumping principle, a test device was surface microachined as shown in Fig. 2. A series of heaters (six in the picture) are placed across the channel top shell. To allow for visual observation, we chose oxide for the shell layer and polysilicon as the sacrificial layer, unlike typical surface micromachining processes.

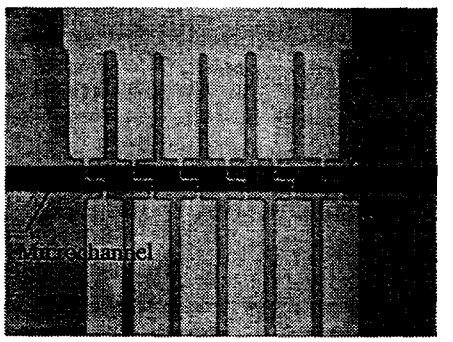


Figure 2. Optical photo of micropump and channel (channel has a hydraulic diameter of ~4 μm)

THEORY

A vapor bubble under asymmetric heating was modeled to understand the nature of the various forces that would be present (Fig. 1(a) case). The relevant forces are gradients in vapor pressure, surface tension (Marangoni forces) and momentum due to evaporation and condensation. The other method of obtaining net pressure in a duct (Fig. 1(b) case) is not analyzed in this section but is tested experimentally. By growing and collapsing a stationary vapor bubble, it could serve as an obstruction against flow in a channel and therefore function as a valve.

The vapor pressure of a liquid can get very large as the temperature of the liquid approaches the critical temperature. Heating a vapor bubble from one end should create a temperature gradient across the bubble. This temperature gradient would imply the presence of a net pressure difference across the end of the vapor bubble since vapor pressure is a strong function of temperature. A simple conduction model would predict a significant temperature difference between the heated and unheated end. However, due to the process of evaporation and condensation, a conduction only model is inadequate for this configuration. Therefore, a conjugate heat transfer model incorporating both heat and mass transfer was used.

The simplified geometry of the bubble and duct are shown in Fig. 3.

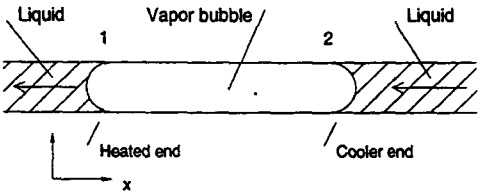


Figure 3. Geometry modeled by heat and mass transfer analysis

0-9640024-1-8 $3.00 © 1996 TRF

The energy equation was applied across the vapor bubble in the duct

$$\rho c\frac{\partial T}{\partial t}+u_b\frac{\partial(\rho cT)}{\partial x}=\frac{\partial}{\partial x}\left(k\frac{\partial T}{\partial x}\right)+\frac{UP}{A_c}(T-T_\infty) \quad (1)$$

where
u_b = average bulk velocity in x direction
ρ = density of fluid
c = specific heat
k = thermal conductivity of fluid
T = temperature of fluid
T_∞ = ambient temperature
U = overall heat transfer coefficient
P = perimeter of duct
A_c = cross sectional area of duct

The justification for eliminating the y and z terms of the energy equation is based on the assumption that the small cross sectional area of the microchannel is at approximately the same temperature. Similar to 1-D fin analysis, heat lost at the surface is accounted for by using a surface boundary condition, which is the last term in Eq. (1). Next the momentum equation is applied to flow in a duct as

$$u\frac{\partial(\rho u)}{\partial x}=-\frac{\partial p}{\partial x}+\mu\frac{\partial}{\partial y}\left[\frac{\partial u}{\partial y}\right] \quad (2)$$

where
u = velocity in x direction
p = pressure in fluid
μ = fluid viscosity

The well known Darcy friction factor solution is

$$f=\frac{(-dp/dx)D_H}{(1/2)\rho u_b^2}=\frac{K}{\mathrm{Re}_D};\frac{dp}{dx}=\frac{\Delta p}{L} \quad (3)$$

where
Re = Reynolds number
K = friction constant (ranges from 64 to 90)
D_H = hydraulic diameter
L = length of vapor bubble

The constant K ranges from 64 for a macro scale circular duct to about 80 for microscale ducts as reported by Pfahler *et al* [2]. From Eq.(2) the pressure gradient in the duct must be linear at steady state. From kinetic theory of gas the evaporation at the hot interface is approximated as

$$\dot{m}''=(p_{sat}-p_{bubble})\sqrt{\frac{1}{2\pi RT_{sat}}} \quad (4)$$

where
p_{sat} = saturation pressure at interface
p_{bubble} = pressure in bubble
R = gas constant
T_{sat} = saturation temperature

Next the interfacial conditions must be satisfied at the vapor-liquid interface. Using an energy and mass balance at the vapor-liquid interface

$$\dot{m}_1''h_{fg}=-k\left[\frac{dT}{dx}\right]_{l_1}+k\left[\frac{dT}{dx}\right]_{v_1} \quad (5)$$

at interface 1 (hot end), and

$$-\dot{m}_2''h_{fg}=-k\left[\frac{dT}{dx}\right]_{v_2}+k\left[\frac{dT}{dx}\right]_{l_2} \quad (6)$$

at interface 2 (cold end). Furthermore

$$\dot{m}_1''=\dot{m}_2''=(\rho_v v_v)_1=(\rho_l v_l)_1 =(\rho_v v_v)_2=(\rho_l v_l)_2 \quad (7)$$

$\dot{m}_1''$ = mass flux at interface 1 (hot end)
$\dot{m}_2''$ = mass flux at interface 2 (cold end)

The temperature boundary and initial conditions are defined as

$$T(x\to\infty)=T_\infty \quad ; \quad T(t=0)=T_o \quad (8)$$

$$T(x=0)=T_{heater} \text{ or } -k\frac{dT}{dx}=q_{heater} \quad (9)$$

$$p_1(x=0)=p_{sat}(T_1) \; ; \; p_2(x=L)=p_{sat}(T_2) \quad (10)$$

Solving Eqs. (1) and (4) simultaneously along with the interfacial and boundary conditions, the temperature, evaporation rate and pressure were found for various heater configuration and vapor bubble lengths. It was found that the pressure difference increased with bubble length as expected.

As stated earlier, there are other factors that may cause a net pressure difference in the duct. Surface tension is a fairly strong function of temperature, hence in the presence of a substantial temperature gradient, the two ends of the vapor bubble can have an interfacial pressure difference known as the Marangoni force [3]. The Marangoni force was approximated as

$$F=\frac{2}{R}\Delta\sigma A \quad (11)$$

R = radius of curvature of vapor bubble
$\Delta\sigma$ = difference of surface tension coefficient across vapor bubble
A = cross sectional area of vapor bubble

The final force that was calculated was the vapor recoil force that was produced due to the evaporation and condensation process and can be high if the evaporation rate is very large. Applying the Reynolds transport theorem for momentum at the liquid and vapor interface, the net force due to evaporation and condensation is

$$F_x=(\rho V^2A)_{vapor}-(\rho V^2A)_{liquid} \quad (12)$$

The effect of condensation is similar to the evaporation process, only now the fluid experiences a deceleration, and so can be modeled in the same way. Using Eq.(12) and typical evaporation rates, the pressure change due to evaporation or condensation was found to be very small (< 1 Pa). As a result, for this analysis, pressure changes due to evaporation or condensation was assumed to be negligible.

FABRICATION

To verify the principles analyzed in the previous section a surface micromachined pumping device was fabricated as shown in Fig. 4. The fabrication process started with the deposition of 2.0 μm of low temperature oxide (LTO) and 2.0 μm of polysilicon using LPCVD. Then the sacrificial polysilicon layer is doped with spin-on phosphorus dopant to make the layer n^{++}. This n^{++} property will later enhance the dry etching rate in XeF_2. This polysilicon layer is then patterned using RIE to form the sacrificial layer that will later become the void space in the microchannel. Next 2.0 μm of LTO is deposited to form the transparent shell of the microchannel (Fig. 4a). Then 0.7 μm of polysilicon is deposited and patterned using RIE to form the heater that lies on top of the microchannel shell (Fig. 4b). Using buffered oxide etchant (BOE), holes are patterned into the oxide layer to expose the sacrificial polysilicon at the ends of the microchannel (Fig. 4c). With photoresist covering the entire wafer except for the oxide holes, the dry etchant XeF_2 is used to etch the sacrificial polysilicon core of the microchannel. The finished devices consist of a hollow oxide shell with polysilicon heaters running along the top surface of the oxide shell as illustrated in Fig. 4d.

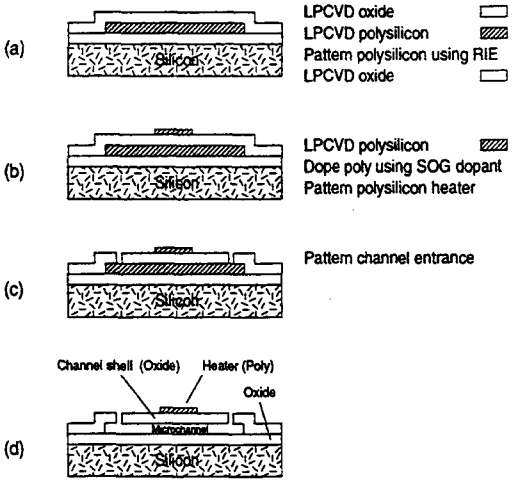


Figure 4. Process flow

TESTING

The micropump was tested to verify the physical principles as described in the theory section. The important pump parameters that were investigated consisted of pumping pressure head, flow velocity and Reynolds number.

The testing was done by powering the heaters in sequence using a pulse signal pattern similar to that shown in Fig. 5. Sample test cases of isopropanol pumping in a 726 μm-long, 29 μm-wide, 2 μm-high microchannel are presented. The three parameters studied were heater pulse amplitude (voltage), pulse width and pulse overlap as shown in Fig. 5. The heater pulse amplitude is simply the heater input voltage and so is directly related to the output power of the line heater. The heater pulse width is the time period that the heater is powered and the pulse percent overlap is defined as the overlap of the power input for adjacent heaters. In this study, these parameters were varied to understand their effect on net fluid flow in a microchannel.

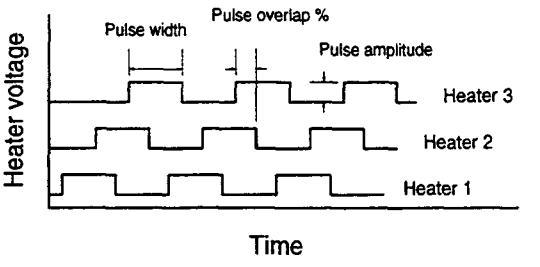


Figure 5. Heater input signal and important parameters

As stated previously, two methods were used to achieve net fluid movement as illustrated in Fig. 1. The first method consisted of heating a single bubble and causing translation of the interface due to the forces mentioned. The second method utilizes multiple bubbles that are generated in sequence from left to right. The existing bubble(s) acts as an anchor during the subsequent nucleation, causing asymmetric liquid flow. The end result of both methods is net fluid displacement in the microchannel.

In this study, three heaters were used to produce net fluid flow in a 3.4 μm hydraulic diameter microchannel. The heater pulse amplitude was varied from 18 to 26 Volts and the pulse width ranged from 0.25 to 1 second. In addition, the pulse overlap ranged from 50 to 90 percent overlap. The three heater signals were produced using a function generator that was controlled using a personal computer. In turn the function generator produced a series of synchronized square pulses with the desired pulse amplitude, width and overlap for each heater as shown in Fig. 5. The testing of the micropump required no special priming signal to initially start the pumping and allowed the same steady state signals to be used from start to finish.

RESULTS AND DISCUSSION

Numerical results of relevant forces

The heat and mass transfer analysis revealed that the two significant forces were variation in vapor pressure and surface tension caused by the heater induced temperature gradients. It was found that at high temperatures, when evaporation rates were high, the vapor pressure forces dominated due to the significant vapor pressure present. However, when the evaporation rates dropped and the temperature between the hot and the cold end got larger, the surface tension became more significant [3]. It was also found that the forces due to vapor recoil were very small for even the largest possible evaporation rates. Table. 1 shows the relative magnitude of the calculated forces.

	Net Pressure
Vapor Pressure	600 Pa
Marangoni Pressure	10 Pa
Vapor Recoil	< 0.5 Pa

Table 1. *Comparison of calculated forces present in microchannel with 130 μm long vapor bubble (three heater lengths) at ~92 °C with ΔT=0.1°C.*

Experimental Results of micropumping device

Fig. 5 illustrates the basic operational envelope of the micropump operating at a heater voltage of 24V. This voltage

translates to 58 mW to each heater, which was found to be the minimum power required to obtain reliable operation of the micropump. The failure mode at power inputs lower than 55 mW (23 Volts) was usually due to failed bubble nucleation at one or more heater sites.

As Fig. 6 shows, the lower and upper limits of the pulse percent overlap was found to be approximately 60 to 85 percent pulse overlap. At pulse overlap greater than about 85% the failure mode was found to be little net displacement of fluid due to the three bubbles nucleating at nearly the same time. The end result was equal displacement of fluid to both sides of the microchannel. This is due to the finite time required for bubble nucleation and growth. Without enough time lag between adjacent heaters, the sequential bubble generation from left to right was lost. To investigate this phenomenon, the input power to the heater was increased in an attempt to shorten the time for bubble generation and growth, thereby getting rapid but sequential bubble generation at the three heaters. It was found, however, that in this case one large symmetrical bubble would nucleate and prevent the pump from operating. At the lower pulse overlap limit, the failure mode was usually due to the premature collapse of the previously generated bubble(s). In the case where the pump was operating in the multiple bubble mode (Fig. 1b) this caused the pump to lose the anchor effect of the earlier bubbles and so operation was less than ideal. For the single bubble operation, premature collapse meant that the single bubble started collapsing on the starting side and thus lost the increased net pressure effect of having a longer bubble.

Testing also revealed that heater power input and pulse width were the parameters that determined whether the pump would operate in single or multiple bubble mode. Heater power, however, was found to be the more significant deciding parameter of the two. When the heater was powered at a high voltage of 26V (62 mW/heater), the pump operated almost exclusively in the single asymmetrically growing bubble mode when testing within the operational window described earlier. However when the pump was operated at a lower voltage of 23V (55 mW/heater) the pump always operated in the multiple bubble mode.

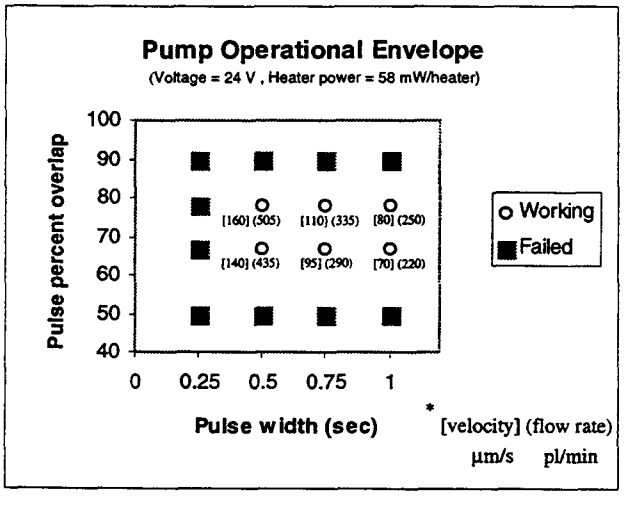


Figure 6. Operation envelope of the micropump

The pulse width also influenced the operational mode of the pump. Testing at 24V and 67 % overlap, the pump operated in the multiple bubble mode for pulse width of 0.5 seconds but changed to the single bubble mode when the pulse width was increased to 1 s. Surprisingly pulse overlap had less of an effect on operational mode determination. Intuitively one would guess that larger pulse overlap would result in single bubble mode of operation but this was not found to be the case at least within the operational window tested.

During testing there was no observable liquid microlayer between the vapor bubbles and the channel walls, which suggests there would be little liquid slipping past the bubble, if any. It was indeed visually confirmed that the motion of the actuated vapor bubble at the center was duplicated by the vapor bubbles away from the heated zone, leading us to conclude that the translating vapor bubble in fact causes net motion of the column of fluid. The above was confirmed when we succeeded in pushing vapor bubbles out of one end of the microchannel. This direct displacement also allowed us to obtain the flow rate from the bubble motion. Assuming that the pump was operating in quasi-steady state, the flow rate and the solution for steady laminar flow in a duct was used to estimate the pressure supplied by the pump.

CONCLUSIONS

The feasibility of pumping fluids in a 3.4 µm hydraulic diameter microchannel using asymmetric heating and traversing bubbles has been demonstrated. A heat and mass transfer analysis has been done to estimate the magnitude of the relevant forces present. The analysis revealed that variations in vapor pressure and surface tension (Marangoni effect) are the two dominant forces present.

The advantage of this present design is the complete absence of any moving mechanical parts or small component assembly. Furthermore the proposed design can be integrated in standard silicon processing..

Test results has shown that our early generation design is already capable of delivering velocities as high as 160 µm/s in a 3.4 µm hydraulic diameter, 726 µm long microchannel. This corresponds to pressure heads as large as 800 Pa, and a flow rate of 0.5 nl/min. With optimization of both pump design and operation significant improvements is expected in the performance of future versions.

ACKNOWLEDGMENT

The authors wish to thank J. Simon for his numerous assistance in processing and Prof. C.-M. Ho and Dr. K. Schadow for various discussions. This work is in part supported by the Office of Naval Research under the Grant N00014-94-1-0536

REFERENCES

1. Koichi Ozaki, "Pumping Mechanism Using Periodic Change of a Fluid", Proc. *IEEE Micro Electro Mechanical Systems*, Amsterdam, Netherlands, Jan.-Feb. 1995, pp.31-36.
2. J. Pfahler, J. Harley and H. Bau, "Gas and Liquid Flow in Small Channels",ASME Proc HTD-Vol.32, *Micromechanical Sensors, Actuators, and Systems*, 1991, pp.49-60.
3. J. L. McGrew, T. L. Rehm and R. G. Griskey, "The Effect of Temperature Induced Surface Tension Gradients on Bubble Mechanics", *Appl. Sci. Res.*, June 1974, pp.195-210.
4. M. Fahndrich, B. Hochwind and A. Zollner, "Fluid Dynamics in Micro Dosing Actuators", The 8th International Conference of Solid-State Sensors and Actuators, *(Transducers '95)*, Stockholm, Sweden, June 1995, pp.295-298.
5. A. F. Mills, "Heat and Mass Transfer", Richard D. Irwin, Inc., Homewood, IL, 1995.
6. V. S. Arpaci, "Conduction Heat Transfer", Addison-Wesley, Reading, MA, 1966.

AN ELECTROSTATICALLY–ACTUATED MICROVALVE FOR SEMICONDUCTOR GAS FLOW CONTROL

Janet K. Robertson and Kensall D. Wise

Center for Integrated Sensors and Circuits
Department of Electrical Engineering and Computer Science
The University of Michigan
Ann Arbor, MI 48109–2122

ABSTRACT

Semiconductor manufacturing presents important challenges in terms of the sensors needed for closed–loop process control. In particular, improved trimmer valves are needed to precisely control gas flows below 10sccm. This paper reports a small electrostatically–actuated microvalve which has been integrated with micromachined flow channels to form an integrated microflow controller. The current valves actuate at 80V, but actuation voltages as low as 40V have been obtained. A gas flow model of the microvalve in the molecular, viscous, and transition flow regimes has been developed. The model has been verified by data obtained from microvalves. The maximum calculated flow rate for eight identical microvalves and flow channels in parallel is 0.69sccm with a minimum predicted leak rate of 3.8×10^{-3}sccm. The response time of the valve is ~100μsec. Assuming a minimum operating pulse width of 10msec, the valve is capable of delivering a 14.5×10^{-6} standard cm^3 pulse of gas at 100Torr applied differential pressure.

INTRODUCTION

Many semiconductor processes such as MOCVD, LPCVD and RIE require precise control of gas flow rates at very low operating pressures. This paper reports an electrostatically–actuated trimmer microvalve capable of high precision gas flow control. Electrostatically–actuated valves offer several advantages over other valve types. They are low power and very small, permitting multi-valve assemblies with small chip sizes and very small dead volumes. The microvalve reported here extends a previously reported device [1] via several structural and process changes. This valve differs significantly from other recently reported valves which have been aimed at higher flow rates and higher operating pressures. [2,3]

Figure 1 shows a perspective drawing of the integrated microflow controller. Gas enters the device through a bulk micromachined gas inlet hole in the silicon cap. A free standing double–end–clamped beam is positioned beneath the gas inlet. The beam acts as a microvalve because it can be electrostatically deflected either upward to seal against the gas inlet, or downward out of the gas flow path. The gas, after entering the device through the inlet, flows parallel to the glass substrate through flow channels micromachined in the silicon cap. The gas exits the device through a 500μm diameter outlet orifice drilled in the glass substrate. A differential pressure sensor can also be fully integrated with the device. Since the pressure sensor was developed previously [4] and is known to function well, it was not integrated with the microvalves reported here, thereby simplifying the fabrication process.

The microflow controller contains eight silicon beams—forming eight microvalves, each with its own flow channel—leading to an outlet hole drilled in the center of the die. The beams are 600μm long, 240μm wide, 3μm thick and are 3μm above the glass substrate. The beams are positioned beneath the inlet orifice in an inlet cavity approximately 7.8μm deep. The flow channels are 400μm wide and 1200μm long. The flow channel height can be varied, but for the device reported here, it was ~12.8μm.

The valve is actively pulled into both the on– and the off–states using the cap, beam, and substrate metallization as three independent electrodes. The valve reported here actuates at 80V, but actuation voltages as low as 40V have been observed. The individual valves have a theoretical operating range of 4.8×10^{-4}sccm (closed) to 8.7×10^{-2}sccm (opened), with a calculated response time of 100μsec.

FABRICATION

A double–end–clamped single-crystal silicon beam is formed using a five–mask dissolved–wafer process with a wafer-level electrostatic bond to glass. Figure 2 is a photograph of a micromachined beam. The new fabrication sequence uses a robust Ti/Pt to Ti/Pt lead transfer between the beam and the glass. The metal to metal contact has a series resistance of less than 10 ohms for a 100μm x 50μm contact.

The cap wafer was formed by first recessing the front side of the wafer, to form the flow channels, and then bulk micromachining the back side of the wafer, to form the gas inlets. The cap wafer was then sawed and affixed to the glass over the previously formed beams. Since the beam and the cap were completely fabricated before final assembly, the flow channels and the microvalve inlet area could be thoroughly cleaned before the beams were encapsulated.

Figure 3 is a SEM of the 120μm diameter gas inlet. The matrix of metal circles visible around the inlet are standoffs which were used to prevent stiction. The octagonal metal o–ring around the edge of the inlet is also visible. Figure 4 is a photograph of two completed die. The eight bulk micromachined gas inlets are clearly visible.

THE GAS FLOW CONDUCTANCE

The gas flow path of the integrated microflow controller can be divided into six parts: the inlet orifice, the beam, the inlet region, the flow channel, the exit region, and the exit port. These regions are shown schematically in Figure 5. The effects of the large inlet orifice and exit port are negligible. In the following sections, the theoretical equations for both the molecular conductance and the viscous conductance for each part of the flow system are reviewed. These equations will then be used to model the conductance of the integrated microflow controller gas flow channels and the microvalve.

0-9640024-1-8 $3.00 © 1996 TRF

Solid-State Sensor and Actuator Workshop
Hilton Head, South Carolina, June 2-6, 1996

The Molecular Conductance

The molecular molar flow rate, Q_m, in sccm, is given by

$$Q_m(SCCM) = k_m \cdot \Delta P, \quad k_m = \frac{\bar{v}A}{4} W \left[\frac{T_s}{P_s T}\right] \tag{1}$$

where k_m is the molecular conductance, W is the transmission coefficient, $\bar{v}$ is the mean molecular speed, and $T = 296$K is the measurement temperature. The last term on the right has been included to convert to standard temperature and pressure, 273K and 1atm, respectively.

The transmission coefficient is a function of the geometry of the flow channel. For rectangular flow channels, W is given by [5]

$$W_R = \sqrt{\pi\delta}\,\frac{R}{L}\left\{0.5 + \ln\left[\frac{2}{\delta}\right]\right\} \tag{2}$$

where L is the length of the flow channel, R is the effective radius defined as $\pi R^2 \equiv ab$, a and b are the side lengths $(a \le b)$ and $\delta = a/b$. Equation 2 is valid for long channels $(L/R > 100)$, and will be used to calculate k_m for the flow channels and the exit region.

The molecular conductance in the inlet region and the conductance between the beam and the cap were modeled as gas flow between two parallel flat plates. The transmission coefficient for two flat plates is given by $W_m = w_1 - w_2$ [6] for

$$w_1 = 0.5\left[1 + \sqrt{1+x^2} - x\right] \tag{3}$$

$$w_2 = \frac{0.66\left\{x - \ln\left[x + \sqrt{x^2+1}\right]\right\}^2}{x^3 + 3x^2 - 4 - \left(x^2+4\right)\sqrt{\left(1+x^2\right)}} \tag{4}$$

where $x = L/h$, and h is the plate separation.

If two flow channels with cross-sectional areas A_1 and A_2, and transmission coefficients W_1 and W_2, are connected in series, the molecular transmission coefficient for the resulting flow channel is given by [7]

$$\frac{1}{W_{total}} = \frac{1}{W_1} + \frac{1}{gW_2} - \frac{1}{g}, \quad g = \frac{A_2}{A_1} \tag{5}$$

where W_{total} is the system transmission coefficient based on an effective cross-sectional area of A_1.

The Viscous Conductance

For compressible viscous flow, the viscous molar flow rate, Q_v, is given by [8]

$$Q_v(SCCM) = k_v \cdot \Delta P, \quad k_v = \frac{AD_h^2}{L\eta\phi} \cdot \frac{2\bar{P}T_s}{P_s T} \tag{6}$$

where k_v is the viscous conductance, D_h is the hydraulic diameter, η is the viscosity, $\Delta P = P_1 - P_2$ is the applied differential pressure, $\bar{P}$ is the average pressure in the flow channel, and ϕ is an empirical constant related to the geometry of the flow channel. When two or more flow channels are connected in series, the viscous conductance of the entire system is given by

$$\frac{1}{k_v(total)} = \frac{1}{k_{v1}} + \frac{1}{k_{v2}} + \frac{1}{k_{v3}} \cdots \tag{7}$$

The Transition Conductance

In the transition region, the conductance can be approximated by [9]

$$k_t = k_v + k_m\left[\frac{1 + 2.507K_n^{-1}}{1 + 3.095K_n^{-1}}\right] \tag{8}$$

where K_n is the Knudsen number.

THE MICROVALVE MODEL

Before the microvalve could be accurately modeled, the conductances of the inlet region, the flow channel, and the exit region had to be determined. To understand the contribution of each of these regions, two gas flow tests were run.

The first test measured the combined conductance of the flow channel and the exit. The inlet of two valves on a cap die was removed with a diamond saw. This modified cap die was then electrostatically bonded to a glass die. Although the glass die was metallized and an exit port had been drilled in the die, no beams were present. The gas flow rate as a function of applied pressure was then measured. The conductance was then modeled using Equations 1,2, 6, and 8.

The second test measured the combined flow conductance of the inlet, the flow channel and the exit. A complete cap die was electrostatically bonded to a metallized glass die. An exit port had been drilled in the glass die, but as with the previous test, the double-end-clamped beams were omitted. The gas flow rate as a function of applied pressure was then measured. The conductance of the inlet region was modeled as flow through two parallel plates. The total conductance was then modeled by combining the previously developed model for the flow channel and the exit region with the inlet region conductance using Equations 5 and 7. The total transition region conductance was then calculated using Equation 8. Figure 6 is a graph of both the measured and the predicted flow conductance for each of the previous tests.

The gas flow rate through a functioning microvalve and its adjoining flow channel was then measured. The effect of the beam deflection on the system conductance was modeled as flow through two parallel plates. The total system conductance was calculated by combining the model developed in the previous tests with the conductance of the deflecting beam. The microvalve model was partially verified by using the microvalve gas flow data and the microvalve model to estimate the plate separation between the encapsulated beam and the inlet as the valve was actuated. The estimated plate separations are shown in Figure 7 for both an actuated microvalve and an undeflected microvalve.

The average estimated plate separation in the undeflected state was 1.95μm, which was consistent with the valve measured dimensions for an undeflected beam. The measured spacing between the cap and the glass was ~7.85μm. The beam was ~3μm from the glass and ~3μm thick. Therefore the beam to inlet spacing for an undeflected beam should have been ~1.85μm, which agrees with the estimated spacing of ~1.95μm.

DISCUSSION

When actuated at 80V to close the valve, the modeled flow indicated that the beams often closed only to within 0.7–1.0μm. Particles from the gas flow test system are suspected to have played a role in this, even though a 0.5μm filter was used to filter the incoming N_2 gas line. This is a well known problem with electrostatic actuation which could be overcome using an alternative thermo–pneumatic valve [10]. Such valves could be retrofit into the same basic flow controller structure and would offer more robust actuation, though at the cost of significantly slower speed, higher power, and greater area.

Figure 8 is a graph of the modeled gas flow rate through a single microvalve and its adjoining flow channel. Figure 8 also graphs the modeled total pressure drop across the valve as a function of the beam to inlet spacing. The pressure drop reported in Figure 8 is the pressure drop across the beam and does not include the head loss due to the inlet region. This graph can be used to estimate the degree of flow modulation possible for various beam to inlet separations. If the beam, when fully opened, is approximately 5μm from the inlet, and closes to within 0.1μm, the flow rate can be modulated by a factor of 182:1 at 100 Torr. For these conditions, the total calculated flow rate for a single microvalve and its adjoining flow channel is 8.7×10^{-2}sccm in the open state and 4.8×10^{-4}sccm in the closed state.

The most attractive feature of these microvalves is their ability to deliver precise gas flows at sub–sccm gas flow rates. For example, with a minimum operating pulse width of 10msec and a maximum system–limited modulation period of 1 second, this microvalve offers better than 7 bits of precision.

ACKNOWLEDGMENTS

This work was funded by the Semiconductor Research Corporation under Contract 95–YC–085.

REFERENCES

1. Robertson, J.K. and K.D. Wise, "A Nested Electrostatically–Actuated Microvalve for an Integrated Microflow Controller", *Proc. IEEE Micro Electro Mechanical Systems*, January (1994), pp. 7–12.

2. Barth, P.W., C.C. Beatty, L.A. Field, J.W. Baker, and G.B. Gordon, "A Robust Normally–Closed Silicon Microvalve", *Solid–State Sensor and Actuator Workshop*, June (1994), pp. 248–250.

3. Zdeblick, M.J., "Thermopneumatically–Actuated Microvalves and Integrated Electro–Fluidic Circuits", *Solid–State Sensor and Actuator Workshop*, June (1994), pp. 251–255.

4. Cho, S.T., K. Najafi, C.L. Lowman, and K.D. Wise, "An Ultrasensitive Silicon Pressure-Based Microflowmeter", *IEDM Technical Digest*, 499-502 (1989).

5. Steckelmacher, W., "The Effect of Cross–Sectional Shape on the Molecular Flow in Long Tubes", *Vacuum, 28*, 269–275 (1978).

6. Berman, A.S., "Free Molecular Transmission Probabilities", *Journal of Applied Physics, 36*, 3356 (1965).

7. Oatley, C.W., "The Flow of Gas through Composite Systems at Very Low Pressures", *British Journal of Applied Physics, 8*, 15-19 (1956).

8. Olson, R.M., *Essentials of Engineering Fluid Mechanics*, 4th edition, Harper & Row, New York (1980).

9. Knudsen, M., *Annln. Phys, 28*, 75 (1909).

10. Bergstrom, P.L., J. Ji, Y. –N. Liu, M. Kaviany, and K.D. Wise, "Thermally Driven Phase–Change Microactuation", *Journal of Micromechanical Systems*, March 1995, Vol 4, pp. 10–17.

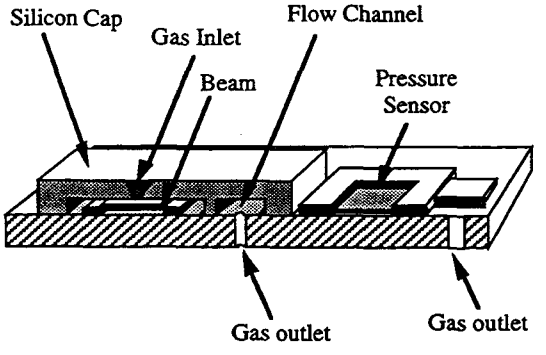


Figure 1. Perspective drawing of the integrated microflow controller. The pressure sensor is formed simultaneously with the beam and is positioned over a separate outlet hole.

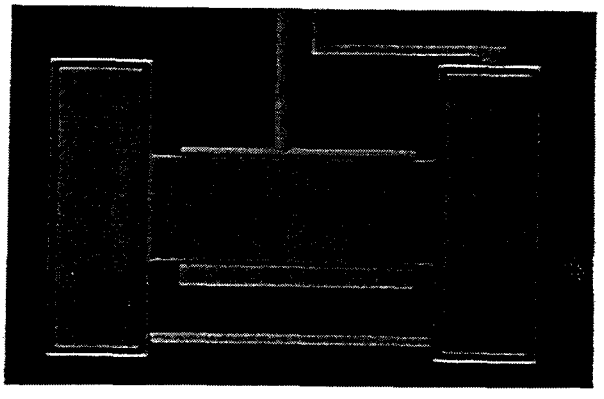


Figure 2. Photograph of a micromachined silicon beam. The beam is 600μm long, 240μm wide and ~3μm thick. The metal pull–down plate is visible beneath the beam.

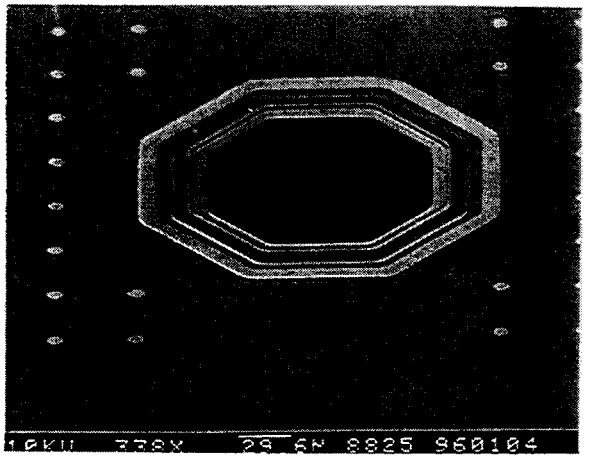


Figure 3. SEM of the microvalve top plate gas inlet.

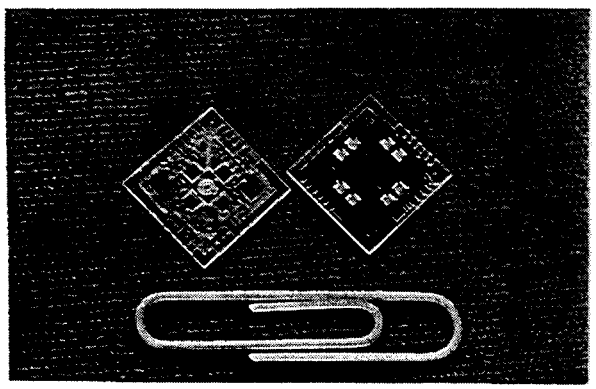


Figure 4. Photograph of two completed microflow controller die. The left die is resting on the silicon cap, and the right die is resting on the glass substrate. The bulk micromachined gas inlets are visible on the right die.

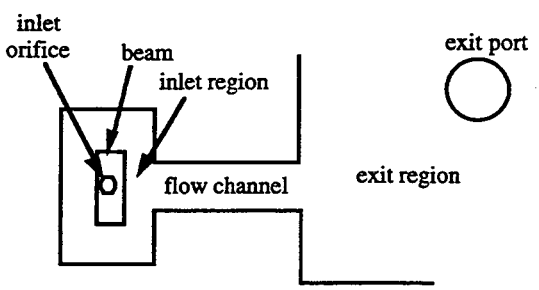


Figure 5. Schematic drawing of a single microvalve and flow channel.

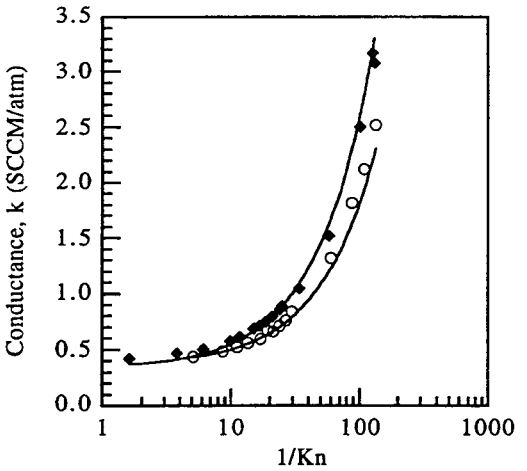


Figure 6. Graph of the measured and modeled conductance for two separate flow tests. ◆ Conductance of the flow channel and the exit region. O Conductance of the flow channel, the exit region and the inlet.

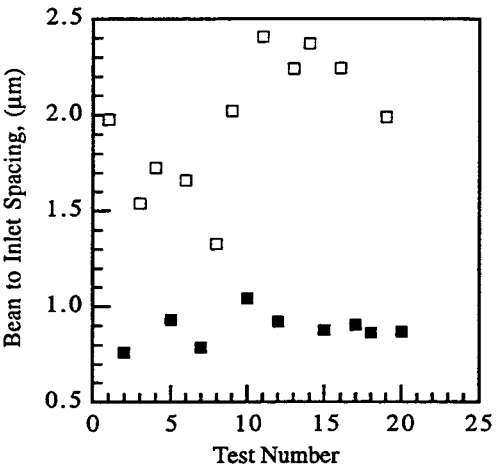


Figure 7. Estimated plate separations inferred from gas flow data taken on undeflected microvalves □, and microvalves actuated between 80V and 112V ■.

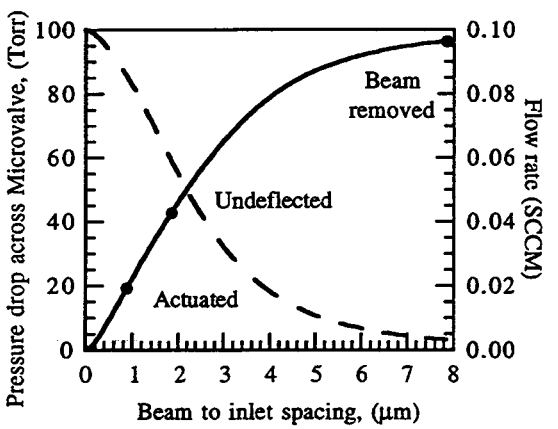


Figure 8. Graph of the modeled microvalve gas flow rate (solid line) and the modeled % pressure drop (dashed line) across the microvalve as a function of beam to inlet spacing. Total applied pressure was 100 Torr. Measured data points are also shown.

LATERAL IN-PLANE DISPLACEMENT MICROACTUATORS WITH COMBINED THERMAL AND ELECTROSTATIC DRIVE

Xi-Qing Sun, Xiaoyi Gu, and W. N. Carr
Microelectronics Research Center, New Jersey Institute of Technology
Room 200-MIC, 323 King Blvd., Newark, NJ07102, USA
sun@hertz.njit.edu, carr@admin.njit.edu

ABSTRACT

Electrostatic actuation using a bimorph structure has been combined with thermal actuation in cantilever devices. The combined electrostatic and electrothermal excitation increases the efficiency of the actuation mechanism and permit a 12 volt voltage supply and 40 mW power. In this work we have optimized the design to create a horizontal movement of $\Delta x = 60$ μm for cantilever of overall initial length L = 1000 μm. The actuation characteristics of cantilevers with only electrostatic actuation, only thermal actuation, and combined actuation are separately presented for comparison. These actuators are designed for direct interface to CMOS driver circuits.

INTRODUCTION

Thermal and electrostatic actuators have mostly been described for out-of-plane displacements. These devices have included the micro tweezers [1,2], micro polymeric-actuators [3], and micromanipulators [4]. The work of Keller and Howe describes one of the few thermally-actuated devices with in-plane actuation [5]. Our present work is with reversed bimorph structures, i.e., the top layer with a smaller thermal expansion coefficient and the bottom layer with a larger value. Following the sacrificial layer release, the cantilever bows upwards due to the built-in stress. Our device exhibits an actuation similar to an "inchworm" to achieve a lateral in-plane displacement of its tip along the plane of the underlying insulation/substrate surface.

When electrostatic actuation is used alone, the device voltage supply is generally 50 volts and higher [1,3,6]. When thermal heating alone is used as the actuating mechanism, the power required is in the range of hundreds of milliwatts [3,4]. These voltage and power requirements limit the interface compatibility for VLSI device circuits especially when low power is a requirement. By combining the thermal and electrostatic drive we have reduced the voltage requirement to 12 volts and the power requirement to 40 mW. These levels are compatible with CMOS single chip drivers.

DESIGN AND FABRICATION

The inchworm actuator is schematically shown in Fig.1. As a cantilever structure with a fixed end, initially the released cantilever bows upwards over hundreds microns of out-of-plane displacement at its center and several tens of in-plane displacement at its tip as illustrated in Fig.2 (a). As an excitation either thermal V_a or electrostatic V_b is imposed on the actuator, the up-bowed cantilever stretches out along the insulation surface in Fig.2 (b), and moves up to the fully closed state in Fig.2 (c) as the excitation exceeds a threshold value. Throughout the whole process, the tip part always keeps contact with the substrate, and thus the inchworm actuator can achieve lateral in-plane displacement. It is because of this point, a friction between the moving tip and the insulation surface is importantly involved in this approach.

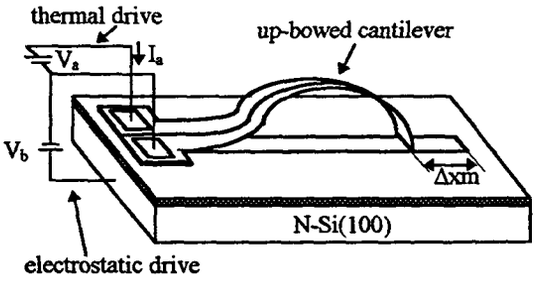


Fig.1. A schematic view of up-bowed actuator.

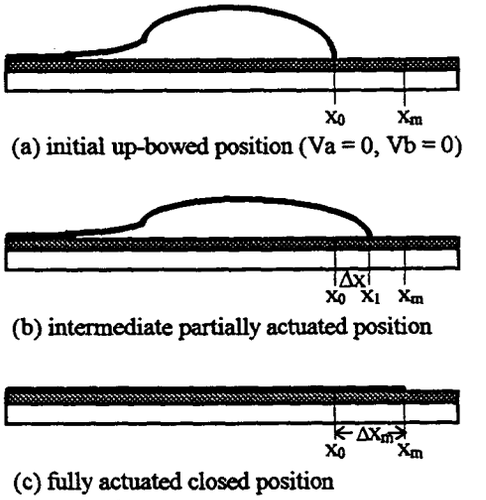


Fig.2. Actuation sequence illustration.

Fig.3 shows the schematic views of the inchworm actuator, which is a polysilicon layer on a $TaSi_2/Si_3N_4$ composite layer to form the bimorph structure. The thermal coefficients of polysilicon, $TaSi_2$ and Si_3N_4 are 2.33 x 10^{-6}/K, 8.8 x 10^{-6}/K and 2.8 x 10^{-6} /K respectively. The high-temperature-balanced bimorph will spring a large curvature at low temperature as soon

as it is released [7]. The polysilicon and $TaSi_2$ layers serve as the heater for thermal actuation and upper electrode for electrostatic actuation. The Si_3N_4 layer offers protection for the upper $TaSi_2$ layer from HF solution. A Si_3N_4/SiO_2 composite layer is used as an insulation layer between the actuator and the substrate, and also provides the stationery support film for the actuator tip motion.

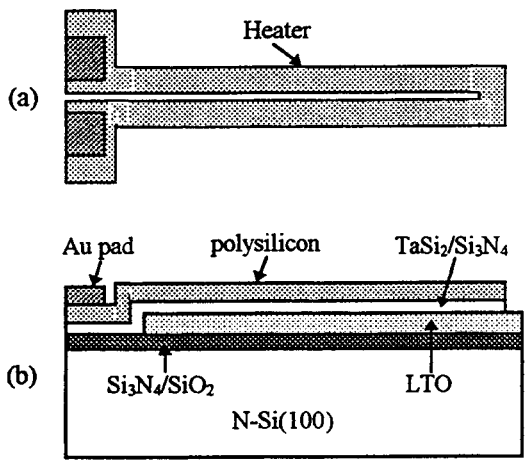


Fig.3. Schematic view: (a) top; (b) side.

The fabrication process started by oxidizing 100 nm thick SiO_2 on a n-type (100) silicon wafer. This is followed by an LPCVD 300 nm Si_3N_4 layer as an insulation and HF-resistant layer. Next, a 2 μm thick LTO was deposited and patterned as

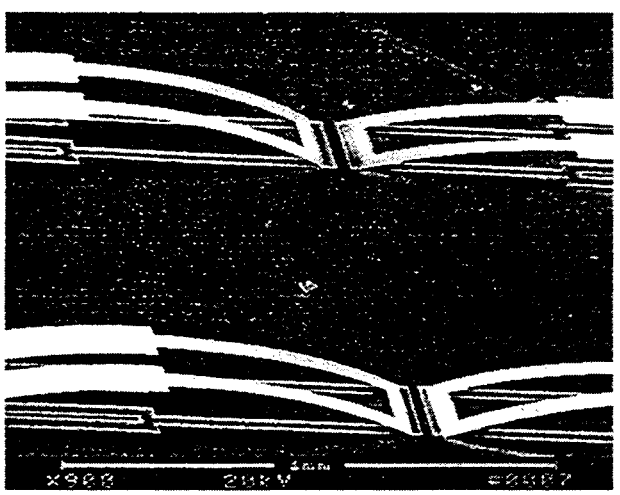


Fig.4. SEM photograph of an up-bowed microactuator.

sacrificial layer. After the 200 nm Si_3N_4 deposition, we followed with a 100 nm $TaSi_2$ sputtering, a 800 nm polysilicon deposition, a subsequent phosphorus diffusion at 950 °C, photolithography and RIE in that order. The $TaSi_2$ material was selected not only for raising the thermal mismatch, but also for lowering the heater resistance to reduce the parasitic electrostatic influence during thermal actuation. The metal contact pads are a 500 nm gold layer. The final release step was done by dipping the wafers into 49% HF solution for over 15 min. The released cantilever bows up and the movable tip

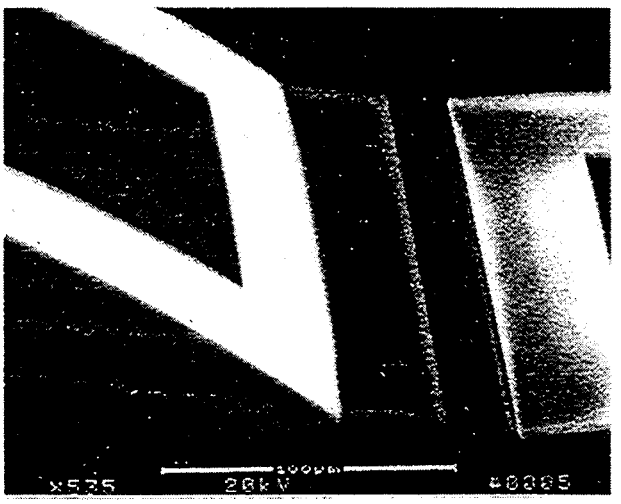


Fig.5. SEM photograph of the tip with an initial lateral in-plane displacement of 35 μm.

retracts backwards on the insulation surface as is shown in the SEM views of Fig.4 and Fig.5. Cantilevers with 1000 μm length and 35 μm width can shrink back 30 to 60 μm lateral in-plane displacement following release. Correspondingly, the vertical out-of-plane displacement of the center part is hundreds microns. So, the up-bowed actuator can generate both the lateral in-plane and the vertical out-of-plane displacements at the same time. In our current design, the inchworm actuator fabrication possesses high yield more than 80% without special release processing.

TEST AND RESULTS

Electrostatic Actuation

A CMOS high-voltage supply [8] was adopted to supply a voltage ranging from 0V to 60V for electrostatic drive. As the applied voltage increases, the up-bowed cantilever is attracted downward under the electrostatic attraction induced from the substrate. Consequently, the tip is pushed laterally along the substrate surface as it overcomes friction with the underlying insulation layer. A cantilever with 1000 μm in length and 30 μm initial lateral displacement was chosen in this test. The measured relationship between the lateral tip position and the applied voltage is shown in Fig. 6. Very similar to the parallel electrode attraction [9], the tip lateral displacement increases steadily as an increase in applied voltage initially. Once the bias exceeds a threshold V_t at 53 V, the cantilever suddenly closes and fully flattens to the insulation surface to achieve the full

lateral displacement. As the bias sweeps down, a hysteresis is observed as shown in Fig. 6 which is contributed mainly by the nature of the parallel electrodes attraction [9], and partly by the friction between the tip and the insulation surface.

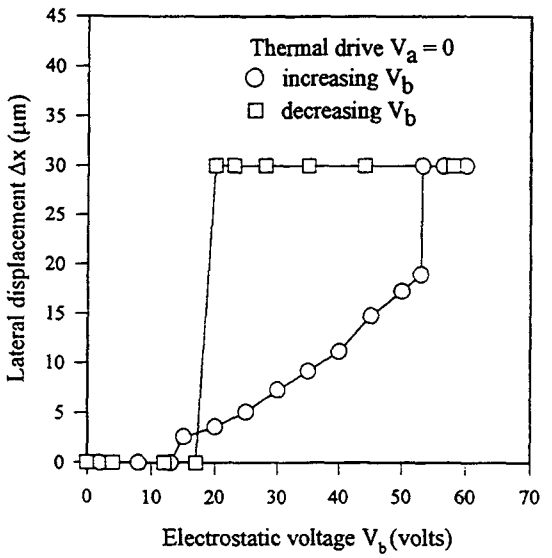


Fig.6. Measured lateral displacement as a function of applied voltage for electrostatic actuation only.

In some aspects, the electrostatically actuated inchworm is similar in principle to the microscratch reported previously[10]. But the up-bowed cantilever structure reported in this design can achieve much larger lateral in-plane displacement.

Thermal Actuation

Quite different from the electrostatic drive, thermal actuation is based on the thermal expansion mismatch of the heated bimorph. The thermal bimorph also provides large force and displacement. A current source I_a was used for thermal actuation. Fig. 7 shows the relationship of tip displacement versus input power. As can been seen, the up-bowed cantilever reaches a displacement of 26 μm at the input power of 86 mW and above at approximate 650 °C.

In Fig. 7 there is a slight hysteresis between the power increasing and the power decreasing processes, probably due to the tip friction against the substrate.

Combined Drive

As described above, separate thermal and electrostatic drives require either high electrostatic voltage up to 53 V or high thermal power up to 86 mW to achieve the maximum motion. In this example, a constant electrostatic bias of 12 V was introduced to assist the thermal actuation. Fig.8 shows the measured results of the combined drive. As can been seen, the input thermal power and the applied electrostatic voltage at which the actuator achieves the full range motion are 40 mW and 12 V respectively.

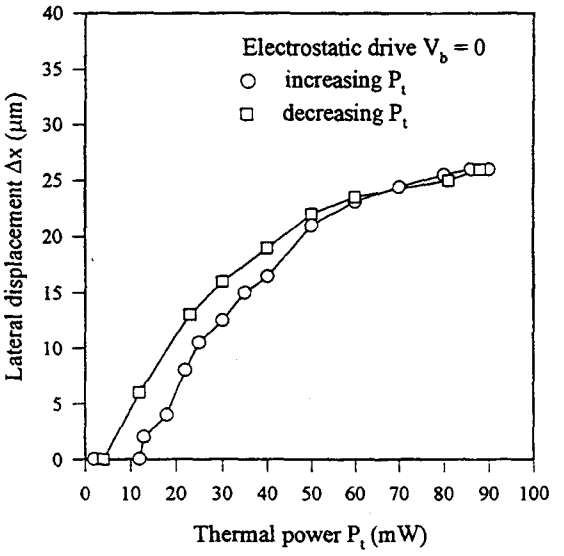


Fig.7. Measured lateral displacement as a function of input power for thermal actuation without electrostatic excitation.

Dynamic behavior of the actuator has been tested with a square-wave signal V_a for the combined actuation. The results show that the actuation response depends largely on the tip motion amplitude. An up-bowed cantilever with an in-plane displacement amplitude of 15 μm can respond at 360 Hz. The cantilever with the full range in-plane amplitude of 30 μm responds at approximate 10 Hz. Our limited amount of fatigue testing shows that an up-bowed actuator with a vibration

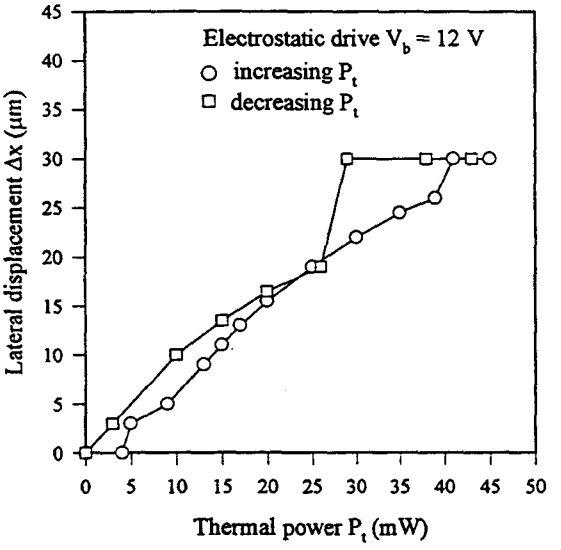


Fig.8. Measured lateral displacement as a function of input thermal power with combined actuation (V_a varies, V_b = 12 V).

amplitude about 15 μm will operate as long as several days continuously without obvious performance degradation;

Actuators with the full range motion have operated only for tens of hours before failure.

Note that all the up-bowed actuators work well under the three drive conditions. The maximum lateral displacement achieved is up to 60 μm, determined by the initial shrunk lateral displacement.

The adhesion between the movable tip part and the insulation layer is strongly affected by room ambient condition. Generally, after a several weeks storage in damp room air, the tip part is stuck to the substrate. In these cases tip performance can be recovered by repeating the release process again.

CONCLUSIONS

An up-bowed microactuator which achieved large lateral in-plane displacement up to 60 μm, was successfully designed and fabricated with standard CMOS fabrication process. Electrostatic, thermal, and the combined actuations were evaluated. The combined drive reduced the power supply requirements to a much lower level. The combined actuation was powered directly with CMOS circuitry. Hysteresis was observed among all three actuations and the tip friction was considered to be the important hysteresis factor. The dynamic response and the lifetime for continuous operation of the devices depended largely on the vibration amplitude. With half range vibration amplitude the devices were operated at 360 Hz for several days without failure.

ACKNOWLEDGMENT

We would like to give thanks to Dr. Robert Marcus and Yanwei Zhang of Microelectronics Research Center, NJIT for helpful suggestions.

REFERENCES

1. W.-H. Chu and M. Mehregany, "Microfabricated tweezers with a large gripping force and a large range of motion", *Technical Digest of the 1994 Solid-State Sensor and Actuator Workshop*, Hilton Head , SC, June 13-16 (1996), pp. 107-110.

2. L. Chen, Z. Zhang, J. Bao, D. C. Thomas, and N. C. Macdonald, " Selective chemical vapor deposition of tungsten for microdynamic structures", *Proc. IEEE MEMS-89*, Salt Lake City, Utah, Fub. 20-22 (1989), pp. 83-87.

3. G. Lin, C.-J. Kim, et al, "Design, fabrication, and testing of a C-shape actuator", *Digest of Transducers'95*, Stockholm, Sweden, June 25-29(1995), pp. 416-419.

4. H. J. Quenzer, A. Maciossek, et al, "Surface micromachined metallic microactuator with buckling characteristics", *Digest of Transducers'95*, Stockholm, Sweden, June 25-29 (1995), pp. 128-131.

5. C. G. Keller and R. T. Howe, "Nickel-filled hexsil thermally actuated tweezers", *Digest of Transducers'95*, Stockholm, Sweden, June 25-29 (1995), pp. 376-379.

6. X.-Q. Sun, Z.-J. Li, and L.-T. Liu, "Some structure improvements in silicon based micro electrostatic motors", *Digest of Transducers'93*, Yokohama, Japan, June 7-10 (1993), pp. 760-762.

7. G. Zheng and W. N. Carr, "Internal stress of thin bimorph microengineered structures", *Proc. ICMPC'94*, Kunming, China, Nov. 7-11 (1994), pp.498-503.

8. H. Jiang and W. N. Carr, "On-chip integration of high-voltage generator circuits for an electrostatic micromotor", *Digest of Transducers'95*, Stockholm, Sweden, June 25-29 (1995), pp. 150-153,.

9. X.-Q. Sun and Z.-J. Li, "A micro polysilicon beam switch vibrator", *Chinese Journal of Semiconductors*, Vol. 14, No. 6, pp. 320-327, 1993.

10. T. Akiyama and K. Shono, "Controlled stepwise motion in polysilicon microstructures", *Journal of Micro Electro Mechanical Systems*, Vol. 2, No. 3, pp. 106-110, 1993.

IMPEDANCE IMAGING FOR HYBRID BIOSENSOR APPLICATIONS

D. A. Borkholder[1], N. I. Maluf[1,2], and G. T. A. Kovacs[1]

[1]Stanford University
Center for Integrated Systems Room 202X
Stanford, CA 94305-4075

[2]Lucas NovaSensor
1055 Mission Court
Fremont, CA 94539

ABSTRACT

The impedance characteristics of individual cell / electrode systems are used to monitor cellular viability, position, adhesion, and response to external stimuli in hybrid biosensor applications. A planar microelectrode array consisting of 36 platinized iridium electrodes (10 μm diameter) is used as a substrate for the culture of mammalian cells. Electrode impedance is monitored across the array as different environmental factors are changed. Maps of electrode impedance have been shown to correlate directly to cell positioning over an electrode and general cellular viability. Exposure to a well known voltage-gated Na^+ channel blocker (tetrodotoxin) provided significant cellular response as compared to control electrodes without cells. The effective use of small electrodes (10 μm diameter) to study single cell / electrode interactions has been demonstrated.

INTRODUCTION

Much of biosensor development aims to utilize the highly selective and sensitive nature of biological compounds to detect a particular analyte. However, there is interest in developing sensors with sensitivity to a wide variety of physical and chemical stimuli. Whole cell based biosensors offer such diversity while maintaining high sensitivity. Utilizing planar microelectrode technology, it is possible to monitor the electrical activity of anchorage dependent cultured cells [1], as well as the impedance characteristics of the electrode/cell system [2-7]. It has been suggested that it is possible to detect the attachment, spreading, movement, and division of cells plated on a relatively large (250 μm diameter) single electrode with the impedance characteristics dependent on the average response of populations of cells [5].

The reaction of anchorage dependent mammalian cells to different environmental factors may be monitored by impedance measurements of the electrode / cell interface. The resulting impedance is dependent on many variables including degree of occlusion of the electrode surface, relative adhesion of the cell to the electrode, and cell membrane ion channel conductance. Thus, it is possible to monitor cell positioning over an electrode, cellular viability and mobility, and the introduction / removal of agents which modulate ion channel conductance (or any of the other factors mentioned above). The technique is potentially usable with a far broader range of cell types than other techniques that require cell generated action potentials for signal transduction.

NG108-15 (neuroblastoma x glioma) hybrid cells [8] have been utilized extensively for many types of morphological and biochemical studies. They express at least four major families of voltage-sensitive channels, including voltage-gated sodium channels, that respond to a variety of ion channel blockers [9]. Tetrodotoxin (TTX) is a highly specific blocker of voltage gated Na^+ channels. TTX binds to the outer mouth of the channel through which sodium ions move, thereby physically blocking ionic current through the channel [10]. Thus it is a useful substance for testing the applicability of impedance imaging for hybrid biosensors.

Chemical and biological warfare (CBW) agents pose a difficult detection problem due to the broad (and ever expanding) spectrum of materials that may be used. Thus, there is a need for detection systems that exhibit high sensitivity while maintaining diverse detection capabilities. Whole cell based biosensors provide this critical combination. The use of impedance imaging allows for a large number of different cell types to be cultured, thereby allowing for a tunable sensor response. This technology promises to be a complementary component to existing CBW detection systems. Also of great interest is the use of such techniques for pharmaceutical screening, both for intended agent effects and for toxicological studies. A variety of basic biological studies could also potentially benefit from this approach. In the latter two cases, this type of assay could reduce the number of animal-based tests required.

MATERIALS AND METHODS

A 6x6 array of 10 μm diameter planar microelectrodes (100 μm pitch) has been fabricated for cultured cell studies. A 500 nm thermal oxide provides electrical isolation from the underlying silicon substrate. Iridium electrodes and gold interconnects and bondpads are deposited utilizing lift-off techniques, followed by deposition of a 1 μm low stress PECVD silicon nitride passivation layer. Vias are etched through the nitride to define the electrodes and bondpad openings. The 9x9 mm chips are packaged in a standard 40-pin ceramic dual in-line package. The bondwires are encapsulated (for mechanical and electrical robustness) in a nontoxic low stress epoxy.

An impedance imaging system (Figure 1) has been developed to monitor the impedance of cell / electrode interfaces across the electrode array. A quadrature synthesizer is used to generate both sine and cosine signals of programmed frequency and amplitude. The 10V P-P sinewave signal is attenuated as needed (from 0dB to -80dB) and connected to an individual electrode via an analog multiplexer. The resulting current is monitored by a transimpedance stage which holds the large reference electrode at virtual ground. Automatic gain control (0dB to 100dB) is used to amplify the voltage output of the transimpedance stage. This amplified signal is multiplied in quadrature by the source signal and low pass filtered (fourth order Butterworth at 10Hz) to provide the real and imaginary components of the measurement. Quadrature multiplication (a common lock-in amplifier technique) allows for signal detection (at the excitation frequency) with high noise immunity. Known value resistors are used to calibrate the system for electrode impedance measurements.

0-9640024-1-8 $3.00 © 1996 TRF

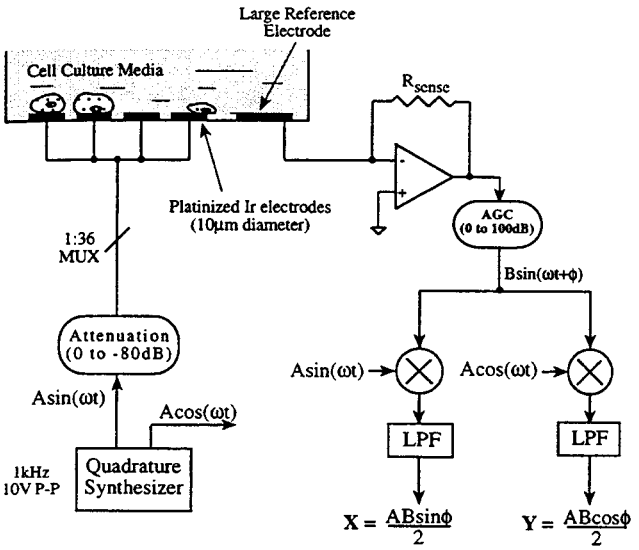


Figure 1. Computer controlled system used for monitoring the impedance of cell / electrode interfaces. An AC voltage is used to drive one small electrode at a time; the resulting current is converted to a voltage, amplified, and homodyned to yield the real and imaginary parts of the electrode impedance. The small electrode dominates the measured impedance.

The temperature of the substrate is regulated by an analog feedback loop using a cold-junction compensated thermocouple and a solid state Peltier device **[11]**. Achieved regulation at 37°C is better than ± 0.3°C.

NG108-15 cells are cultured on the electrode arrays in serum containing media using standard cell culture protocols **[9]**. The chips are placed in a CO_2 incubator (37°C, 10% CO_2) for 24 to 48 hours (long enough to ensure adequate adhesion to the electrodes). For impedance measurements, the chips are removed from the incubator and placed into the measurement system where the temperature is maintained at 37°C. While atmospheric CO_2 concentration is not controlled, buffers in the culture medium adequately maintain the pH for the duration of the impedance measurements.

A 1kHz, 100mV P-P sinusoidal wave (small enough to ensure linear electrode characteristics **[12]**) is applied to a single electrode in the array. The resulting current flowing through a large reference electrode into a virtual ground is used to calculate the effective impedance of the cell / electrode interface combination. Electrode impedance is monitored across the array at approximately three minute intervals. The culture media is exchanged for new media with the same formulation in order to "normalize" the solution pH for subsequent addition of toxins prepared in the same solution (Significant pH changes can occur during the 24 to 48 hour culture period due to cell metabolic processes; this effect must be normalized out to perform meaningful cellular response studies). Culture media containing the substance to be tested is exchanged for the existing media and the resulting impedance changes are monitored over time. After a leveling of impedance response, the test solution is rinsed twice and replenished with new cell culture media. Again, the resulting impedance is monitored. Finally, the cells are rinsed an additional time to ensure full removal of any remaining test solution. It is important to note that this may not guarantee full reversal of a toxin's effects.

Impedance measurements across the array allow for determination of cell position and coupling to electrodes when the measurements are normalized to data taken prior to cell plating. Where detection of environmental factors or other external stimuli is desired, the impedances are normalized to the first measurement of a series.

This entire sequence was repeated for an identical control chip which underwent all treatments and culture protocols without the addition of cells. This data ensures impedance changes observed with changing environmental factors are due to biological effects rather than changes in adsorbed substances on the electrodes themselves.

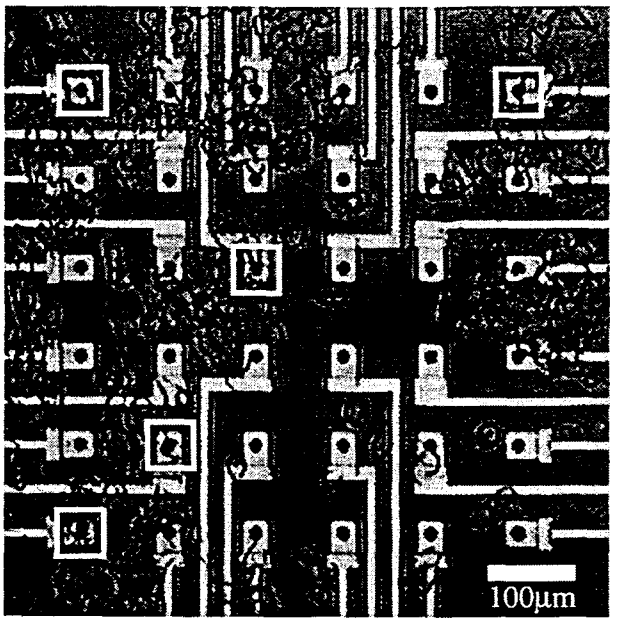
RESULTS

System performance and variability were evaluated by measuring a 1MΩ, 1% metal film resistor in place of the electrode array (measured electrode impedances at 1kHz typically range from 10kΩ to 2MΩ depending on cell coupling). With twenty trials, maximum single channel variations of 0.11% and channel-to-channel variations of 0.12% in impedance magnitude were observed. This is indicative of the anticipated measurement noise due to the electronics.

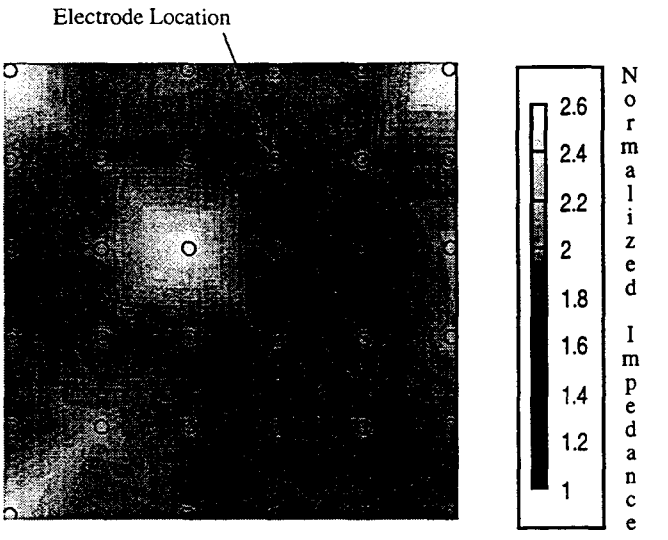
Impedance measurements normalized to the bare chip impedances were used to directly correlate cell placement over electrodes. Figure 2 presents impedance data for healthy cells growing on the electrode array. Regions with cells positioned directly over an electrode (highlighted with boxes in the optical image) correspond to regions of high impedance as can be seen from the 2-D impedance map in Figure 2b. In applications where action potential signals are to be measured from the cultured cells, tight cell / electrode coupling is desirable for improved signal-to-noise ratio (SNR) **[13]**. Thus, impedance imaging may be used as a complementary technology to determine which electrodes are covered by tightly coupled cells and should be monitored for optimal SNR.

Lysing (causing cellular death by generation of different osmolality inside and outside of the cell, resulting in rupture of the cellular membrane) the cells by perfusion with a hyperosmolar solution resulted in the data presented in Figure 3. A clear decrease in measured impedance across the array is evidenced, corresponding to lifting of cells and partial removal of adsorbed biological substances from the electrodes. While this is a drastic demonstration of cellular viability assessment, it also illustrates the sensitivity of the impedance measurements to adsorbed materials and the necessity of controls.

To examine the more subtle effects of ion channel blockage, single electrode impedances were monitored over time as such toxins were added. Optical images of several representative electrodes are shown in Figures 4a and 5a at the beginning of the experiment. The impedance magnitude of the indicated electrodes are tracked in Figures 4b and 5b. Small effects were observed as the culture media was replaced with fresh medium, probably due to pH differences between the two solutions. Cellular metabolism results in acidification of the culture media and becomes more pronounced over time as buffering action is depleted. Thus, perfusion with fresh solution

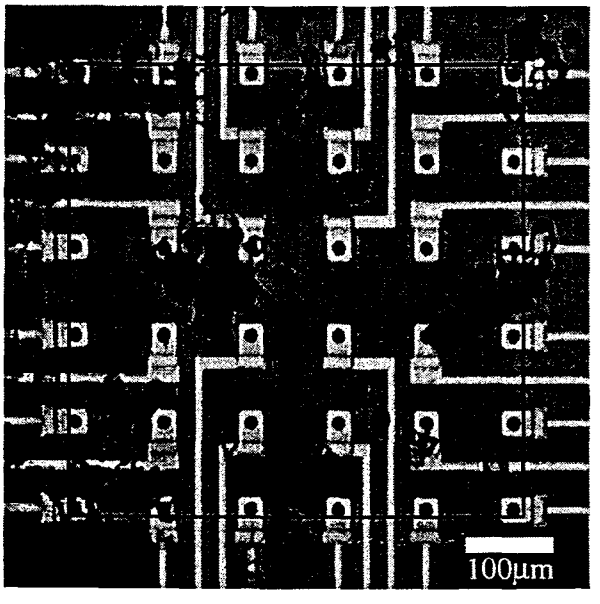


(a)

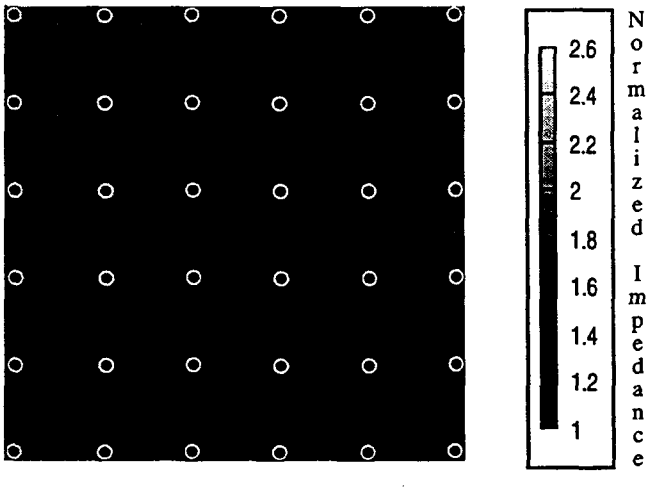


(b)

Figure 2. Cultured NG108-15 hybrid cells. a). Optical photograph of cells growing on the electrode array. Boxes indicate electrodes with apparent tight cell / electrode coupling as seen by the impedance map in (b). b). 2-D impedance magnitude map normalized to the bare chip impedance. Interpolation was used to form contours between discrete data points.



(a)



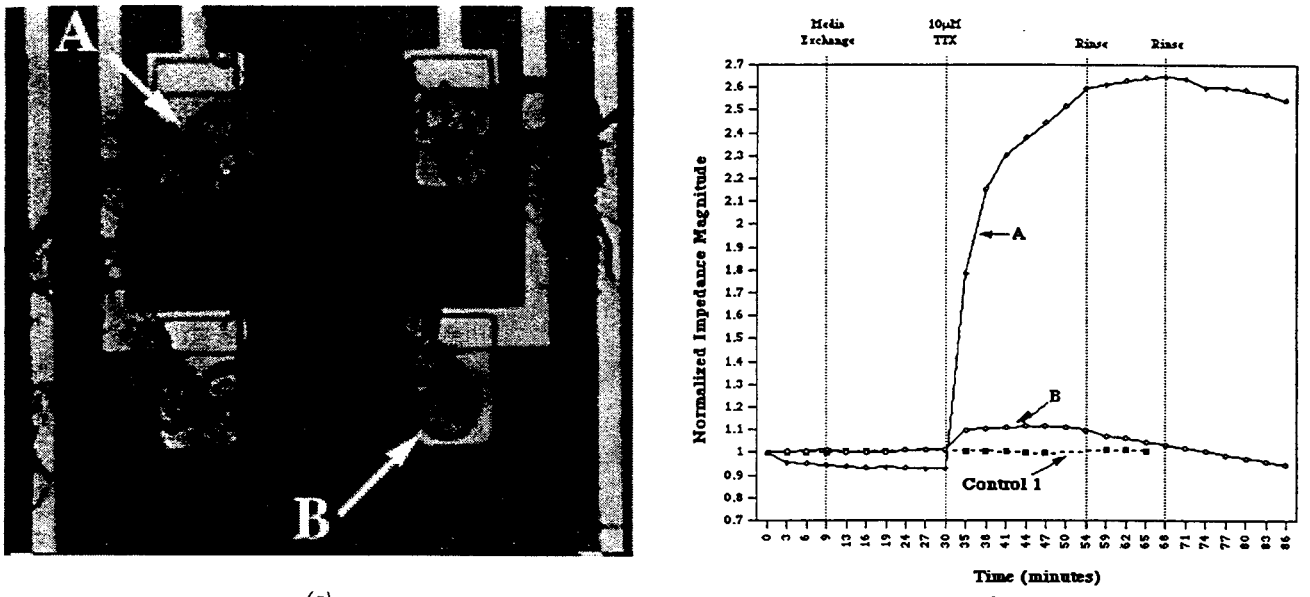
(b)

Figure 3. NG108-15 cells following lysing in a hyperosmolar solution for five minutes. a). Optical photograph of cells on the array. b). 2-D impedance magnitude map normalized to the bare chip impedance. Note overall decrease in impedance due to lifting of cells and partial removal of adsorbed biological substances.

results in increased pH which is evidenced by the change in measured impedance.

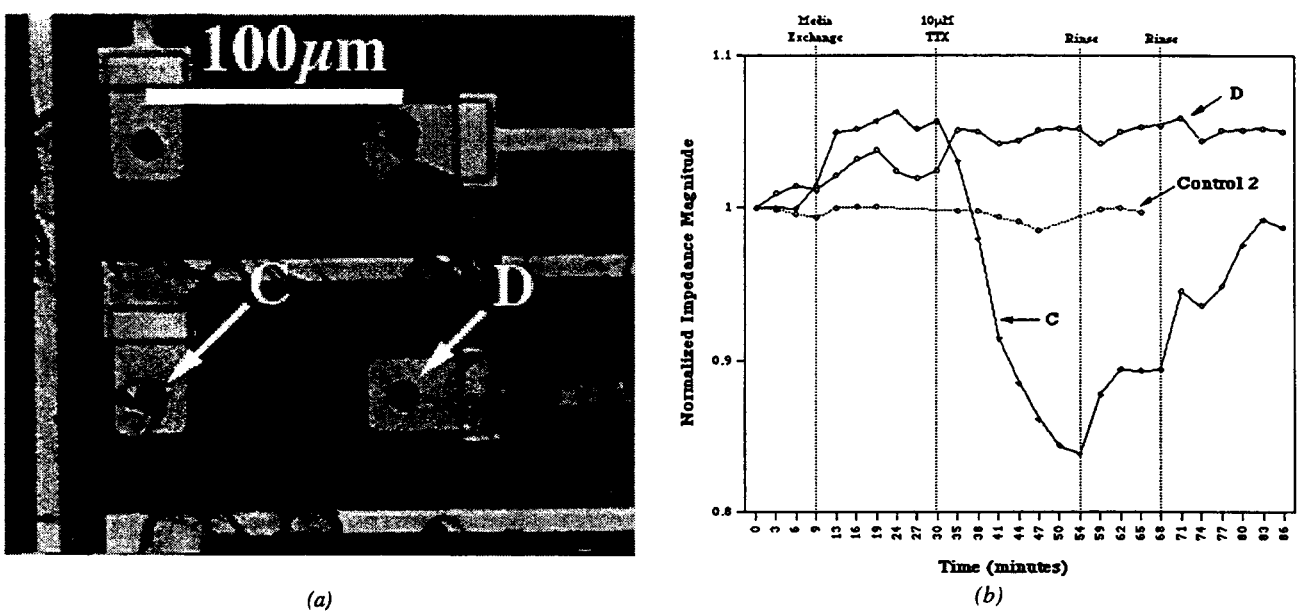
The addition of the sodium channel blocker tetrodotoxin (TTX) resulted in qualitatively different responses from different electrode / cell combinations. The most pronounced effects may be seen with traces A (Figure 4b) and C (Figure 5b). With the addition of TTX, the cell clump over electrode A (Figure 4a) responded with a drastic increase (180%) in measured impedance magnitude, most likely corresponding to blockage of voltage-gated Na^+ channels. As these channels are blocked, the resulting impedance is expected to increase due to decreased resistive conductance through them. This particular clumped cell response did not recover upon removal of the TTX solution.

The optical image of Figure 4a shows complete electrode coverage by a single cell on electrode B. The response in Figure 4b follows that described by trace A except that the cell recovers from the TTX exposure. It is important to note that statistically significant data has not been gathered on variability of cell



(a) (b)

Figure 4. (a). *Optical image of cells plated over electrodes A and B. Note multiple cells positioned over electrode A while a single cell completely covers electrode B. (b) Normalized impedance magnitude for electrodes A, B and Control 1. Note near constant response for Control 1 (representative response of control electrodes without cells), while both A and B respond with an increase in impedance magnitude with the introduction of TTX. Impedance magnitudes were on the order of 10-50kΩ.*



(a) (b)

Figure 5. (a). *Optical image of cells plated over electrodes C and D. Note single cell partially covering electrode C and no cells covering electrode D. (b). Normalized impedance magnitude for electrodes C, D and Control 2. Note near constant response for Control 2 (representative response of control electrodes without cells) , minimal response for D (no cell over electrode) and a decrease in impedance magnitude for electrode C corresponding to the introduction of TTX. Rinsing the TTX away results in a net rise in electrode C impedance magnitude. Impedance magnitudes were on the order of 50 kΩ to 200 kΩ.*

sensitivity to TTX and its reversible effects as related to impedance measurements of this type.

Electrode C exhibited a qualitatively different response to the TTX addition, with a resulting impedance magnitude decrease of 20%. Removal of the TTX solution resulted in an impedance magnitude rise to a constant level 10% below the pre-TTX value. A second rinse resulted in increased impedance to within 1.6% of the pre-TTX value. The mechanism for this behavior is not yet understood. Note minimal changes on an adjacent electrode without coupled cells (electrode D).

DISCUSSION AND SUMMARY

A system has been presented that monitors cellular viability and position by impedance imaging, demonstrating potential use in a variety of sensors. Maps of electrode impedance have been shown to correlate directly to cell positioning over an electrode array. While a thorough investigation of the mechanisms affecting cellular response to a well known voltage-gated Na^+ channel blocker (TTX) has not been undertaken to date, the results presented indicate significant cellular response as compared to control electrodes without cells. The effective use of electrodes small enough (10 μm diameter) to study single cell / electrode coupling has been demonstrated. Investigations of this and other agents along with tight control (and purposeful variation) of environmental factors utilizing this system's full capabilities (frequency variation over five decades and signal source amplitude variation) should allow for development of detailed models of cell / electrode coupling and the mechanisms affecting cell membrane impedance.

By utilizing different cell types, it is possible to tune the sensitivity of the sensor to a specific application. Non-electrically active cell types such as hepatic cells could be used for toxin detection. For sensors utilizing electrically active cells, impedance imaging may be used to determine which electrodes are covered by cells (and therefore should be monitored), and as a secondary monitor of cellular viability. Possible application areas include: neurotoxin detection, cell cytotoxicity testing, pharmaceutical screening, and environmental monitoring.

ACKNOWLEDGMENTS

This work was funded through a combination of G. Kovacs' NSF NYI Award (No. ECS-9358289) and Terman Fellowship (courtesy of Mr. D. Packard and Mr. W. Hewlett) and a grant from the Office of Naval Research (No. N00014-93-1-0759). D. Borkholder is supported by an Air Force Fellowship.

REFERENCES

[1] G.W. Gross, "Simultaneous single unit recordings in vitro with photoetched laser deinsulated gold multimicroelectrode surface", *IEEE Trans. Biomed. Eng.* 26, 273-279 (1979).

[2] M. Kowolenko, C.R. Keese, D.A. Lawrence, I. Giaever, "Measurement of macrophage adherence and spreading with weak electric fields", *Journal of Immunological Methods*, Vol. 127, 71-77 (1990).

[3] C.R. Keese, I. Giaever, "A whole cell biosensor based on cell-substrate interactions", *Annual International conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 12, No. 2, 1990.

[4] P.M. Ghosh, C.R. Keese, I. Giaever, "Monitoring electropermeabilization in the plasma membrane of adherent mammalian cells", *Biophys. Journal*, Vol. 64, May 1993, 1602-1609.

[5] C.R. Keese, I. Giaever, "A biosensor that monitors cell morphology with electrical fields", *IEEE Engineering in Medicine and Biology*, June/July 1994, 402-408.

[6] I. Giaever, C.R. Keese, "Use of electric fields to monitor the dynamical aspect of cell behavior in tissue culture", *IEEE Transactions on Biomedical Engineering*, Vol. BME-33, No. 2, Feb. 1986.

[7] P. Connolly, G.R. Moores, W. Monaghan, J. Shen, "Microelectronic and nanoelectronic interfacing techniques for biological systems", *Sensors and Actuators B*, 6 (1992) 113-121.

[8] C.N. Christian, et al., "Synapse formation between two clonal cell lines", *Science*, 196: pp.995-998 (1977).

[9] V.C. Kowtha, et al., "Comparative electrophysiological properties of NG108-15 cells in serum-containing and serum-free media", *Neurosci. Lett.*, 164: pp.129-133 (1993).

[10] B. Hill, *Prog. Biophys. Mol. Biol.* 21: 1-32 (1970).

[11] N.I. Maluf, et al., "A thermal signal generator probe for the study of neural thermal transduction", *IEEE Transactions on Biomedical Engineering* (July 1994) vol.41, no.7, pp. 649-55.

[12] G.T.A. Kovacs, "Microelectrode models for neural interfaces", pp.121-165, in D. A. Stenger, T. M. McKenna (Eds) (1994), "Enabling Technologies for Cultured Neural Networks", Academic Press, San Diego, CA.

[13] Nicholls, Martin, and Wallace, "From Neuron to Brain", (1992) pp. 32-34, Sinauer Associates, Sunderland, MA .

USING VIDEO MICROSCOPY TO CHARACTERIZE MICROMECHANICS OF BIOLOGICAL AND MAN-MADE MICROMACHINES

Dennis M. Freeman freeman@mit.edu
C. Quentin Davis qdavis@mit.edu

Department of Electrical Engineering and Computer Science and
Research Laboratory of Electronics,
Massachusetts Institute of Technology, Cambridge, Massachusetts 02139;
Eaton-Peabody Laboratory of Auditory Physiology,
Massachusetts Eye and Ear Infirmary, Boston, Massachusetts 02114.

ABSTRACT

Microscopic mechanics at the scale of individual sensory receptor cells play an important signal processing role in the inner ear. However, there are few direct measurements — in large part because the measurements are technically challenging. The important structures are small (micrometers) and the motions are even smaller (nanometers) as well as fast (audio frequencies from 20 Hz to 20 kHz). Furthermore, to relate the motions of different structures, simultaneous measurements from many targets are important.

We have developed a new method for visualizing and measuring the sound-induced motions of inner-ear structures. We use (1) a light microscope to magnify the image of the inner ear and to project the image onto a CCD camera, (2) optical sectioning to obtain sequences of images that characterize the entire 3D structure, and (3) stroboscopic illumination to create slow motion video sequences in response to audio-frequency stimuli. Even after optical magnification, sound-induced displacements of inner-ear structures are small compared to the pixel spacing of modern CCD cameras. Nevertheless, subpixel motions modulate the brightness field. We have modified robot vision algorithms to determine motion from changes in the brightness field.

Although this system was designed to study mechanics of the inner ear, it has broader application. We illustrate possible applications in MEMS with measurements of the three-dimensional motions of the teeth in a microfabricated comb drive.

BACKGROUND

During the past 15 years, microelectromechanical systems have evolved from scientific curiosities to viable commercial products, holding promise for revolutionizing the fields of sensor and actuator design. Enormous progress in molecular biology and cell physiology during those same 15 years has enabled scientists to further understand the pervasiveness and elegance of design of biological microelectromechanical systems. Our research group is studying one such system — the cochlea, which is the part of the inner ear that responds to sound.

The cochlea is a remarkable signal processing unit. It has a single mechanical input: the piston-like motion of the stapes, which is the last in the chain of middle ear bones that move when sound strikes the eardrum. Therefore, one could view the cochlea as a simple displacement sensor — albeit, one with extraordinary sensitivity [31; 8]. Using neural signals generated by the cochlea, humans can reliably detect sounds that cause motions of the stapes on the order of picometers — i.e., on the order of the diameter of a hydrogen atom. However, the cochlea is more than a simple displacement sensor [26]. Its neural outputs manifest considerable signal processing. The cochlea performs a high-quality frequency analysis of low-level sounds: $Q_{10dB} > 600$ has been recorded for motions in the saline-filled cochlea of the bat [22]. Furthermore, cochlear signal processing is profoundly nonlinear, even at the threshold of hearing. One important nonlinearity is compression. Humans detect sounds differing in intensity over a 120 decibel range. This range of intensities is compressed as it is encoded for neurons with considerably smaller (20-50 dB) dynamic ranges. Our group is interested in understanding the micromechanical mechanisms underlying these signal processing capabilities of the cochlea.

Cochlear mechanics are organized hierarchically. Sensory reception in the cochlea is mediated by hair cells [16; 17], which are specialized for mechanoreception by a microscopic bundle of mechanically sensitive hairs (Figure 1). Rotations of a hair bundle cause changes in the intracellular potential of the hair cell, thereby stimulating the release of a neurotransmitter that causes action potentials of cochlear neurons.

The mechanisms by which motions of hair bundles change intracellular potential involve more microscopic mechanical events. Like all cells [32; 33], hair cells are surrounded by a cell membrane. The cell membrane consists of a lipid bilayer that acts as an electrical and chemical barrier between the inside and outside of the cell. In ad-

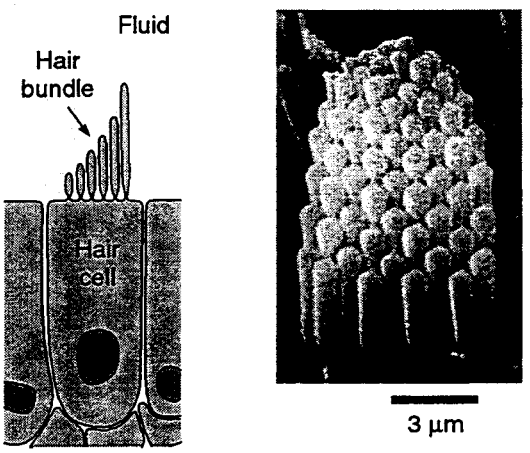


Figure 1. Structure of a hair cell. The left panel shows a schematic diagram of a hair cell and its mechanically sensitive hair bundle. The scanning electron micrograph in the right panel shows a surface view of a hair bundle in the cochlea of the lizard viewed in the direction of the arrow. Each bundle contains 50 to 100 sensory hairs.

0-9640024-1-8 $3.00 © 1996 TRF

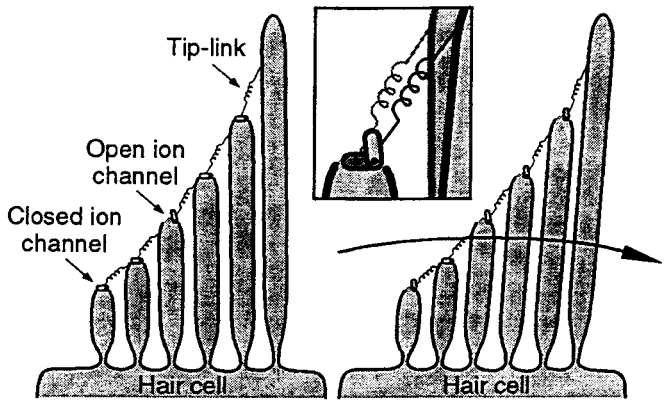


Figure 2. Schematic representation of stress-gated ion channels in hair cells. Rotation of the hair bundle (right) changes the mechanical stress in tip-links, which are represented here as springs connecting adjacent sensory hairs. The change in tip-link stress modulates the ionic conductivity of ion channels, schematized here as trap-doors. The inset shows a magnified view near one tip-link. The gray and black images were taken from the left and right panels, respectively.

dition, the cell membrane contains integral membrane proteins. One class of membrane proteins, the ion channels, act as pores through which ionic currents can pass between the inside and outside of the cell. Hair cells contain ion channels whose state of ionic conductivity is modulated by mechanical stress. The mechanical stress that acts on these stress-gated ion channels is believed to be generated by filaments called tip-links [**27; 18**], which connect the tips of adjacent sensory hairs (Figure 2).

Fluid dynamics plays a key role in cochlear micromechanics. Hair cells do not operate in isolation but are organized along with other cells and tissues to form a sensory epithelium (Figure 3). Displacements of the stapes cause pressures in the fluids that act to deform the sensory epithelium. As the tissues deform, shearing motions between the hair cells and overlying tectorial membrane cause rotations of hair bundles. Forces of fluid origin play a key role in determining motions of this coupled, hydromechanical system. Interaction of mechanical properties of the epithelium with inertial fluid forces accounts for much of the frequency selectivity exhibited in motion of the epithelium [**25; 29; 30**]. Fluid forces play a key role in determining the motions of hair bundles [**9; 5**] and contribute to a micromechanical resonance of some hair bundles of lizards [**10**].

Cochlear structures are extensively cross-coupled. Although organized hierarchically, the levels are not independent. For example, the state (open or closed) of stress-gated ion channels (nanometer scale) has been shown to affect the stiffness of the hair bundle (micrometer scale) [**15**]. Motions of hair bundles are coupled both mechanically and hydrodynamically to one another and to the sensory epithelium (millimeter scale). Through the surrounding fluids, hair cells can also interact chemically and electrically.

In humans, the cochlear micromachine consists of approximately 15,000 hair cells, each with more than 50 sensory hairs and associated tip-links in a saline-filled environment. The challenge is to understand how these parts — more than a million in each ear — interact to form a system with the remarkable signal processing capabilities of the ear.

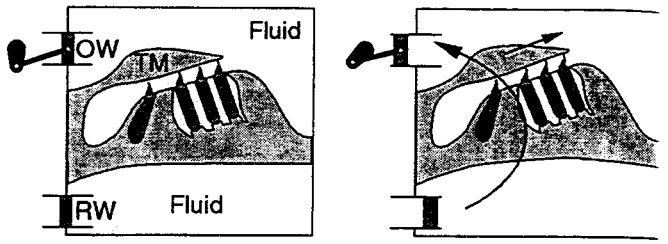


Figure 3. Schematic cross-section of the cochlea. The sensory epithelium and overlying tectorial membrane (TM) separate two fluid-filled compartments. One is coupled to the oval window (OW), into which the piston-like foot of the stapes fits; the other is connected to the round window (RW), which is covered by a thin membrane. The right panel illustrates deformations caused by motions of the stapes.

METHODS

We have developed a video system to measure sound-induced motions of cochlear structures in an in vitro preparation of the alligator lizard cochlea [**6**]. The cochlea is mounted in an experimental chamber so that it can be viewed with a microscope while simultaneously being stimulated by an underwater pressure source (Figure 4). The chamber is mounted on the stage of a light microscope (Zeiss Axioplan, New York, NY) that projects magnified images of cochlear structures onto a scientific grade CCD camera (Photometrics 200 series with a Thompson 7883 CCD, Tucson, AZ). The camera has 576 × 384 pixels with 23 μm spacing. Total magnification between the target and camera can be adjusted from 5 to 600.

Optical sectioning. Many biological tissues, including hair cells, are nearly transparent. Images can be obtained from any part of such a tissue by altering the microscope's focus control. Thus images from a series of focal depths can be combined into a three-dimensional image that characterizes all of the structures in the cochlea, including hair cells, hair bundles, and the tectorial membrane. Optical sectioning can also be used to generate three-dimensional images of opaque

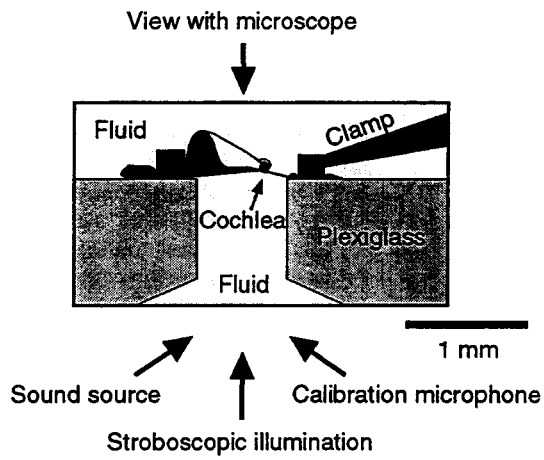


Figure 4. Cross-sectional view of the experimental chamber. The cochlea is clamped so that it seals a circular opening between two fluid-filled compartments. The cochlea is stimulated by a piezoelectric disk that generates hydrodynamic pressures in the bottom compartment; the top compartment is vented to the atmosphere. Sound pressure is calibrated using an underwater microphone inserted just under the cochlea.

targets. However, only unoccluded surface structures are visible. To automate the imaging of three-dimensional structure, a computer-controlled stepper motor is used to adjust the focus control of the microscope with a resolution of 1/11 μm per step.

Stroboscopic illumination. The important motions of cochlear structures are at audio frequencies, i.e., from 20 Hz to 20 kHz. These frequencies exceed the maximum rates of scientific-grade (12-bit) video imagers (it takes 3 seconds for our system to digitize a full frame image and save it to the disk). Therefore, stroboscopic illumination is used to slow the apparent rate of motion. We typically use a gas discharge strobe lamp (Chadwick-Helmuth Co. Inc., El Monte, CA) which generates high intensity illumination. However, the light from a gas discharge strobe is not spatially uniform and pulse-to-pulse variations can cause motion artifacts. Therefore, light from the strobe is coupled to the microscope through a fiber optic scrambler (Technical Video Ltd., Woods Hole, MA). The scrambler also isolates our system from the considerable electromagnetic interference that is generated by the strobe. The duration of the strobe pulse (8 μs) limits the maximum frequency of motion that our system can measure. To relax this constraint, we have recently developed a stroboscopic light source using an LED (AND Division of Purdy Electronics Corp., Sunnyvale, CA). Although they generate less intense light, LEDs radiate much less electromagnetic interference and are orders of magnitude faster.

Subpixel motion estimates from video images. Fundamental limitations on image resolution and on motion detection from images are different. Image resolution is limited by the optics [12]. To a first approximation, images are blurred by convolution with a point-spread function [11; 2], which limits resolution to distances on the order of the wavelength of light (typically 550 nm). However, one can detect motions of a blurred image that are significantly smaller. The fundamental limitations on motion detection are determined by the quantum nature of light (shot noise) and by the video imager.

Subpixel displacements can be estimated from changes in brightness (Figure 5) using methods originally developed for robot vision [13]. Generally we identify regions of interest that contain a single moving structure. Images at successive times are then analyzed to determine the best rigid-body translation in a least-squares sense [14].

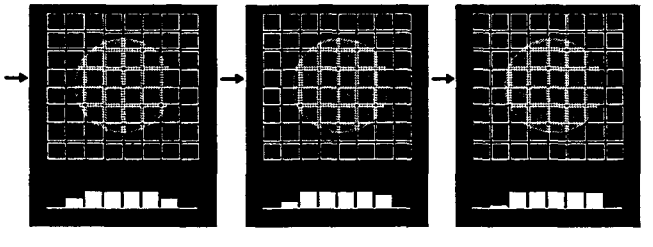


Figure 5. Variations in brightness caused by subpixel motions. The array of squares in each panel represents the pixels of a video camera; the gray disk represents the image of a simple scene; and the bar plots represent the brightness of the row of pixels indicated by the arrow. The three panels illustrate rightward motion of the disk. Although the displacements are subpixel, information about the motion is available from changes in the brightness of the pixels near the edges.

RESULTS

Sound-Induced Motions of Cochlear Structures

Figure 6 shows the cochlea of an alligator lizard as imaged with our system. The hair bundles of hair cells and the overlying tectorial membrane are clearly visible. Generally 30 to 60 hair cells are contained in the field of view. Approximately 20 can be seen in Figure 6, which shows only a portion of the camera's field of view.

Because the cochlea is nearly transparent, one can view structures both above and below those shown in Figure 6 by changing the plane of focus. In this experiment, images were acquired at 13 planes of focus separated by 3 μm. In general, one can obtain sufficiently many images to determine the three-dimensional structure of any visible structure in the cochlea. Figure 7 illustrates how the sequence of two-dimensional images can be combined to obtain a three-dimensional image of a hair bundle and overlying tectorial membrane.

Three-dimensional images of the cochlea were obtained at 8 phases of the sinusoidal stimulus. The resulting data set consisted of 104 images of 576 x 382 pixels (13 planes times 8 phases) and total data collection time was 5 minutes. This amount of time is large compared to the time needed to determine the motion of a single structure using other measurement methods, such as laser interferometry. However, from the resulting 104 images, one can estimate the three-dimensional displacement waveform for *every* resolvable structure in the field of view. Thus the total measurement time is in fact extraordinarily short compared to the time it would take to make motion measurements of each structure using another method. We illustrate the utility of this feature by showing how two separate analyses of the same data set may be used to address two issues at different spatial scales — the relation between hair bundle motion and that of the overlying tectorial membrane, and relations among the motions of neighboring sensory hairs within a hair bundle.

To determine motions of the hair bundle and overlying tectorial

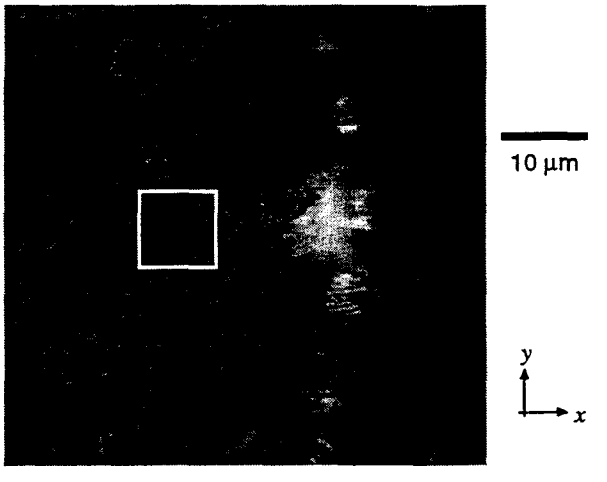


Figure 6. Cochlear structures in the alligator lizard. This figure shows a portion (240 x 240 pixels) of a video image of the lizard cochlea. The white square surrounds one hair bundle; several other hair bundles are also visible. The faint vertical structure at the right edge of the image is the edge of the tectorial membrane (see Figure 3). Total optical magnification from specimen to video camera was 100x.

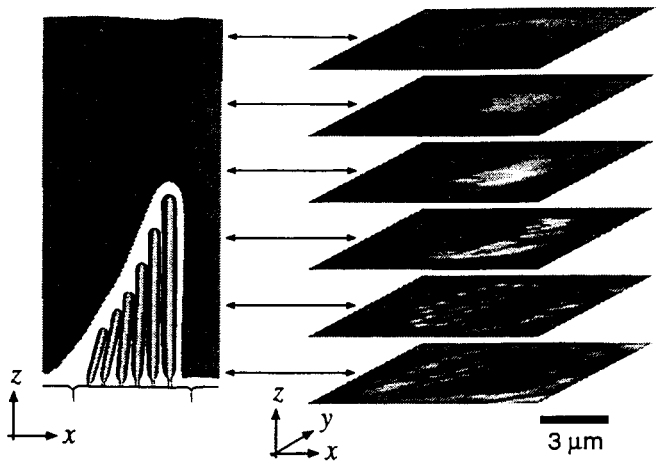


Figure 7. Three-dimensional image of a hair bundle. The bottom 2D image shows a magnified and tilted version of the highlighted part of Figure 6. The other images are from a sequence of planes, each 3 µm above the previous one. The double-headed arrows associate each image with the corresponding plane of section in the schematic drawing on the left. The scale bar applies to both the x and z directions; y is foreshortened.

membrane, we analyzed motions of the structures in the 6 images of Figure 7. Displacements near the base of the hair bundle were larger than those near the tip (Figure 8), which demonstrates that the hair bundle is rotating about its point of attachment to the hair cell. Displacements of the tectorial membrane were generally smaller than those of either the base or tip of the hair bundle. Furthermore, as distance from the base of the hair bundle increases, the motion shows increasing phase lag.

To get a quantitative estimate of hair bundle rotation, each displacement waveform was subtracted from the displacement waveform for the base. The results show that the displacement of the tip of the hair bundle relative to the base is 0.3 µm peak-to-peak. Since the planes through the tip and base are separated by 6 µm, hair bundle rotation is approximately 3° peak-to-peak. Motions of five other hair bundles in this data set were analyzed and the results were qualitatively similar to those shown in Figure 8.

In all of our 12 experiments to date, displacements of the tips and bases of hair bundles have been larger than those of the overlying tectorial membrane for frequencies throughout the lizard's auditory range. This result is consistent with the idea that inertial forces arising from the mass of the tectorial membrane and surrounding fluid tend to resist motion of the tectorial membrane [24]. This result is not consistent with other theories that suggest that the tectorial membrane is a resonant structure [1]. However, our results should be regarded as preliminary. Control experiments must be performed to assure that the cochlea is not damaged during experimentation. Regardless of the ultimate resolution of this particular issue, the important point is that video microscopy can be used to make measurements to directly test theoretical conceptions.

The same data set used in the previous analysis can also be used to study mechanical events at the level of individual sensory hairs. The sinusoidal stimulus generates nearly sinusoidal motions of all the sensory hairs (top right panel of Figure 9). To determine if there is relative motion between the sensory hairs, we applied our motion detection algorithm to estimate the motion of the right edge of the bundle and shifted the images to compensate for that motion. The resulting images show relative motions between sensory hairs (bottom right panel of Figure 9).

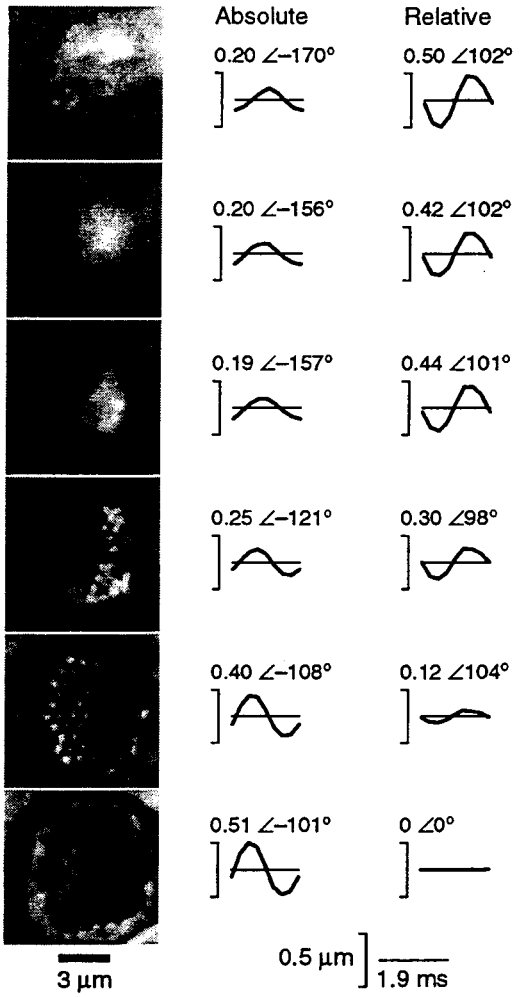


Figure 8. Sound-induced motions at six planes of focus through a hair bundle and overlying tectorial membrane. Images (left column) corresponding to those illustrated in Figure 7 were obtained at 8 phases of the stimulus period and were analyzed to quantify motion at each plane of section. The waveforms in the center column show average displacements in the x direction during one cycle of the stimulus. The associated numbers are the peak-to-peak magnitude (in µm) and angle (in degrees) of the fundamental component of the displacement. The waveforms in the right column show differences between the waveforms in the center column and the bottom waveform in the center column. The intensity of the 513 Hz stimulus (115 dB SPL in the fluid adjacent to the basilar membrane) corresponds to approximately 89 dB SPL at the eardrum.

Although tip-links are too small to be resolved in a light microscope, the results in Figure 9 show that we can directly measure the motions of the sensory hairs to which the tip-links are attached. The relative motions of the tips of the sensory hairs give direct estimates of tip-link stretch, which is the physiologically important mechanical input for transduction mechanisms at the nanometer scale. Although preliminary, our results suggest that tip-links are stretched not only because the distance between insertion points increases when the hairs are rotated by equal angles (as illustrated in Figure 2), but also because the angular displacement of shorter hairs is greater than that of longer hairs. Motions of individual sensory hairs have been

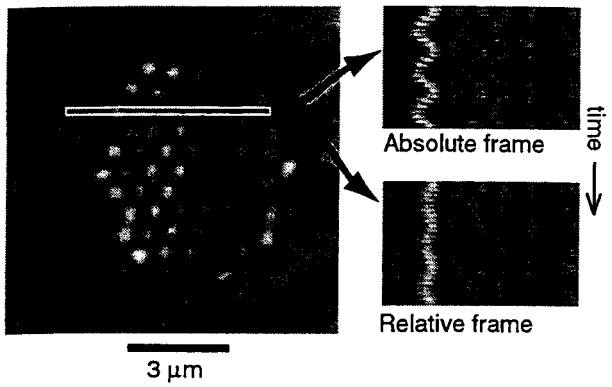


Figure 9. Motions of sensory hairs within a bundle. The left panel is an enlarged image of the hair bundle in Figure 7, at a plane of section near the center of the bundle (second from bottom). The highlighted region was extracted from each of eight images acquired during the stimulus period and positioned so that regions from earlier phases are on top of regions from later phases. The resulting montage, which illustrates motions during one period, is repeated three times to generate the images in the right panels. The upper and lower right panels are similar, except that the images used to generate the lower panel were first shifted to cancel the motion of the right part of the hair bundle.

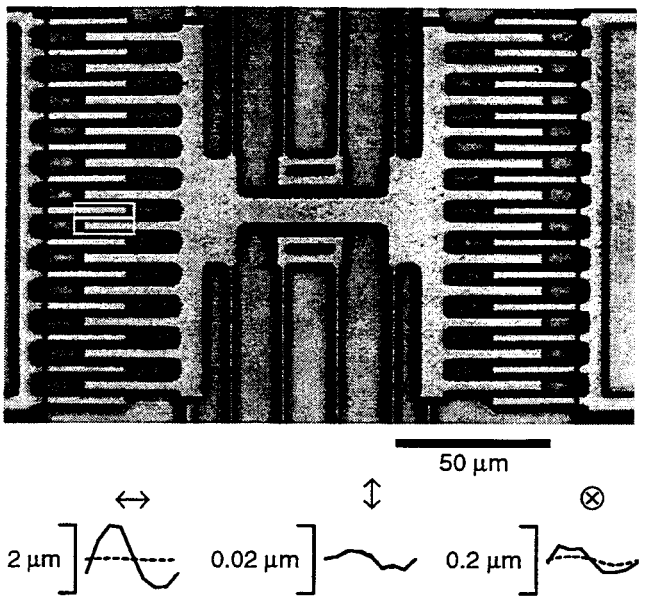


Figure 10. Motions of teeth in a comb drive. The field of view of this image (64× magnification) of a microfabricated accelerometer is focused on the central shuttle and associated comb drives. Most of the folded-beam cantilever that supports the shuttle lies above and below the field of view; only a small portion of it is visible near the center of the image. The portions of the images in the white boxes were analyzed to determine the motions of a stationary (top) and moving (bottom) tooth of the comb drive, and results are shown in the three plots as dashed and solid curves, respectively. The left, center, and right panels show average displacements in the horizontal, vertical, and out-of-plane directions, respectively.

measured elsewhere [7; 23] but only for hair bundles that were stimulated with a water jet; the membranes through which natural stimuli are delivered had been removed.

Three-Dimensional Motions of a Comb Drive

Although originally developed to measure motions of cochlear structures, our video system has broader application. We have performed several pilot experiments to determine if it can be used to measure the motions of man-made micromechanical devices.

Figure 10 shows an image of a microfabricated accelerometer obtained using our system. The comb drive was driven with an 8.8 Hz sinusoidal voltage, 44 volts peak-to-peak with an 80 volt DC offset. Images were obtained at 45 planes of focus with 0.45 μm spacing to generate a three-dimensional image. Three-dimensional images taken at 8 phases of the sinusoidal stimulus were analyzed to determine the motions of the stationary and fixed teeth highlighted in Figure 10. Displacements of the moving tooth were approximately 2 μm peak-to-peak in the horizontal direction (left plot, solid line). These displacements are significantly larger than the corresponding displacements of the fixed tooth, which serve as control measurements to assess the stability of the measurement system. Vertical motions (center plot) were about 6 nm peak-to-peak for both the fixed and moving teeth, which is near the noise floor of our measurement system (without averaging).

Our system can also measure out-of-plane motions. Although one cannot use optical sectioning to "see through" opaque structures, three-dimensional images of opaque structures still provide information about three-dimensional structure and motion, as illustrated in Figure 11. Image brightness is greatest when the tooth is "in focus" and decreases as the plane of focus is moved higher or lower. Our motion detection algorithm takes advantage of these changes in brightness to align three-dimensional images of the same tooth taken at different phases of the sinusoidal stimulus. Results indicate that the out-of-plane displacement of the moving tooth was 80 nm peak-to-peak. The out-of-plane displacement of the stationary tooth was much smaller (26 nm peak-to-peak).

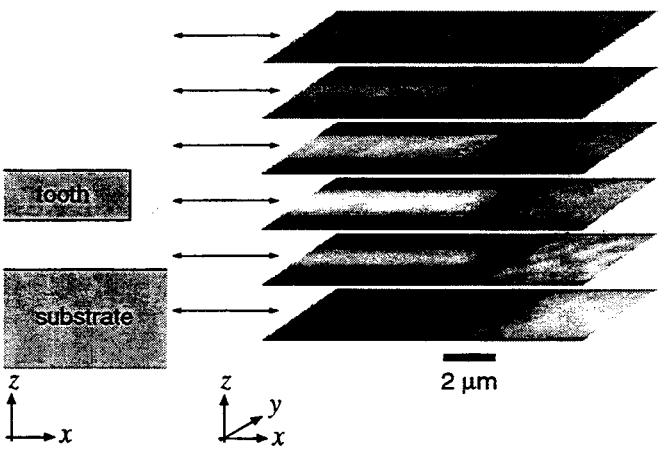


Figure 11. Three-dimensional image of a tooth in a comb drive. This figure illustrates 6 images of the stationary tooth highlighted in Figure 10. The images are at different planes of focus, separated by 2.27 μm. The double-headed arrows associate each image with the corresponding plane of section in the schematic drawing on the left. The scale bar applies to both the x and z directions; y is foreshortened.

DISCUSSION

We have developed a video system that can be used to characterize the motions of both biological and man-made micromechanical systems. Using stroboscopic illumination, we observed motions over the entire audio range of frequencies, and can in principle observe motions at much higher frequencies. Using the optical sectioning property of a light microscope, we measured three-dimensional motions of both transparent (biological) and opaque (silicon) structures. Quantitative results were obtained directly from the video images, based on motion detection algorithms originally developed for robot vision.

Implications for hearing research

Using the video system, we have obtained the first direct experimental observations of the relation between sound-induced motions of the tectorial membrane and hair bundle displacement. We have also obtained the first direct experimental observations of sound-induced motions of sensory hairs within a bundle. Although these results must be regarded as preliminary until control experiments have tested the viability of our preparations, they demonstrate how video microscopy can be used to probe important cochlear mechanisms at a variety of spatial scales: from ion channels (nanometers) to hair bundles (micrometers) to the entire cochlea (millimeters). Video microscopy has opened the experimental door to understanding key relations in cochlear mechanics that have previously been addressed only theoretically.

Implications for MEMS

Tools for characterizing the performance of man-made micromachines are primitive by comparison to those available for electronic devices. For many micromachines, tools for assessing internal failure modes do not exist, and only the terminal behaviors can be tested. The inability to characterize internal failure modes has important consequences for both the designer and fabricator. For the designer, it can result in large numbers of prototypes. For the fabricator, it means reduced yields because common techniques such as laser trimming cannot be applied. As the complexity of micromachines increases and as their production becomes a commercial endeavor, developing better tools for characterizing the performance of micromachines will become crucial. Results in this paper demonstrate the potential use of video methods for characterizing the performance of MEMS devices.

Comparison to other motion measurement systems

A number of systems have been developed to measure micromechanical motions. Our method is different from and has several advantages over previous methods, including the Mössbauer method [19], photodiode method [4], capacitive probe [34], and laser interferometry [20; 28]. (1) Our motion estimates are derived directly from sequences of images. Thus all of our quantitative results can be checked qualitatively by simply viewing the slow-motion video sequences. For example, to test that the displacement estimates in Figure 8 were qualitatively correct, each image was shifted by the estimated displacement in a direction to compensate for the motion. The resulting sequence of images, which should be stationary, provided a test that motions were estimated properly. (2) In the video method, the identity of the target whose motion is measured is unambiguous since motion is determined directly from images of the target. All the other methods listed above require additional observations to determine which motion is being measured. (3) Unlike any of the other methods listed above, the video method allows simultaneous measurement of motions of many structures in the same specimen. This is a significant advantage when determining interrelations among motions of different structures. (4) The video method provides estimates of all three components of rigid body translation for every structure in the image. The other methods listed above provide an estimate of just one component of motion. (5) In contrast to some methods (e.g., the Mössbauer method, some laser interferometric methods [3] and some photodiode methods [4]) our method does not require any mechanical contact with the moving specimen. Mechanical contact has been shown to affect motions of cochlear structures [21], and it seems inescapable that similar degradations would occur in high-performance man-made micromachines.

The video method has one important disadvantage: it is particularly sensitive to vibrations of the microscope and camera. To reduce vibrations, we use a pneumatic vibration isolation table, which very effectively reduces vibrations at high frequencies but has little effect for low frequencies (<10 Hz). Although low frequency vibrations contaminate the output signals of all motion detectors, the contamination can generally be filtered out — provided that the frequencies of interest are sufficiently different from the contamination frequencies. However, in the video method, images are acquired at a low rate, typically 0.3 Hz. Low frequency vibrations of the microscope or camera cause displacements of the images that cannot be distinguished from motions near the frequency of interest. These vibrations determine the noise floor of our measurements, which under optimal conditions (e.g., high contrast target, no nearby structures with confounding motions) is approximately 5 nm without averaging — i.e., 1% of the wavelength of the light that generates the images. With averaging we have reduced the noise to less than 1 nm. We are currently working on other methods to reduce this noise floor.

ACKNOWLEDGMENTS

The authors thank Thomas F. Weiss for his contributions throughout this project, RuthAnne Eatock who prepared the scanning electron micrograph shown in Figure 1, and A. J. Aranyosi for helpful comments on the manuscript. This work was supported by grants from the NSF and NIH (R01 DC00238 and P01 DC00119).

REFERENCES

1. J. B. Allen. Cochlear micromechanics – a physical model of transduction. *J. Acoust. Soc. Am.*, 68:1660–1670, 1980.

2. M. Born and E. Wolf. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Pergamon Press, New York, 1975.

3. N.P. Cooper and W.S. Rhode. Basilar membrane mechanics in the hook region of the cat and guinea pig cochleae: Sharp tuning and nonlinearity in the absence of baseline position shifts. *Hearing Res.*, 63:163–190, 1992.

4. A. C. Crawford and R. Fettiplace. The mechanical properties of ciliary bundles of turtle cochlear hair cells. *J. Physiol.*, 364:359–379, 1985.

5. P. Dallos, M. C. Billone, J. D. Durrant, C. Y. Wang, and S. Raynor. Cochlear inner and outer hair cells: Functional differences. *Science*, 177:356–358, 1972.

6. C. Q. Davis and D. M. Freeman. Direct observations of sound-induced motions of the reticular lamina, tectorial membrane, hair bundles, and individual stereocilia. In *Abstracts of the Eighteenth Midwinter Research Meeting*, St. Petersburg Beach, Florida, February 1995. Association for Research in Otolaryngology.

7. R. K. Duncan, H. N. Hernandez, and J. C. Saunders. Relative stereocilia motion of chick cochlear hair cells during high-frequency water-jet stimulation. *Aud. Neurosci.*, 1:321–329, 1995.

8. R. R. Fay and A. N. Popper, editors. *Comparative Hearing: Mammals* in *Springer Handbook of Auditory Research*. Springer-Verlag, New York, 1994.

9. D. M. Freeman and T. F. Weiss. The role of fluid inertia in mechanical stimulation of hair cells. *Hearing Res.*, 35:201–208, 1988.

10. D. M. Freeman and T. F. Weiss. Hydrodynamic analysis of a two-dimensional model for micromechanical resonance of free-standing hair bundles. *Hearing Res.*, 48:37–68, 1990.

11. J. W. Goodman. *Introduction to Fourier Optics*. McGraw-Hill Book Company, New York, 1968.

12. H. H. Hopkins and P. M. Barham. The influence of the condenser on microscopic resolution. *Proc. Phys. Soc.*, 63:737–744, 1950.

13. B. K. P. Horn. *Robot Vision*. MIT Press, Cambridge, MA, 1986.

14. B. K. P. Horn and E.J. Weldon, Jr. Direct methods for recovering motion. *Internatl. J. of Computer Vision*, 2:51–76, 1988.

15. J. Howard and A. J. Hudspeth. Compliance of the hair bundle associated with gating of mechanoelectrical transduction channels in the bullfrog's saccular hair cell. *Neuron*, 1:189–199, 1988.

16. A. J. Hudspeth. The hair cells of the inner ear. *Sci. Am.*, 248:54–64, 1983.

17. A. J. Hudspeth. How the ear's works work. *Nature*, 341:397–404, 1989.

18. A. J. Hudspeth and P. G. Gillespie. Pulling strings to tune transduction: adaptation by hair cells. *Neuron*, 12:1–9, 1994.

19. B.M. Johnstone and A.J.F. Boyle. Basilar membrane vibration examined with the Mössbauer technique. *Science*, 158:389–390, 1967.

20. S. M. Khanna and D. G. B. Leonard. Basilar membrane tuning in the cat cochlea. *Science*, 215:305–306, 1982.

21. S. M. Khanna and D. G. B. Leonard. Measurement of basilar membrane vibrations and evaluation of the cochlear condition. *Hearing Res.*, 23:37–53, 1986.

22. M. Kössl and I. J. Russell. Basilar membrane displacement in the cochlea of the mustached bat. In *Abstracts of the Eighteenth Midwinter Research Meeting*, St. Petersburg Beach, Florida, February 1995. Association for Research in Otolaryngology.

23. R. B. MacDonald and D. P. Corey. Stereocilia bundles of the bullfrog sacculus hair cells do not splay in response to stimulation. In *Abstracts of the Nineteenth Midwinter Research Meeting*, St. Petersburg Beach, Florida, February 1996. Association for Research in Otolaryngology.

24. F. Mammano and R. Nobili. Biophysics of the cochlea: Linear approximation. *J. Acoust. Soc. Am.*, 93:3320–3332, 1993.

25. S. T. Neely. Finite difference solution of a two-dimensional mathematical model of the cochlea. *J. Acoust. Soc. Am.*, 69(5):1386–1393, 1981.

26. J. O. Pickles. *An Introduction to the Physiology of Hearing*. Academic Press, San Diego, CA, second edition, 1988.

27. J. O. Pickles and D. P. Corey. Mechanoelectric transduction by hair cells. *Trends in Neurosci.*, 15:254–259, 1992.

28. M. A. Ruggero and N. C. Rich. Application of a commercially-manufactured Doppler-shift laser velocimeter to the measurement of basilar-membrane vibration. *Hearing Res.*, 51:215–230, 1991.

29. W. M. Siebert. Ranke revisited – a simple short-wave cochlear model. *J. Acoust. Soc. Am.*, 56:594–600, 1974.

30. M.A. Viergever and R.J. Diependaal. Quantitative validation of cochlear models using the Liouville-Green approximation. *Hearing Res.*, 21:1–15, 1986.

31. G. von Békésy. *Experiments in Hearing*. McGraw-Hill, New York, 1960.

32. T. F. Weiss. *Cellular Biophysics: Transport*. MIT Press, Cambridge, MA, 1996.

33. T. F. Weiss. *Cellular Biophysics: Electrical Properties*. MIT Press, Cambridge, MA, 1996.

34. J. P. Wilson and J. R. Johnstone. Basilar membrane and middle-ear vibration in guinea pig measured by capacitive probe. *J. Acoust. Soc. Am.*, 57:705–723, 1975.

COMBINED ORGANIC THERMAL AND ELECTROSTATIC OMNIDIRECTIONAL CILIARY MICROACTUATOR ARRAY FOR OBJECT POSITIONING AND INSPECTION

John W. Suh, Steven F. Glander, Robert B. Darling*, Christopher W. Storment and Gregory T. A. Kovacs

Stanford University, Department of Electrical Engineering
Room CISX 202, M/C 4075, Stanford, CA 94305
*Department of Electrical Engineering, Box 352500
University of Washington, Seattle, WA 98185

ABSTRACT

An organic thin film ciliary microactuator array using independent thermal and electrostatic actuation is described. A polyimide thermal bimorph structure provides for large angle deflection with high load capacity. Electrostatic plates provide a low power hold-down, capacitive sensing and feedback control capability. Integrating four orthogonally oriented actuators into a unit cell and replicating this cell into an array allows for precise movement of small objects in arbitrary directions. The array has immediate applications for precision positioning, alignment, inspection, and assembly of small parts, such as IC dice, with micron scale resolution.

INTRODUCTION

Several groups have described efforts to apply micromachined actuators to component positioning, inspection, and assembly [1-6]. These applications often require considerable excess force and large movement of the actuator. Despite the large number of papers on micromachined actuators, only a few have discussed the fabrication of microactuator arrays which can move more than their own weight.

The limitations of candidate actuation schemes can be lessened by combining two or more. Our approach is to use thermal actuation for the initial large-angle movement of each actuator and then transition to electrostatic actuation for a low-power hold-down mode. While thermal bimorph actuators are well known, it appears that the only previous attempt to incorporate electrostatic actuation in an organic thermal actuator was that of Lin, et al. [7]. Their approach was to apply an electrostatic pull-down potential to an electrically non-insulated Au resistive heater loop of a thermal bimorph actuator. The devices presented herein contain fully independent (and electrically insulated from the external environment) thermal and electrostatic actuators within each element, allowing for uncomplicated implementation of electrostatic actuation and/or sensing in combination with simultaneous thermal actuation. In this paper, the design, fabrication, and testing of such thermal/electrostatic actuators is presented.

FABRICATION

Surface micromachining techniques were used to create the actuators using polyimide as the primary structural material and aluminum as the sacrificial layer. The fabrication process was designed to be compatible with CMOS or BiCMOS circuits which could be pre-fabricated on a silicon substrate. The process consists of nine lithographic steps using eight masks. Process cross-sections are shown in Fig. 1.

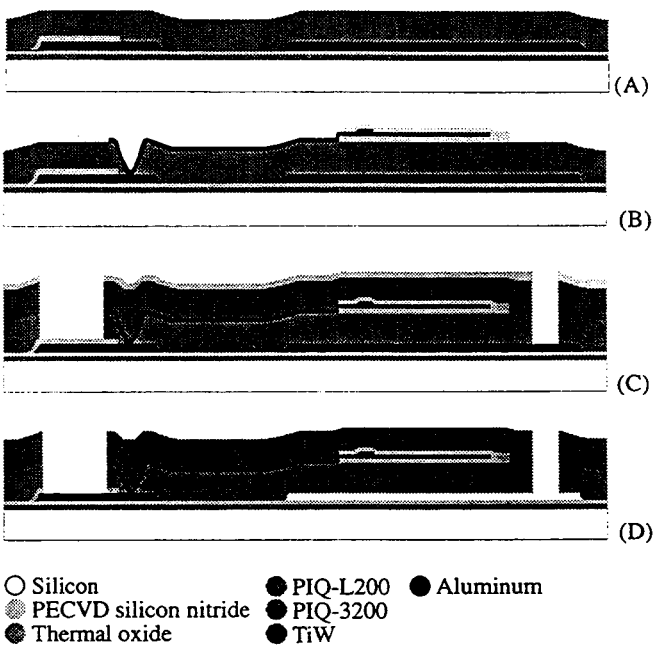


Figure 1. Process flow cross-sections for organic thermal/electrostatic actuator.

The steps up to and including the first polyimide layer are shown in Fig. 1(A). Starting with a thermally oxidized Si wafer, 0.25 μm of Al was sputtered and patterned to form alignment keys (Mask 1). Next, 0.5 μm of LTO was deposited to encapsulate the Al alignment keys. Another 1.6 μm of Al was sputtered and patterned to form sacrificial mesas, interconnects, and bond pads (Mask 2). Since this Al layer would later be exposed to a wet Al etch, a 0.5 μm PECVD silicon nitride masking layer was deposited and patterned to cover the exposed bond pads (Mask 3). An adhesion promoter (Hitachi PI-coupler-3) was spun on and cured at 350°C for 30 minutes followed by spin on of the first (low thermal expansion coefficient) Hitachi PIQ-L200 polyimide. The polyimide was cured in a nitrogen purged oven using the a three step profile: 20 minutes at 105°C, 30 minutes at 160°C, 60 minutes at 350°C. The final, cured lower polyimide thickness was 4 μm.

The following steps describe the processes performed on top of the first polyimide layer shown in Fig. 1(B). A 150 nm PECVD silicon nitride layer was deposited and patterned as the etch mask for the Al interconnect vias (Mask 4). An isotropic oxygen plasma etch was used to etch vias in the first polyimide to the underlying Al. The PECVD silicon nitride was re-patterned to form the lower encapsulation layer for all subsequent metallizations (Mask 5). The silicon nitride also functions as an internal stiffener for the actuator. Next, 100 nm

0-9640024-1-8 $3.00 © 1996 TRF

Solid-State Sensor and Actuator Workshop
Hilton Head, South Carolina, June 2-6, 1996

of 50:50 Ti:W was sputtered and patterned to form the resistive heaters (Mask 6). Electrical connections between the Ti:W and the first Al layer were made by sputtering and patterning 750 nm of Al (Mask 7). Another 150 nm of PECVD silicon nitride was deposited and patterned (reusing Mask 5) to complete metal encapsulation.

The remaining process steps are as follows. The second (high thermal expansion coefficient) Hitachi PIQ-3200 polyimide was spun on and cured in the same manner as the first. The final thickness was also 4 μm after curing. Finally, 250 nm of PECVD silicon nitride was deposited and patterned (Mask 8) to form the actuator outlines, as shown in Fig. 1(C). The wafer was then diced. Afterwards, a RIE oxygen plasma was used to etch through the polyimide and stop either on the PECVD silicon nitride (over the bond pads) or on the sacrificial aluminum mesas. Individual chips were then placed in an aluminum wet etch to release the actuators from the substrate. A final plasma etch was done to remove the PECVD silicon nitride on the top polyimide layer as well as on the bond pads. The final profile is shown in Fig. 1(D).

Because the polyimides were cured at high temperature and because the upper polyimide layer has the larger thermal expansion coefficient, the actuators assume an upward, out-of-plane curl when the process is complete. Heating by the Ti:W resistors flattens them back towards the substrate. The natural tendency for the actuators to curl upwards assists the release process by reducing the chance of sticking and by allowing the etchant to more easily access the area underneath the actuator.

ACTUATOR AND ARRAY DESIGN

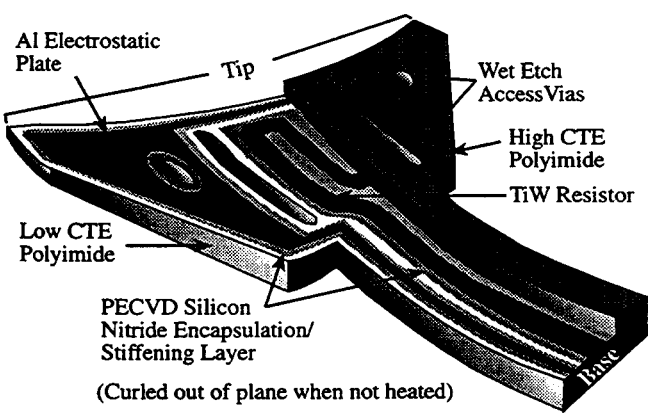


Figure 2. Illustration of a combined thermal and electrostatic ciliary microactuator.

Actuator elements. A single actuator is illustrated in Fig. 2, with half of the upper polyimide and PECVD nitride layers shown removed. The two wing shaped regions contain the electrostatic plates. These wings attach to the tip and mid-section of the central, thermally actuated region.

The radius of curvature of a multi-layered structure as a result of thermal strain is

$$R = \frac{h}{\varepsilon_T - \varepsilon_B} \tag{1}$$

where ε_T and ε_B are the strains on the top and bottom layer of the actuator and h is the combined thickness of the structure. Because the actuator has curvature along its length and width, the strain solutions used were those of a disk exhibiting axially symmetric strains [8, 9]. Relative to its location before release from the substrate, the deflection of the actuator tip involves both vertical and horizontal displacements,

$$\delta_V = R\left(1 - \cos\left(\frac{L}{R}\right)\right) \tag{2}$$

$$\delta_H = L - R\sin\left(\frac{L}{R}\right) \tag{3}$$

where L = 430 μm is the length along the actuator's axis of symmetry when flat. Following release, the actuators assumed a radius of curvature of approximately R = 706 μm, yielding displacements of δ_V = 127 μm and δ_H = 26 μm.

The lifting capacity of an actuator can be estimated as the force required to deflect the actuator's tip to the substrate,

$$F_L = \frac{3}{L^3}(EI)_{equiv}\delta_V \tag{4}$$

where $(EI)_{equiv}$ is the equivalent flexural rigidity of the actuator [10,11]. Using the nominal parameters of the actuator layers listed in Table 1, an actuator load capacity of F_L = 76 μN is obtained.

Table 1. *Layer parameters of fabricated actuator*

Layer #	1	2	3	3	5	6
Material	PIQ-L200	PECVD Silicon Nitride	Ti:W	Al	PECVD Silicon Nitride	PIQ-3200
Thickness	4 μm	150 nm	100 nm	780 nm	150 nm	4 μm
E (GPa)	8.63	70	N/A	7	70	2.95
CTE (ppm/°C)	20	0.8	N/A	23.6	0.8	54

The Ti:W heating elements were designed for a nominal resistance of 1500 Ω, dissipating 16.7 mW for an applied 5.0 V. The wing electrodes were designed for an area of 24,400 μm^2, which yields a capacitance of 300 fF for each actuator, when the actuator is fully flattened onto the substrate.

Chip layout. The omnidirectional ciliary array is composed of a two dimensional array of cells having four orthogonally oriented actuators each (Fig. 3). This 8 x 8 array of cells (each 1 x 1 mm) has a total active area of 0.64 cm^2.

The four actuators of each cell are independently activated by 4 thermal and 4 electrostatic control lines. Within each column of 8 cells, the thermal elements for a given actuator orientation are wired in series, whereas the electrostatic elements for a given actuator orientation are wired in parallel. The eight columns are then wired in parallel for both the thermal and electrostatic elements. Each electrostatic control line (E1 - E4) applies a potential to the wing electrodes of all 64 actuators of a given orientation. Similarly, each thermal control line (T1 - T4) runs a current through the Ti:W heaters of all 64 actuators of a given orientation, such that the load resistance for the overall array is the same as that of a single element (≈ 1500 Ω). A common ground bus provides a current return path for the thermal elements, and a substrate contact provides an opposing plate for the electrostatic elements.

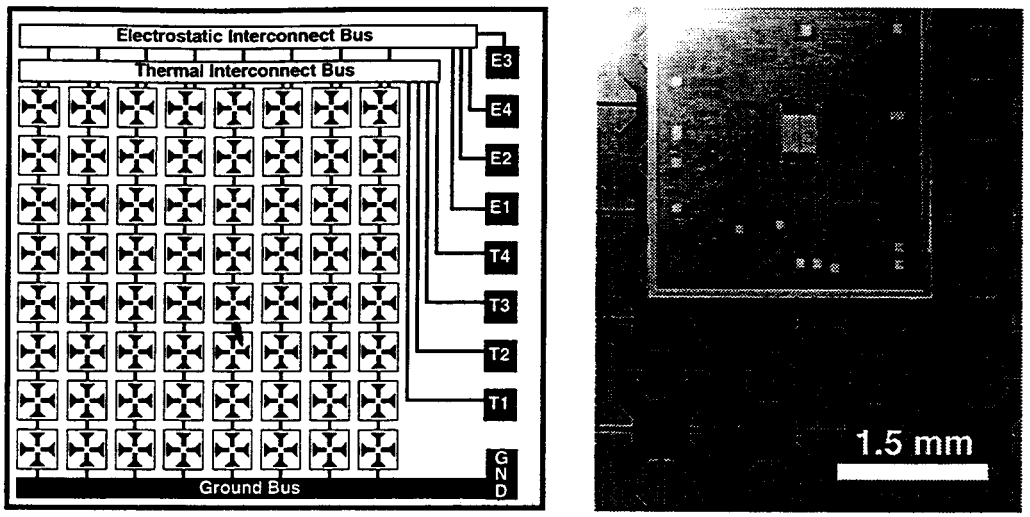


Figure 3. Schematic (not to scale) of the omnidirectional ciliary microactuator array chip (left). The SEM photograph on the right shows part of the array with an ADXL50 accelerometer (courtesy of Dr. R. Payne, Analog Devices, Inc.)

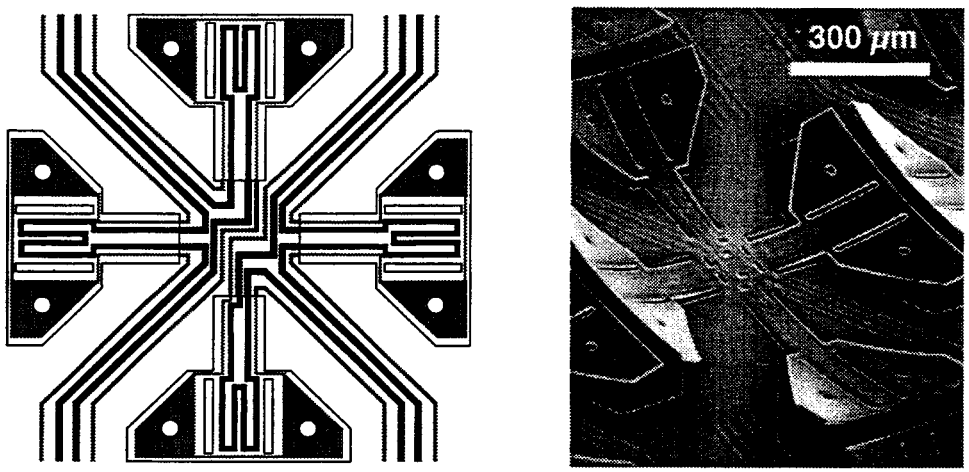


Figure 4. Illustration of a motion pixel (left). Four actuators in a common center configuration make up each pixel. Thicker, darker lines represent Al interconnects for the Ti:W resistors. Thinner, lighter lines represent Al interconnects for the electrostatic plates. The SEM photo on the right shows an oblique view of a motion pixel.

EXPERIMENTAL RESULTS

Lifting capacity. Lifting capacity tests were performed to determine how massive an object the ciliary array could convey. The array successfully moved objects with different masses at widely variable step rates which were independent of their mechanical resonances. The objects included a 3 x 3 mm Si IC (8.6 mg), a 1.0 x 1.5 x 1.5 mm glass block (15 mg), a # 4-40 stainless steel nut (0.16 g) and an 8-pin plastic DIP package (0.5 g). The DIP package exerted nearly 69 $\mu N/mm^2$ which is near the maximum theoretical lifting capacity of a single actuator (76 μN) (even though there are four actuators in a motion pixel, at times only one actuator in a pixel supports an object).

Four phase gait and arbitrary motion direction. Lateral movement of an object across the array is accomplished by applying a coordinated gait to differently oriented actuators. Figure 5 illustrates the basic four phase gait required to move the supported object in a rightward direction. Initially, all actuators are un-deflected and contact the object. Next, the up-down actuator pair and the left actuators are deflected downward to allow the object to rest only on the right actuators. The basic gait is established by applying two square waves to the left-right actuator pair with a relative phase delay of one quarter period. Deflecting the right actuators downward lowers the object by δ_V and translates it to the left by δ_H. De-energizing the left actuators raises the object by δ_V and translates it to the right by another δ_H increment. De-energizing the remaining actuators completes the cycle. The size of the horizontal and vertical displacement steps δ_H and δ_V are controlled by the voltage applied to the thermal actuators.

The above sequence, with appropriate rotation, describes object movement in each of the four primary (up, down, left, and right) directions. By actuating adjacent pairs with the same drive signals, motion along the diagonals can also be directly produced. Motion in any arbitrary direction was produced by a vector-to-raster algorithm based on that of Foley and Van Dam [13]. A PC-controlled system for producing the correct gait combinations and driving the array was developed and tested. The system allows for speed control via the drive pulse frequency and arbitrary directional control via keyboard or mouse. This system was used for performing the tests detailed below.

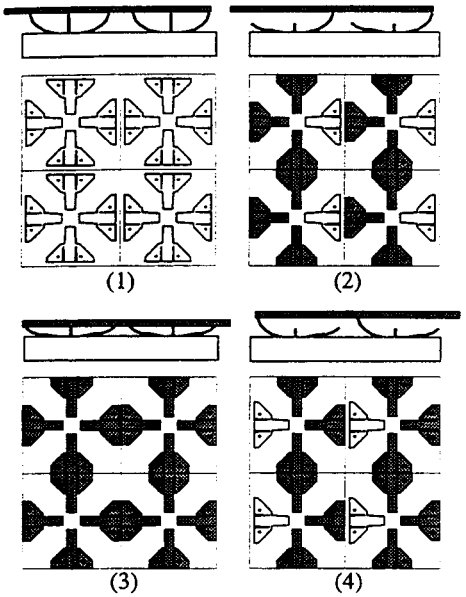


Figure 5. Side and plan views of four phase gait. Side views show object movement over actuators. In the plan view, dark colored actuators are thermally deflected downward.

Four-phase, linear gaits and step size. The actuators were operated at widely variable step rates. The effect this had on the linear movement of the 3 x 3 mm silicon IC was examined and the results are summarized in Fig. 6 below. When the time taken for all four phases (gait period) illustrated in Fig. 5 above is below 80 msec, the distance the object moves in one gait or distance per step varies nearly linear with gait period. Above ≈120 msec the distance the object moved per gait plateaued at an average of 9 μm per cycle. This behavior is related to the thermal time constant of the actuator. Longer gait periods allow the actuator to approach its steady state temperature. When the gait period is shorter, the actuator does not deflect as much and so the resulting distance per cycle is less.

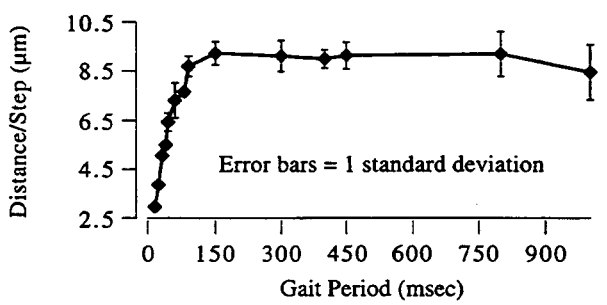
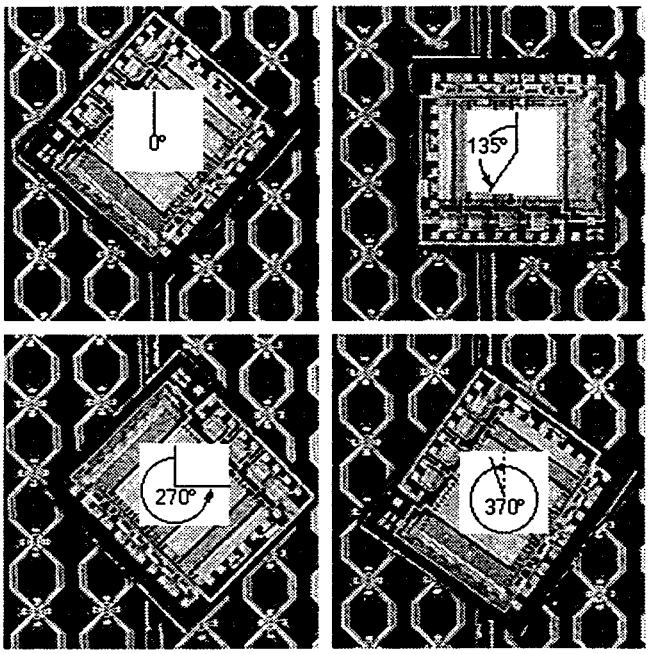
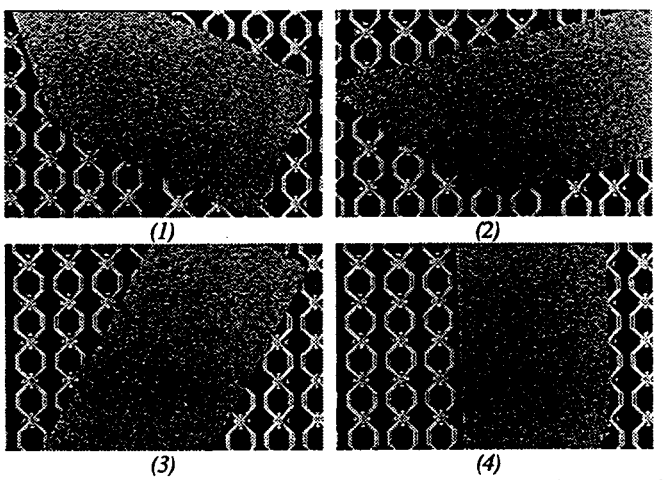


Figure 6. Plot of average distance a 3 x 3 mm silicon IC moved per gait as gait period is increased from 16 msec to 1 sec. The error bars for periods ≤ 40 msec are smaller than the dots representing their averages.

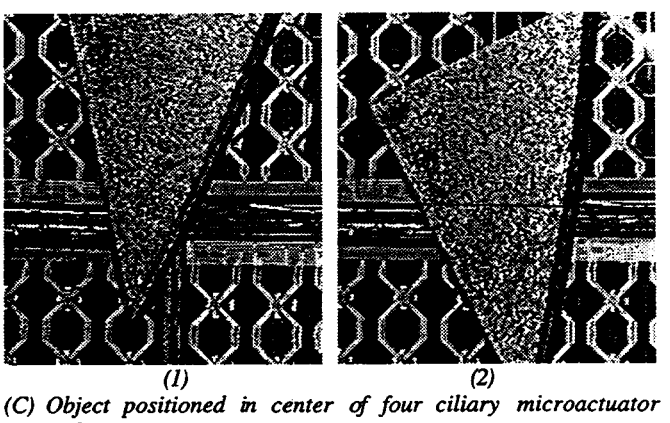
Rotation, alignment, and centering. A set of four ciliary microactuator arrays was also packaged and tested. Takeshima and Fujita [12] and Böhringer, et al.[1] have discussed the potential for a distributed micro-motion system to manipulate objects using sensorless methods. They describe how distributed micro motion systems can perform such tasks as object rotation, alignment to a given axis, and centering objects at the center of an array. With this four chip set, these manipulation tasks have been demonstrated and are shown in Fig. 7.



(A) Object rotation through 370°



(B) Object's longitudinal axis aligned to vertical axis of microactuator array.



(C) Object positioned in center of four ciliary microactuator array chips.

Figure 7. *Video frames of four omnidirectional ciliary microactuator arrays performing object rotation, alignment, and centering.*

Electrostatic Actuation. As thermal actuators are well known to have high power requirements, a low-power electrostatic hold-down mode becomes very attractive. The Al electrostatic plates are embedded symmetrically within the layers of nitride and polyimide of each actuator, and can be used to hold the actuator down on the substrate, or to hold an object onto a pair of actuators by the application of a voltage between the pair.

The capacitance of an actuator relative to the Si substrate involves a stack of several dielectric layers. When an actuator is flattened against the substrate (zero air gap), the layers include 0.15 μm of PECVD nitride (ε_r = 7.5), 4 μm of polyimide (ε_r = 3.4), 0.5 μm of LTO (ε_r = 7.5), and 0.5 μm of thermal oxide (ε_r = 7.5). The capacitance of this stack is

$$C = \frac{A}{\frac{d_1}{\varepsilon_1} + \frac{d_2}{\varepsilon_2} + \frac{d_3}{\varepsilon_3} + \frac{d_4}{\varepsilon_4}} \tag{5}$$

If fringing fields are neglected, the hold-down voltage is

$$V \approx \sqrt{\frac{2F_L\varepsilon_0}{C}\left(\frac{d_1}{\varepsilon_1} + \frac{d_2}{\varepsilon_2} + \frac{d_3}{\varepsilon_3} + \frac{d_4}{\varepsilon_4}\right)} \tag{6}$$

where F_L is the force required to flatten the actuator and ε_0 is the permittivity of the air gap. The above parameters predict a hold-down voltage of 27 V.

Figure 8 illustrates the deflection characteristics of the actuators to both thermal and electrostatic actuation. When the actuators are in their full upward position, their capacitance to the substrate is very low and 500 V of applied bias only produces a downward deflection of 15 μm. In comparison, 16.7 mW applied to the thermal elements produces a downward tip deflection of nearly 30 μm. Predeflecting the actuators downward with the thermal elements greatly reduces the electrostatic voltage required to pull down and hold the actuators. Pull-in voltages of 300 to 500 V were observed for 16 to 20 mW of initial thermal power per actuator. Once the actuators were snapped down, holding voltages of only 70 V were required, both with and without simultaneous thermal actuation. The observed 70 V hold-down voltage is larger than the predicted 27 V value, probably because the Si substrate provides a less than perfect ground plane.

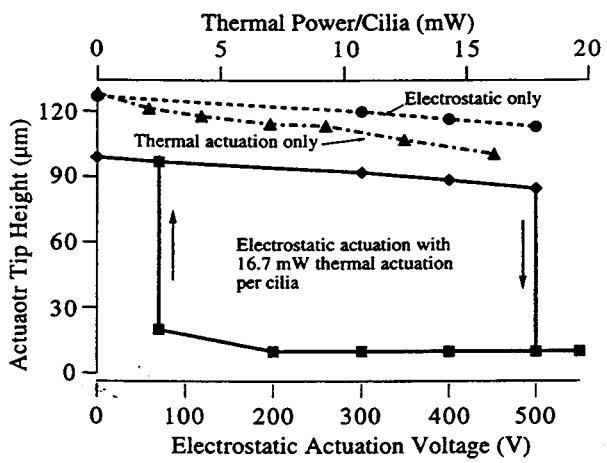


Figure 8. *Deflection characteristics for electrostatic only, for thermal only, and for combined thermal (at 16.7 mW per cilia) and electrostatic actuation.*

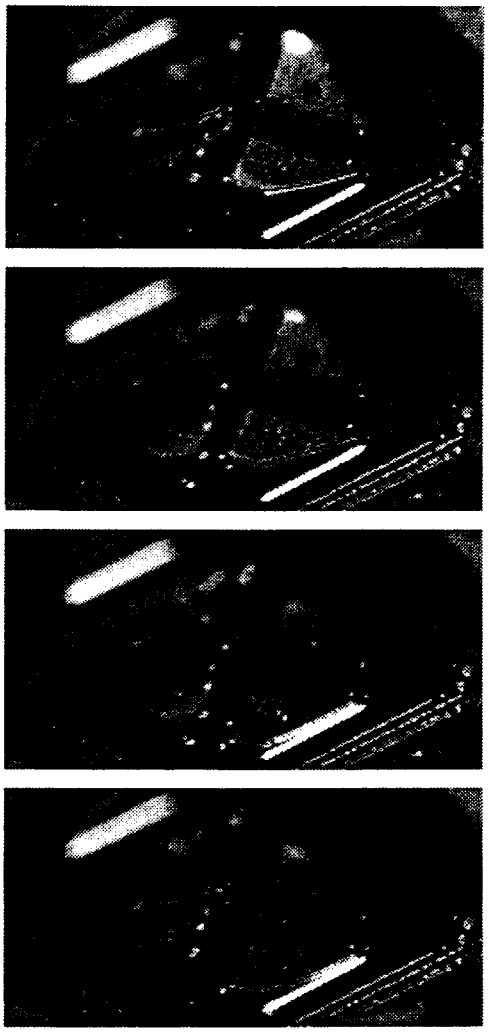


Figure 9. *Video frames of actuators during combined mode thermal and electrostatic actuation. A) Unpowered, initially curled actuators. B) 16.7 mW of applied power to both actuators. C) Right actuator with 500V applied to electrostatic plates. D) Left actuator unpowered, returns to same position as in (A).*

By applying a differential voltage between neighboring actuators, conductive objects can be drawn down onto the actuators. This causes the actuator to flatten out against the object, drawing both the actuator and the object closer to the substrate, as shown in Fig. 10.

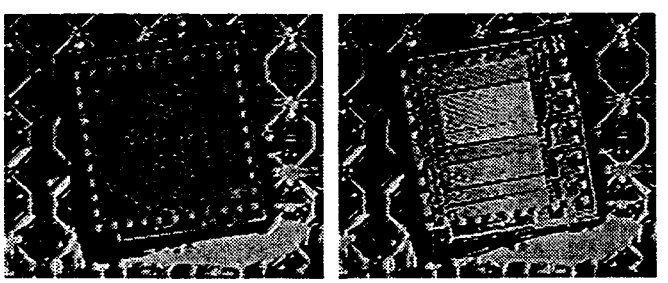


Figure 10. *Optical video frames showing electrostatic hold down of an IC die in deionized water. Tilting of the chip as it is drawn down on top of the actuators is seen as a change in light reflected from its surface.*

Capacitive Sensing of Objects. If the array were constructed to allow independent addressing of each actuator, electrostatic sensing could be used for object weighing and/or location. The feasibility of this concept was examined. A circuit sensitive to capacitive changes resulting from changes in actuator position was built. A simple, balanced race condition circuit was designed using 4000 series CMOS gates which produces glitch pulses whose width is proportional to the nodal capacitance. Low pass filtering of this pulse train produces a DC voltage proportional to the capacitance. The response of this circuit was recorded while manually loading and then unloading all the actuators. The results are summarized in Fig. 11. below. The results indicate that sensing object weight/location by capacitive sensing is possible.

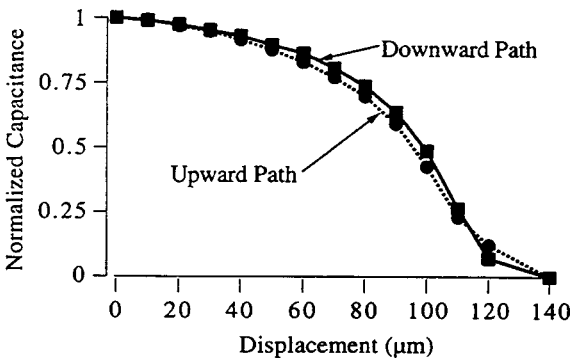


Figure 11. Vertical deflection of actuators (produced manually through micrometer) and change in voltage (proportional to capacitance). The actuator's tip is undeflected when displacement is set to zero.

CONCLUSIONS AND OUTLOOK

A ciliary microactuator array using independent thermal and electrostatic actuation has been described. The fabrication is based upon polyimide thermal bimorphs with aluminum sacrificial release layers and is compatible with prefabricated CMOS circuits. The combination of thermal and electrostatic actuation provides the simultaneous features of a high lifting capacity, large angle deflection mode with a low power hold-down mode. Capacitive sensing of the actuator position can be used for position feedback control and object sensing. The 8 x 8 array has demonstrated applicability for precision positioning, alignment, and inspection of IC dice with micron resolution.

Further system demonstrations and enhancements are planned. Design efforts are underway to lower electrostatic actuation voltages to practical levels ($\leq$ 100V) by redesign of the shape and position of the electrostatic plates as well as layer thicknesses. In addition, capacitive sensing for feedback control and object measurement is being investigated. A ciliary microactuator array with pre-fabricated CMOS circuits, which would enable control of individual cilia, is also planned.

ACKNOWLEDGMENTS

This research was supported by grant no. N00014-92-J-1940 from the Advanced Projects Research Agency (ARPA), managed by the Office of Naval Research. Special thanks are extended to Drs. Ken Gabriel and Teresa McMullen for supporting this effort. The authors give thanks to Karl Böhringer for his much appreciated contributions during testing. Thanks also go to Ken Honer and Dario Falquier.

REFERENCES

1. K.-F. Böhringer, B. R. Donald, R. Mihailovich, N. C. MacDonald, "A Theory of Manipulation and Control for Microfabricated Actuator Arrays," in *Proc. IEEE Micro Electro Mechanical Systems, MEMS '94*, pp. 102-107, 1994.
2. C. Liu, T. Tsao, Y.-C. Tai, W. Liu, P. Will, C.-M. Ho, "A Micromachined Permalloy Magnetic Actuator Array for Micro Robotics Assembly Systems," in *Transducers – Digest 8th Int. Conf. on Solid-State Sensors and Actuators/Eurosensors IX*, vol. 1, pp. 328-331, Stockholm, Sweden, June 1995.
3. N. Takeshima and H. Fujita, "Polyimide Bimorph Actuators for a Ciliary Motion System." in *ASME WAM, Symp. Micromech. Sensors, Actuators, and Systems*, DSC-vol. 32, pp. 203-209, 1991.
4. K. S. J. Pister, R. Fearing, and R. Howe, "A Planar Air Levitated Electrostatic Actuator System." in *Proc. IEEE Workshop on Micro Electro Mechanical Systems*, pp. 67-71, Napa Valley, CA, February 1990.
5. S. Konishi and H. Fujita, "A Conveyance System Using Air Flow Based on the Concept of Distributed Micro Motion Systems," in *Transducers–Digest 7th Int. Conf. on Solid-State Sensors and Actuators*, pp. 28-31, Pacifico, Yokohama, Japan, June 1993.
6. T. Furuhata, T. Hirano, and H. Fujita, "Array-Driven Ultrasonic Microactuators," in *Transducers – Dig. Int. Conf. on Solid-State Sensors and Actuators*, pp. 1056-1059, Montreux, France, June 1991.
7. G. Lin, C.-J. Kim, S. Konishi, and H. Fujita, "Design, Fabrication, and Testing of a C-Shape Actuator," in *Transducers – Digest 8th Int. Conf. on Solid-State Sensors and Actuators/Eurosensors IX*, , pp. 416-419, 1995.
8. P. M. Hall, F. L. Howland, Y. S. Kim, and L. H. Herring, "Strains in Aluminum-Adhesive-Ceramic Tri Layers," *ASME J. Electronic Packaging*, vol. 112, pp. 288-302, 1990.
9. P. M. Hall, F. L. Howland, Y. S. Kim, and L. H. Herring, "Errata", *ASME J. Electronic Packaging*, vol. 113, pp. 26, 1991.
10. W.-H. Chu, M. Mehregany, and R. L. Mullen, "Analysis of Tip Deflection and Force of a Bimetallic Cantilever Microactuator," *J. Micromech. Microeng.*, vol. 3, pp. 4-7, 1993.
11. J. Söderkvist, "Similarities Between Piezoelectric, Thermal and Other Internal Means of Exciting Vibrations," *J. Micromech. Microeng.*, vol. 3, pp. 24-31, 1993.
12. N. Takeshima and H. Fujita, "Design and Control of Systems with Microactuator Arrays," in *Proc. IEEE Workshop on Advanced Motion Control*, pp. 219-232, Yokohama, Japan, March 1993.
13. J. D. Foley and A. Van Dam, *Computer Graphics: Principles and Practice*, 2nd Edition, 1996.

SURFACE MICROMACHINED POLYSILICON THERMAL ACTUATOR ARRAYS AND APPLICATIONS

John H. Comtois and Victor M. Bright
Air Force Institute of Technology
Department of Electrical and Computer Engineering
Wright-Patterson AFB, OH 45433-7765

ABSTRACT

Thermal actuator arrays have been developed to complement the capabilities of electrostatic actuators. This paper demonstrates how arrays of two or more thermal actuators can be applied to a variety of basic building-block micromechanical devices such as stepper motors, optical component positioners, and grippers.

INTRODUCTION

Electrostatic actuators can operate at low power and high frequencies, but they require close tolerances or high voltages to achieve large forces, and typically have small deflections. Surface micromachined polysilicon thermal actuators are capable of large deflections and high forces in a current/voltage regime that is compatible with standard integrated circuit electronics. The devices presented in this paper were designed and fabricated in polysilicon using the ARPA-sponsored Multi-User MEMS Process, but could be re-created in any MEMS process that provides at least one releasable, current carrying layer.

ACTUATOR DESIGN

Thermal expansion is a strong force, but a single-material actuator based solely on the thermal expansion of a beam would have a small deflection. The actuators presented in this paper use a deflection multiplying structure based on a 'thermo-magnetic' actuator originally demonstrated in the LIGA process [1]. A typical 'U' shaped lateral thermal actuator or 'heatuator' is shown in Figure 1. Current is passed through the 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.

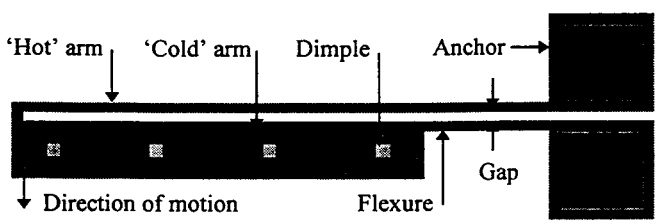


Figure 1. Basic layout of the lateral thermal actuator. Typical dimensions are: 'hot' arm 2.5 μm wide, 240 μm long; 'cold' arm 16 μm wide, 200 μm long; flexure 2.5 μm wide, 40 μm long, and gap 2 μm wide. The 4 μm square dimples prevent stiction. Deflections of 16 μm can be achieved at 3 V and 3.5 mA in 2 μm thick polysilicon.

The only requirement to achieve lateral motion with this actuator structure is to have one arm expand more than the other. So the actuator design can take on a wide range of geometries, including actuators with vertical motion or external heating. Close tolerances are not required, and this extremely simple design can be fabricated in any MEMS process that includes at least one releasable, current carrying layer such as SCREAM, *i*MEMS™, and many other proprietary technologies. The thermal actuators presented in this paper were fabricated in the ARPA-sponsored Multi-User MEMS Process (MUMPS) [2].

Structural features which increase the temperature difference between the hot and cold arms will increase the efficiency of the actuator. The simplest method is to make the cold arm wider or taller to decrease the current density or increase the surface area for cooling. Other investigators have shown that a trench etched under the hot arm of a metal actuator increased the thermal isolation from the substrate, thereby decreasing the heat lost from the hot arm and increasing the deflection versus current ratio [3].

Polysilicon has a higher resistivity than metal, which allows the actuator to operate in a voltage and current regime more closely matched to standard integrated circuit electronics. For example, a 220 μm long, 2 μm thick actuator with a 2.5 μm wide hot arm produced a force of 4.4 μN at an input power of 10.8 mW, delivered at 2.94 V, 3.68 mA, which is completely compatible with standard CMOS circuitry. These actuators can deflect up to 16 μm at the tip when unloaded, and the 4.4 μN force was delivered at an 8 μm deflection. Isolated actuators will operate at maximum deflection up to 300 Hz, and in one particular optical application, were observed to operate up to 1 kHz when driven to a 1.75 μm deflection [4]. Other details of actuator testing and design considerations have been presented elsewhere [5].

These actuators exhibit a back-bending characteristic when driven past the point where they stop deflecting in the forward direction. The mechanism is a deformation of the hot arm which shortens it. When the current is removed, the actuator is left permanently back-bent from its originally fabricated position. The amount of back-bending depends on the amount of over-driving. Large back-bending deflections are slightly less predictable than forward deflections, but small amounts of back-bending can be controlled [5]. This permanent back-bend of an unpowered actuator is useful for assembling and positioning MEMS devices.

In forward deflection against a stiff structure, a long, thin hot arm may bow rather than push; but in back-bending the hot arm is in tension, so it can deliver more force. For example, the 220 μm long actuator mentioned above can deliver 15.5 μN when back-bending. To back-bend that actuator, 7 V, 5.8 mA was applied for 5 seconds, but this is a one-time power requirement. After back-bending the actuator will operate as before, but now starting from a 'negative' deflection relative to its original position. This is a beneficial design option, as the actuator applications will show.

ACTUATOR ARRAYS

A longer polysilicon actuator is capable of greater deflection and force, but the trade-offs are that the electrical resistance

0-9640024-1-8 $3.00 © 1996 TRF

increases with the actuator's length, as does chance of the device failing due to stiction; and a longer hot arm is likely to deliver less force due to bowing. A more practical approach to obtaining higher forces is to connect several actuators with a common structure: a flexible yoke attached perpendicularly at the free ends of parallel thermal actuators, with flexures to accommodate their arc-like motion. Other options for arraying lateral thermal actuators are possible, but the yoke concept has proven to be the most compact, both with a single row of actuators and two rows of actuators set opposite to each other. An added benefit of arraying two or more actuators opposite to each other is that the yoke cancels out their arcing and expanding motions, producing purely linear motion, in addition to combining their forces. An array of ten actuators with a yoke can be seen on the right side of Figure 2.

With the yoke design, the force of the array can be increased by simply adding actuators to the same yoke. Since each actuator has its own attaching flexure, adding more actuators does not increase the stiffness of the yoke. Arrays of 60 actuators on a single yoke have been successfully operated, and more could easily be added. The actuators in the arrays shown below are wired in parallel. There was some initial concern over current sharing through so many parallel paths, but the resistance of the polysilicon increases with temperature, providing built-in current leveling. Any actuator which draws more current will heat more, and as its resistance increases its current draw will decrease.

An important consideration for using these arrays is that they produce a larger deflection when they have been given an initial back-bend. This allows the actuators to move through their zero position, where the least amount of force is lost in flexing the yoke and tethers. Provision must be made at the point where the array is attached to a driven mechanism if back-bending is to be used.

APPLICATIONS

The following applications are presented as examples of how these actuator arrays can be applied to a variety of basic building-block micromechanical devices such as stepper motors, optical component positioners, and grippers. The first application uses an actuator array to position and modulate a hinged mirror in a corner cube retroreflector (CCR). A CCR has three mutually perpendicular, mirrored walls. This mirror arrangement will reflect light back in the direction of its incoming path, and is commonly used in roadside reflectors. The device was proposed as a possible MEMS communication link [6], though no design specifics were given. A recently reported design uses hinged polysilicon plates for the mirrors [7]. Designs developed at the Air Force Institute of Technology (AFIT) took a similar hinged-mirror approach, but with electrostatically or thermally actuated mirrors.

The thermally actuated design, shown in Figure 2, has a static gold mirror on the substrate and two hinged, gold-plated mirrors. One hinged mirror is positioned and modulated with a thermal actuator array. The notched end of the array tether slides into a keyhole when the plate is lifted. The tether end can have several notches to set the initial angle of the plate. The other hinged mirror is held by a slotted locking plate designed to position the mirror within 3.5 mrad of perpendicularity with the substrate. The hinged mirror plates are made as thick as possible to counter stress in the gold layer, keeping the mirror flat. They are constructed of two releasable polysilicon layers with a layer of silicon dioxide trapped between them, for a total thickness of 4.25 μm [8].

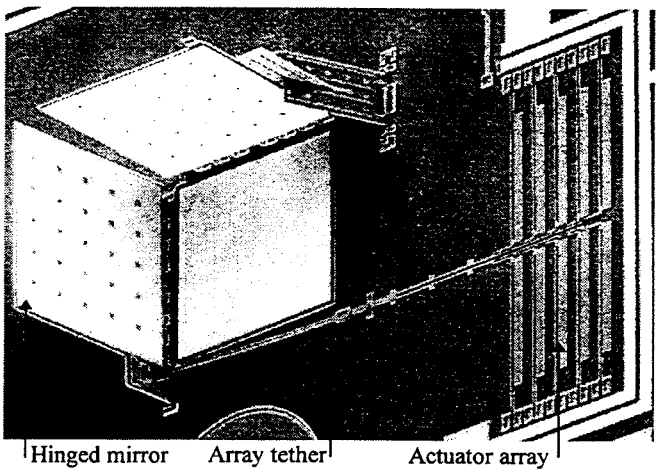


Figure 2. Corner Cube Retroreflector with an actuator array to position and modulate a hinged plate mirror. Mirrors are 280 μm square, actuators are 240 μm long, 20.5 μm wide, and 2 μm thick.

The next application is a gripper design, shown in Figure 3. The actuators are slid over the edge of the die after fabrication and overhang the die edge by 350 μm. Electrical connection to the actuator pair is made via 19 μm wide hinged-wires. These create gold-on-gold contacts with only 5.5 Ω of contact resistance each, which is small compared to the 645 Ω resistance of the actuators. The wires are 1.5 μm thick polysilicon with 0.5 μm thick gold on top. When they are flipped over, the gold on top is now on the bottom, where it shorts across gold contacts on the gripper frame and the substrate wires.

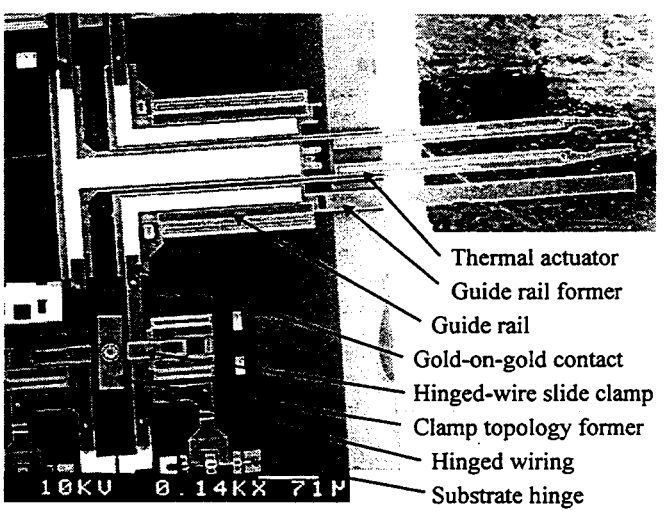


Figure 3. Grippers over edge of the die and wired with flip-over hinged-wires. Clamps hold wires in tension against gripper and substrate wiring. Topology former for the sliding clamps rotates out of the way. Gripper closes a full 16 μm at 2.7 V, 3.3 mA.

The wires are held in place by clamps whose topology is shaped by removable formers [8] to allow enough space under the clamp face to easily slide over the wire. The formers are rotated out of the way before the wires are flipped over. The wires are flexed to provide the spring force to make low resistance contacts, but they must be pressed down manually for the clamp to slide over them. The gripper frame also has formers to give the guide rails the proper topology so the frame can slide under them. These formers can be removed after the grippers are assembled; one is shown removed in Figure 3.

A method of increasing the deflection of the thermal actuator arrays is to build them into rotary or linear stepper motors. Figure 4 shows a rotary stepper motor with a 200 μm diameter rotor. A main array of ten 240 μm long thermal actuators is attached to a drive pawl, which is engaged and disengaged from the rotor by a second array of three pusher actuators set at 90° to the main array. This arrangement of actuators leaves much of the rotor edge free to connect to other structures such as gear trains or cam levers.

Fewer actuators can be used; motors with a single pusher and two drive actuators have been operated. The force delivered through the rotor has not been measured, but similar actuator arrays have been measured delivering 1.6 μN of force per actuator when forward bending, and 3.5 μN per actuator when back-bending against a static beam. Therefore, the amount of force delivered through the rotor will depend on the number of actuators in the main array, and whether the motor is being run clockwise or counterclockwise.

The actuators in this motor are operated from a slightly back-bent position, as shown in Figure 4. The back bending appears as a slight bow in the hot arms of the actuators which, due to the current leveling property of these actuators, all get the same amount of bend. For example, on this motor the pusher array was back-bent with a 5 sec application of 7.5 V, but operates at 3.7 V.

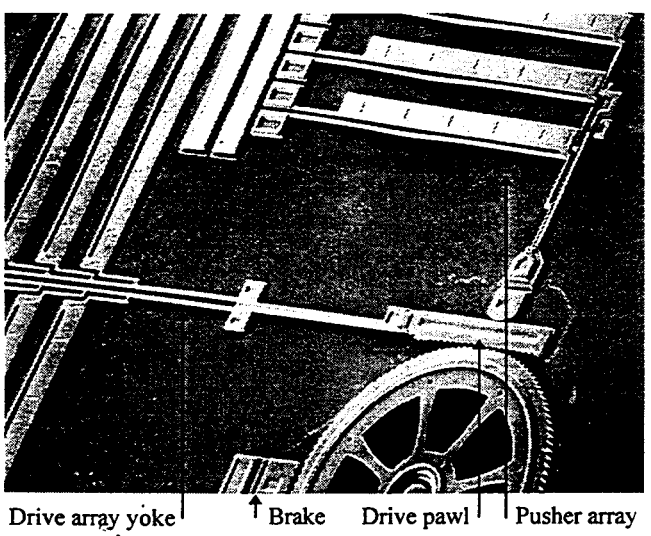


Figure 4. Rotary stepper motor with thermal actuator arrays back-bent into their operating position, with the drive pawl engaged. The three pusher actuators that engage the drive pawl, and the ten actuators that move the rotor are 240 μm long, 21 μm wide, and 2 μm thick. The rotor is 200 μm in diameter.

The pusher array is back-bent to place the drive pawl in contact with the rotor, which is fabricated in the disengaged position. The pawl is left in the engaged position as a brake when no power is applied. A second brake attached to a single actuator is also available for holding the rotor in position while the drive pawl is disengaged, in case there is a backwards force from the structure being driven by the motor. The drive pawl length allows variation in the amount of main drive back bend.

One important detail is the thickness of the teeth on the pawls and rotor. In the MUMPS process it is possible to create 3.5 μm thick structures by stacking the two releasable polysilicon layers [8]. The extra thickness prevents the pawls from sliding over the rotor when driven. The drive pawl can also be tethered directly to the pusher array with a thin flexure to prevent slide-over.

For testing, the actuator arrays were sequenced with a digital timing circuit which operates reed relays to apply power to the actuators. A typical timing sequence for driving the motor of Figure 4 counter-clockwise is shown in Figure 5. This circuit can drive the arrays up to 200 Hz; so with the 64 teeth on the rotor being engaged two at a time, the motor will rotate at 375 rpm. For positioning optical components, the device to be positioned can be placed directly on the rotor as shown in Figure 6, a device currently in fabrication, or the rotor can be geared to the device.

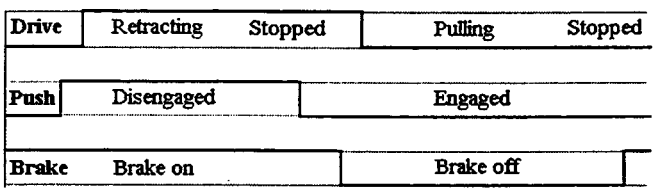


Figure 5. Typical timing sequence for driving the rotary stepper motor of Figure 4 in the counter-clockwise direction.

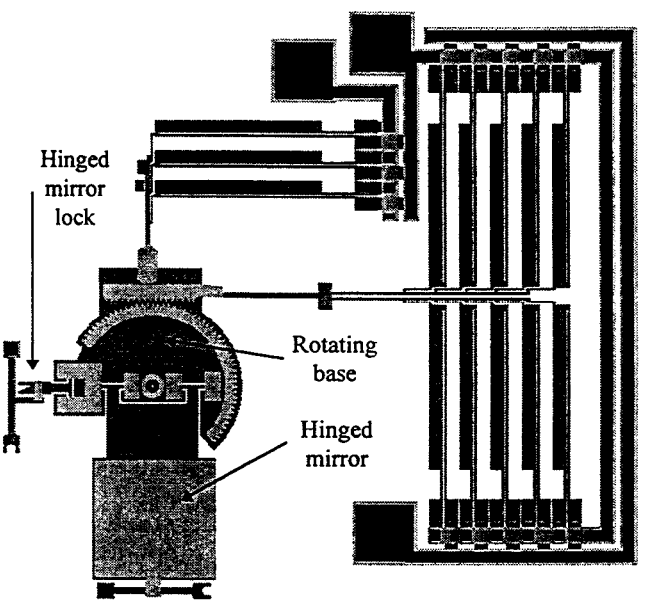


Figure 6. Rotary stepper motor built directly into a hinged, rotating mirror base. Motor could also be geared to the base to provide finer positioning increments. The base is 200 μm across.

The same two-array actuation scheme can be applied to linear motors by replacing the rotor with a sliding, toothed rack. Figure 7 shows a linear stepper motor driven by one-sided arrays. Another configuration which takes up less space uses one pair of arrays which push the rack against a stationary wall to engage it. In the test motor of Figure 7 the rack can be moved 200 μm, but for applications such as sliding grippers off the die, the rack could be as long as needed and incorporated in the device being positioned.

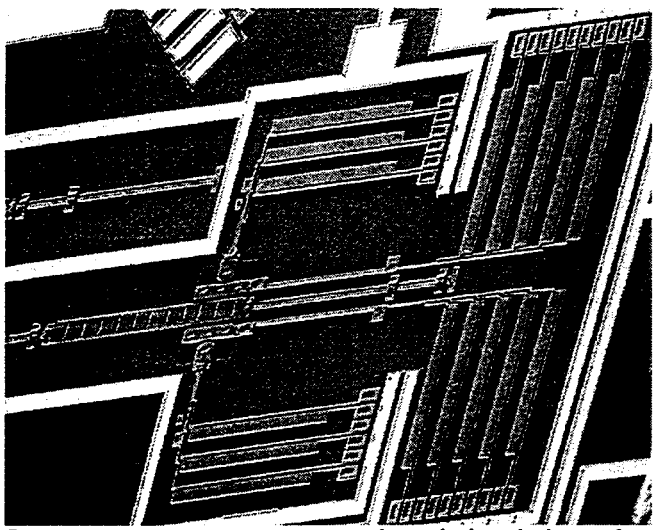


Figure 7. Linear stepper motor with toothed rack driven from both sides by arrays of 240 μm long, 2 μm thick actuators. The rack on this motor can slide 200 μm.

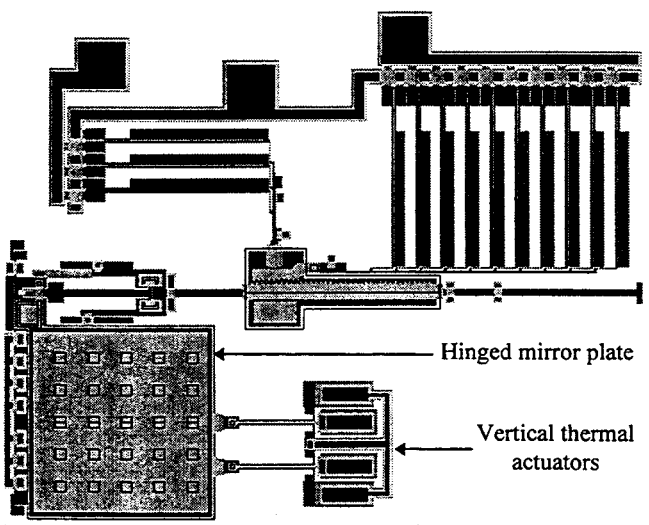


Figure 8. Linear stepper motor arranged to lift a hinged mirror plate for self-assembly of optical devices. Vertical thermal actuators provide the initial lift.

Another potentially useful application of stepper motors is the self-assembly of hinged structures. Such structures are typically assembled manually, a time-consuming and delicate process. Thermally actuated stepper motors have sufficient force to raise and position hinged plates. Figure 8 shows the layout of a motor designed to lift a hinged mirror. In this design, currently in fabrication, vertically acting thermal actuators provide an initial lift to prevent the linear stepper motor from pushing against the hinges without lifting the plate.

CONCLUSION

These applications demonstrate how thermal actuator arrays can expand the microelectromechanical actuator design space. Their high force, large deflections, and compatibility with standard integrated circuit voltages complement the characteristics of electrostatic actuators, and make them ideal choices where compact actuators with high forces or deflections are of greater concern than their power consumption and operating frequency. Their simple design and modest lithography and fabrication requirements mean they could be fabricated in any MEMS process that includes a releasable, current carrying layer.

ACKNOWLEDGEMENT

This work was sponsored by the USAF Phillips Laboratory PL/VTEE, Kirtland AFB, New Mexico.

REFERENCES

1. H. Guckel, J. Klein, T. Christenson, K. Skrobis, M. Laudon, and E. Lovell, "Thermo-magnetic Metal Flexure Actuators," *Technical Digest, Solid State Sensors and Actuators Workshop,* Hilton Head, SC, 13-16 June 1992, pp. 73-75.
2. D. Koester, R. Mahedevan, and K. Marcus, *Multi-User MEMS Processes (MUMPS) Introduction and Design Rules*, rev. 3, Oct 1994, MCNC MEMS Technology Applications Center, 3021 Cornwallis Road, Research Triangle Park, NC, 27709.
3. L. Field, D. Buriesci, P. Robrish, and R. Ruby, "Micro-machined 1x2 Optical Fiber Switch," *Proc. Transducers '95/Eurosensors IX*, Stockholm, Sweden, 1995, pp. 344-347.
4. D. Sene, J. Grantham, V. Bright, and J. Comtois, "Development and Characterization of Micro-mechanical Gratings for Optical Modulation," *Proc. IEEE MEMS-96 Workshop,* San Diego, CA, 11-15 Feb. 1996, pp. 222-227.
5. J. Comtois, V. Bright, and M. Phipps, "Thermal Micro-actuators for Surface-micromachining Processes," *Proc. SPIE: Micromachined Devices and Components*, 23-24 October 1995, vol. 2642, pp. 10-21.
6. K. Brendley, and R. Steeb, "Military Applications of Microelectromechanical Systems," RAND report to the Office of the Secretary of Defense, United States Air Force, United States Army, RAND, Santa Monica, CA, 1993.
7. D. Gunawan, L. Lin, and K. Pister, "Micromachined Corner Cube Reflectors as a Communication Link," *Sensors and Actuators A*, 46-47, pp. 580-583, 1995.
8. J. Comtois and V. Bright, "Design Techniques for Surface-Micromachining MEMS Processes," *Proc. SPIE: Micro-machining and Microfabrication Process Technology*, vol. 2639, 23-24 Oct. 1995, pp. 211-222.

MONOLITHIC GEARED-MECHANISMS DRIVEN BY A POLYSILICON SURFACE-MICROMACHINED ON-CHIP ELECTROSTATIC MICROENGINE

Jeffry J. Sniegowski, Sam L. Miller, Glenn F. LaVigne, M. Steven Rodgers and Paul J. McWhorter

Integrated Micromechanics, Microsensors, & CMOS Technology Department
Sandia National Laboratories
Albuquerque, NM 87185-1080

ABSTRACT

We have previously described a practical micromachined power source: the polysilicon, surface-micromachined, electrostatically-actuated microengine [1]. Here we report on three aspects of implementing the microengine. First, we discuss demonstrations of the first-generation microengine actuating geared micromechanisms including gear trains with elements having dimensions comparable to the drive gear (approximately 50 μm) and a relatively large (1600-μm-diameter) rotating optical shutter element. These configurations span expected operating extremes for the microengine, and address the coupling and loading issues for very-low-aspect-ratio micromechanisms which are common to the design of surface-micromachined devices. Second, we report on a second-generation of designs that utilize improved gear teeth design, a gear speed-reduction unit, and higher force-per-unit-area electrostatic comb drives. The speed-reduction unit produces an overall angular speed reduction of 9.63 and requires dual-level compound gears. Third, we discuss a dynamics model developed to accomplish three objectives: 1) drive inertial loads in a controlled fashion, i.e. specify and achieve a desired time-dependent angular position of the drive gear, 2) minimize stress and frictional forces during operation, and 3) as a function of time, experimentally determine forces associated with the drive gear, such as load torque being applied to the drive gear due to friction.

INTRODUCTION

Preliminary demonstrations of on-chip powered geared mechanisms have been recently reported [2,3]. In the case of on-chip electrostatic actuation of geared mechanisms, the on-chip power source must overcome the forces of stiction [4]. The forces generated by on-chip means are limited and relatively small compared to externally applied forces. The result is that initial start-up can be problematic and unreliable. In contrast, start-up and operation issues related to release and in-use stiction are more readily overcome in mechanical systems that are actuated on by external means [5]. We report here the full function operation of various geared mechanisms driven by on-chip means and discuss system dynamics which lead to tailored input signals which enhance start-up and operation. Details of the fabrication of these devices are reported elsewhere [1,6].

For our initial proof-of-concept demonstration, we chose two principle configurations of the micro-engine and geared micromechanism system. The first, exhibited in Figure 1, consists of the microengine loaded by comparably-sized gears and provides fundamental information on the basic design of the microengine and gears. The second involves the actuation of a relatively large (1600-μm-diameter) rotating optical shutter element shown in Figures 2 and 3. Design details for these two configurations are described in a previous report [7]. These configurations span expected operating extremes for the microengine and address the coupling and loading issues for very-low-aspect-ratio micromechanisms which are common to the design of surface-micromachined devices. We then report on a second-generation of designs that utilize improved gear teeth design, a gear speed-reduction unit, and higher force-per-unit-area electrostatic comb drives. Finally, optimization of endurance and of high-speed start-up of large elements through dynamical modeling is discussed.

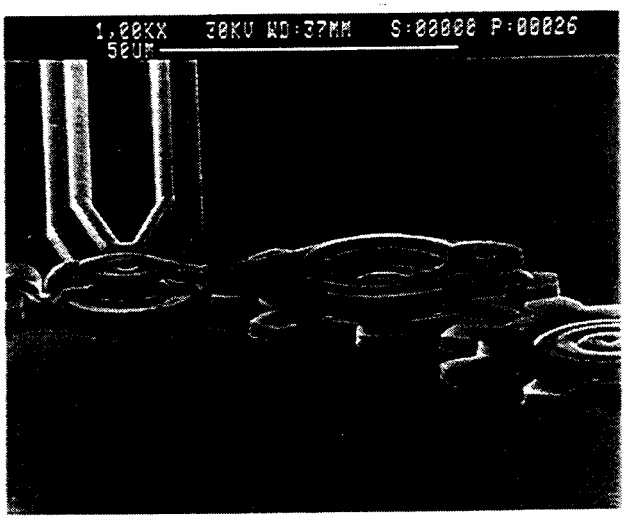


Figure 1. SEM perspective view of the microengine drive gear and comparably sized driven gear. Gear extreme diameter is approximately 60 μm and gear thickness is 2.45 μm.

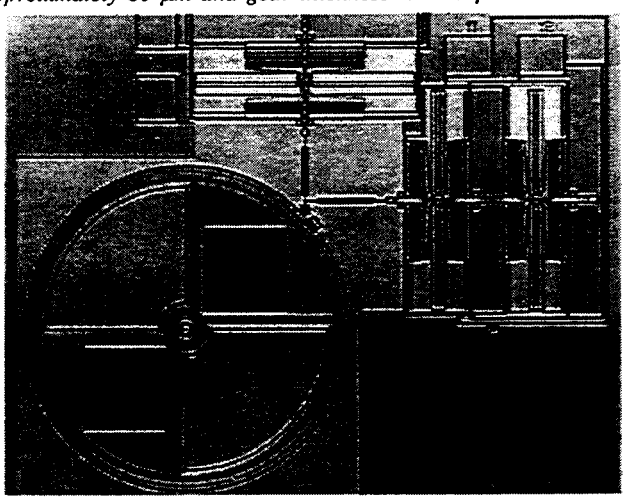


Figure 2. SEM micrograph overview of the microengine coupled to drive a 1600-μm-diameter optical shutter gear. The two linear drives are seen as the two orthogonal electrostatic comb drives in the upper right-hand corner of the micrograph. They are connected by linkages, seen close-up in Figure 3, to the drive gear which is coupled to the driven element by gear teeth.

As a quick review, the principle of operation for the microengine is well described by analogy to two orthogonal pistons connected to a crank shaft, i.e. two linear drive elements connected to a rotary element to produce rotational output. Referring to Figure 2, the linear drive elements, which are electrostatic comb drives [8], are connected by linkages to the rotary drive gear. This drive gear is then coupled directly to

0-9640024-1-8 $3.00 © 1996 TRF

the large gear through the gear teeth. The close-up in Figure 3 more clearly illustrates the linkages from the linear drives connected to the drive gear.

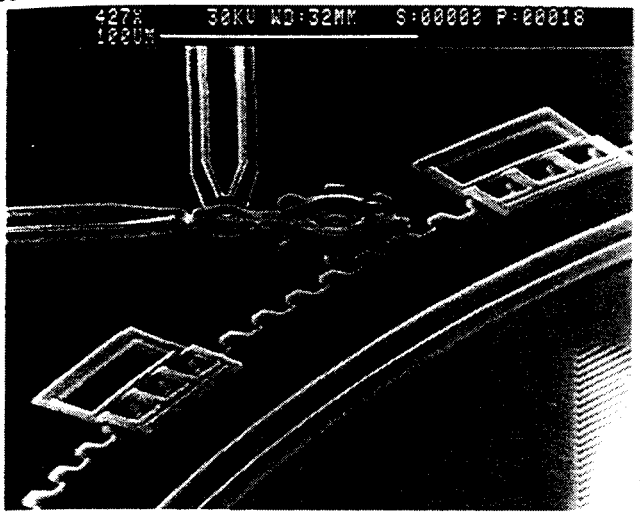


Figure 3. SEM micrograph of the microengine drive gear coupled to a 1600-μm-diameter optical shutter gear. This gear element is 30 times the diameter of the microengine drive gear. Also present in this photo are a set of retainers used to maintain vertical alignment of the sets of gear teeth, which are only 2.45 μm thick.

DEVICE DEMONSTRATIONS

First-Generation Devices - The two devices discussed in this section consist of the microengine driving gears of comparable dimension (Figures 1 and 4), and the microengine driving a 1600-μm-diameter gear which functions as an optical shutter (Figures 2 and 3).

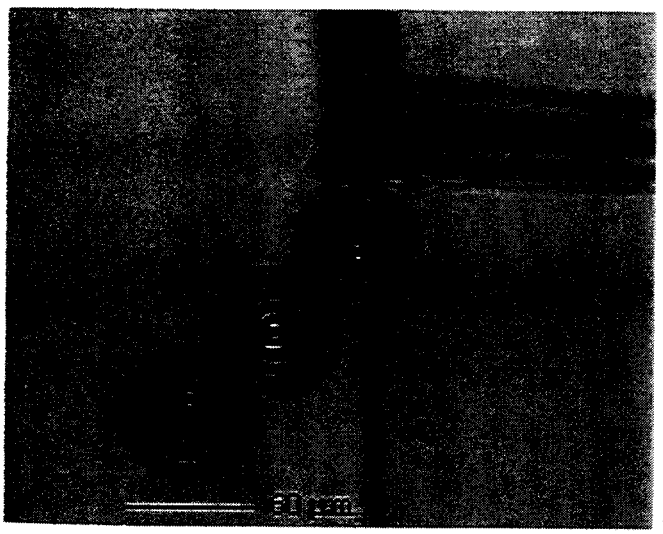


Figure 4. Optical micrograph taken during the operation of the microengine driving two additional gears at 200,000 rpm. Gear extreme diameter is 60 μm.

The rotation of the drive gear and gear train was achieved by the application of 90 V square-wave inputs to the comb-drive actuators. These tests were performed in ambient air and full rotation of single and multiple gear configurations was demonstrated. During tests conducted using low-frequency square waves, the motion observed is similar to that of a stepper motor. With each step in the applied signals, the drive gear moves to an equilibrium position and remains in that position until the next change of the applied signals. Due to the low inertia of these structures, the movement to the equilibrium positions occurs on a sub-millisecond time scale. Changing the phase relationship between X and Y actuators allows for reversal of the microengine rotation. Several devices were operated at square-wave frequencies ranging from 0.5 Hz to 3333 Hz which corresponds to 30 revolutions per minute (rpm) to 200,000 rpm. Figure 4 is an optical micrograph of the drive gear and gear train operating at 200,000 rpm.

The dynamical modeling accounts for the inertial load of the drive gears. In the case of the large optical shutter gear, a simple step ramp of the drive signals to the engine produces a maximum speed of 600 rpm for the 1600 μm-diameter-gear, while a modified ramp of the drive signals as predicted by the modeling has allowed a speed of 4800 rpm to be obtained. These large gear speeds correspond to ~20,000 rpm and 146,000 rpm for the drive gear, respectively. Additional discussion of this topic is covered in the section on dynamical operation model.

Second-Generation Designs - By driving a 1600-μm-diameter gear, the first-generation microengines have proven that electrostatically driven assemblies can produce significantly useful work. Further improvements, however, will be required if similar devices are to propel the more complicated micromachines envisioned. We are addressing these needs by developing a new family of microengines that utilize improved gear teeth design [9], elaborate gear-speed-reduction units, and higher force-per-unit-area comb drives. Together these enhancements can significantly increase the torque available for driving mechanical loads.

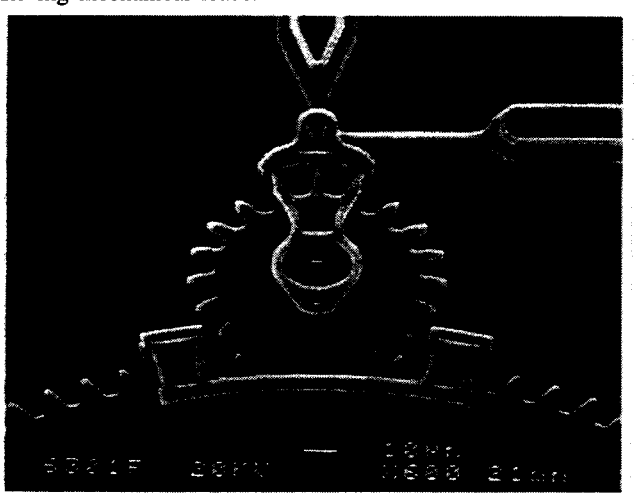


Figure 5. SEM of second-generation tooth design. The drive gear pitch circle radius is increased from 26 μm to 38 μm to accommodate the deeper tooth design and additional teeth. Note the deeper and multiple tooth engagement relative to the gears in Figure 1.

Although the first-generation gear tooth shape is involute [7], (refer to Figures 1 and 3) several shortcomings were readily observed. Process constraints on the ratio of the hub-to-gear spacing to the overall gear dimension allowed relative motion of the gears to significantly change the tooth pressure angle and rolling behavior. To reduce this, a primary modification to the design included a larger drive gear diameter to accommodate deeper tooth design and additional teeth for greater engagement and continuous engagement of multiple teeth. The depth was changed from approximately 7 μm to 9.5 μm. A secondary change modified the manner in which the link runs over the gear edge such that the constraint on the tooth tip width was eliminated from the first design. This allowed fabrication of a more ideal tooth shape. Comparison of Figure 5 to Figure 1

illustrates these changes. These changes eased start-up and produced much smoother operation, particularly at low speed.

A second structure preserved and only slightly modified from the early design is the retainer for out-of-plane gear-tooth alignment. This is the over-clamp at the edge of the large gear in both Figures 3 and 5. A simple evaluation by either including or not including the retainers in the operation of the devices demonstrated that they are highly effective at preventing out-of-plane disengagement of the gears while not appreciably adding to the friction load.

New Second-Generation Devices - The aforementioned design changes led us to consider a gear-speed-reduction unit, i.e. a transmission. To construct the unit required design of dual-level compound gears. This was accomplished by full utilization of the three levels of mechanical polysilicon in the basic process. A dual-level gear-speed-reduction unit that drives a rack back and forth is shown in figure 6.

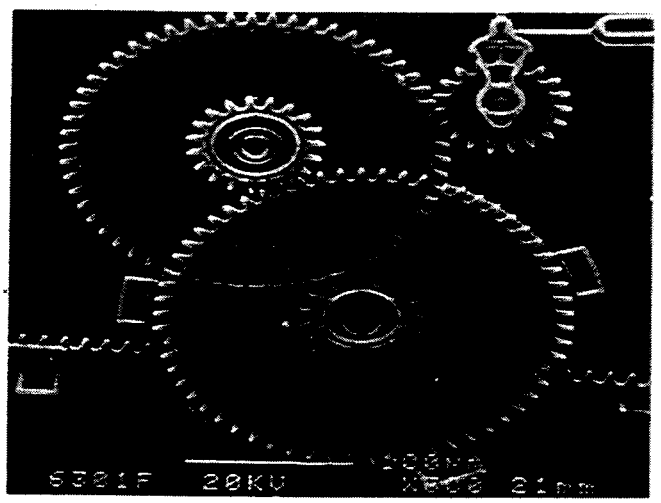


Figure 6 *is an SEM of the completed gear-speed-reduction unit and linear rack. The linear speed of the rack is approximately one tenth that of the linear tooth velocity of the drive gear. The rack can be used to drive a folding mirror, for example.*

This assembly utilizes a microengine with a 19 tooth pinion [**10**] to drive two compound gears that have teeth fabricated in level 2 and level 3 polysilicon. The hubs for all the gears are formed from level 1 and level 2 polysilicon. Each compound gear has a center gear that is about 1/3 the diameter of its outer unit. If we take the gears in order of transmission starting with the drive (pinion) gear, the tooth numbers are 19, 57, 19, 61, and 17 respectively. Thus, the linear speed of the rack driven by the speed-reduction unit is 1/10.77 of what it would be if driven directly by the pinion gear. The lower (level 2 poly) components can be seen in Figure 7. The large flat areas under the final gear are to provide a planar surface for fabrication of the large level 3 gear. The appearance of gear teeth openings in the large top gear in Figure 6 are artifacts that compensate for possible mechanical interferences caused by the conformal nature of the film depositions in the non-planarized process. Such design constraints have been discussed in earlier articles on the microengine [**1,6**]. The use of Chemical-Mechanical Polishing for planarization as presented in [**11**] removes such design constraints.

Bi-directional drive of the rack by the reduction unit has been demonstrated. In addition, removal of the rack from the structure allows continuous high-speed operation. More significant mechanical loads will be coupled to similar microengine driven gear-speed-reduction assemblies in the near future.

Further utilization of all three levels of mechanical polysilicon was made by including a bi-level set of fingers and bi-level support springs on the electrostatic comb drives used to power the rotary drive gear (see Figure 8). The net result is analogous to overlaying two comb drives which effectively increases the net output of the linear drives. However, using bi-level springs separated by a two micrometer thick sacrificial oxide essentially triples the effective out-of-plane thickness of the springs. Since the z-direction stiffness is proportional to the thickness cubed, the threefold increase corresponds to approximately a twenty-fold increase in out-of-plane stiffness.

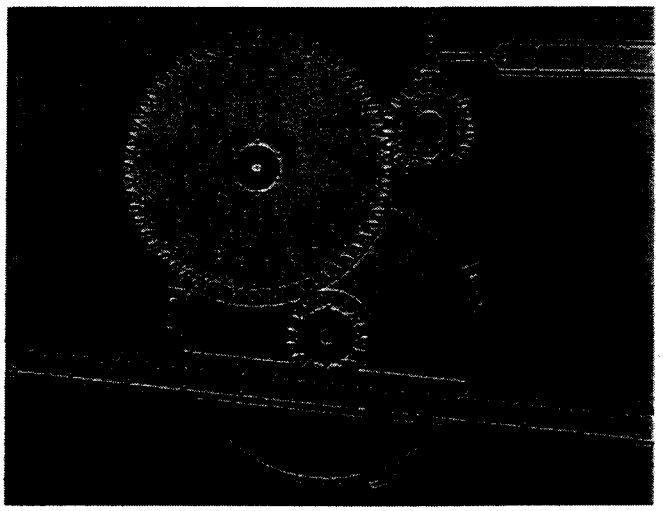


Figure 7. *Shown in this micrograph are the three lower level gears and rack in the gear-speed-reduction unit.*

This produces a more robust structure that is less susceptible to stiction problems associated with the springs as well as the shuttle. In addition, a more aggressive layout and design has reduced this comb drive's footprint to 2/3 that of its predecessor.

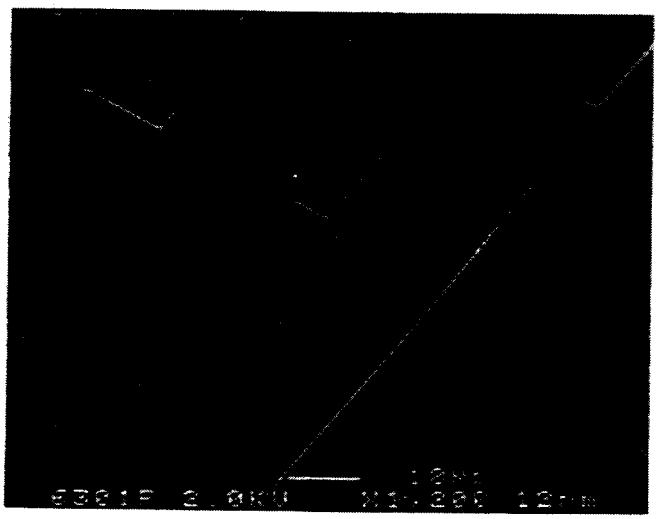


Figure 8. *Utilization of multiple levels of mechanical polysilicon is seen in this SEM image of bi-level comb drive fingers and support springs. Each layer is approximately 2 μm in thickness. In this case, the bi-level nature of the structures arises from the sacrificial oxide layer that was present between polysilicon layers during fabrication.*

Further Design Issues - The dominant failure mechanism of our second-generation drive is off-axis instability of the electrostatic comb drives leading to a lateral buckling failure under high-speed/high-load operation. The instability has its origin in the basic configuration of a linear comb drive. The linear force generated by the comb structure is derived from the fringing fields at the ends of the comb fingers [1]. These fingers, however, form interdigitated plates of parallel capacitors. Each of these fingers, in its nominal position, sits on an electrostatic potential hill along the cross-axis direction. Therefore, the fingers prefer to collapse into one or the other finger to each side. In the comb drives, supporting springs and/or mechanical stops normally prevent this lateral collapse up to a point. In the case of the current design, the onset of the lateral buckling occurs for an applied voltage of approximately 90V under loaded conditions. If there are no connected elements which generate off-axis loads, the onset has not occurred with our latest design at 145 V which is the current limit of our test set-up.

In order to drive the microengine to higher speeds and larger loads, we have had reasonable success shrinking the spacing to the mechanical stops. In particular, putting a taper on the central shaft of the comb drive in the area of the mechanical stops produces a gap of only a few tenths of a micrometer when the comb is energized. Referring to Figure 9, the gaps on each side of the central shaft are diminished as the central shaft moves between the two stops. Also, we have found that adding any additional lateral stiffness to the entire comb drive is worth the effort. This includes adequate stiffness to the comb fingers. An early change was to increase the lateral width of the fingers from 2 μm to 3 μm with 2 μm-finger gaps increased to 3 μm. Although we have seen improvements in controlling the lateral buckling, we are continuing to modify both the mechanical stops and the lateral stiffness.

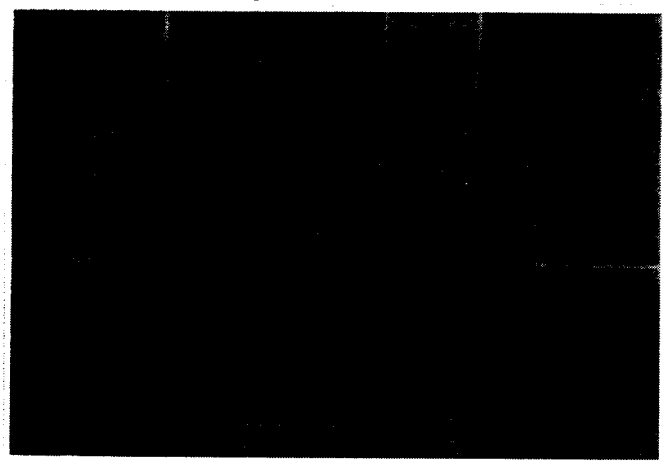


Figure 9. SEM close-up of the central shaft of a linear comb drive. The shaft is the vertical element photo while the stops are seen at both sides. As the shaft moves from this position, it is seen that the spacing to the stops is decreased by the tapers in the shaft.

DYNAMICAL OPERATION MODEL

The dynamical operation of the microengine is governed by the complex relationship between forces, masses, and the physically constrained geometry associated with its moving elements. Relevant forces include the electrostatic forces originating from the applied comb-drive voltages, restoring forces due to the comb support springs, viscous-damping forces, frictional forces between rubbing surfaces, and inertial forces associated with the linear motion of the comb drives and rotational motion of the gears. Though square-wave or sine-wave drive signals can be used to demonstrate functionality and achieve limited operation of the engine, they result in excessive and fluctuating frictional and inertial forces. Such forces can lead to premature failure, and do not result in well-controlled motion of loads being driven by the engine. Specifically, constant angular speed cannot be achieved without the application of specialized drive signals. In addition, rapid acceleration of inertial loads with minimum stress is not possible without properly engineered drive signals.

To enable controlled operation while minimizing parasitic forces, a dynamical model of the engine/load system has been developed and successfully applied. The results are achieved by analytically solving Newton's equations of motion. The model relates the applied voltages, physical device parameters (e.g. force constants, masses and geometry), frictional forces and load torques, and the time-dependent angular position of the drive gear. Application of the model permits the direct experimental determination of the electrostatic, damping, , and spring force constants from functioning engines. The resulting value of Young's modulus is $E \sim 170 \times 10^9$ N/m^2. The measured electrostatic force constant is typically within ten percent of that theoretically expected for the present geometry comb drives [1].

Once the force constants are experimentally determined, the dynamical model is used to experimentally measure the load forces acting on the drive gear. This is possible by measuring the time-dependent angular position of the drive gear, and knowing the applied drive voltages. From a series of different loading conditions, the coefficient of friction between the gear and hub is experimentally measured to be approximately 0.5.

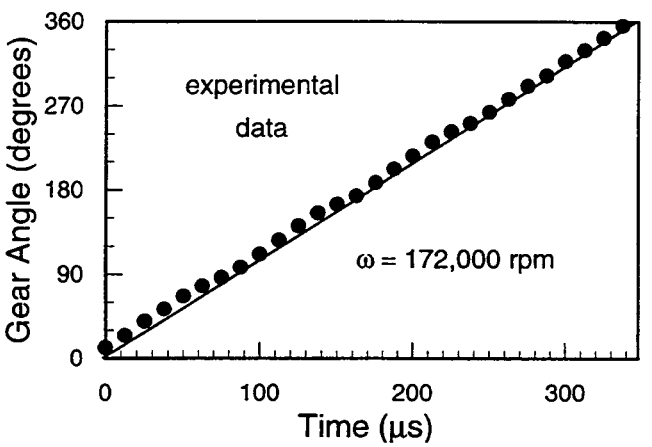


Figure 10. Operation using theoretically computed drive signals results in uniform rotation at speeds in excess of the resonant frequency (104,000 rpm) of the comb drives. Measurements are made using a time-delayed strobe technique.

Controlled operation of the engine can be achieved using the model. To do this, the voltages to be applied to the comb drives are directly computed for a desired angular speed. The experimentally measured angular position of the drive gear as a function of time is shown in Figure 10 for an engine running at a speed of 172,000 revolutions per minute (rpm) operated with drive signals created by the dynamical model. At this speed (nearly twice the resonant frequency of the comb drives), the linear inertial forces of the comb drives significantly impact the dynamics of the engine, and are properly accounted for by the model as evidenced by the approximately constant measured angular speed. The maximum speed of an engine thus far achieved is 300,000 rpm. The data shown in Figure 10, and the speed record, were obtained with no load attached to the drive

gear. At a constant speed, i.e. in the absence of angular acceleration, the rotational inertia of load gears does not impact the engine dynamics.

Rotational inertial effects can be significant during rapid starting and stopping of large gear loads. For example, the 1600-μm-diameter optical shutter shown in Figure 2 can only be operated up to a maximum speed of ~20,000 rpm (speed of 50-μm-diameter drive gear) using abruptly starting square-wave drive signals. The speed limitation associated with abrupt starting is caused by the fact that the transfer of electrostatic energy to rotational kinetic energy requires a finite amount of time. In contrast, one can drive the shutter at speeds up to 146,000 rpm by using engineered drive signals created with the dynamical model. Quickly accelerating a load to this speed requires drive signals of complex form that account for both linear and rotational inertial effects. To illustrate, Figure 11 shows drive signals designed to ramp to an engine speed of 140,000 rpm in two revolutions of the drive gear by linearly increasing the instantaneous speed, while at the same time driving the large load gear. For clarity, only the voltages on two of the four comb drives are shown.

Rapid starting and stopping of the engine can result in excessive parasitic forces when improper drive signals are used, even in the absence of a load with rotational inertia. The reason is that the motion of the engine components are physically constrained by the drive gear and linkages. As a result, the hub on which the drive gear rotates provides forces to the accelerating comb drives to compensate for the non-ideal force history provided by the electrostatic comb drives. Such forces can impact the lifetime of the engine. The endurance record achieved thus far for a single engine is $3.2(10^9)$ revolutions while operating at a speed of 146,000 rpm. The engine also accrued 66,300 start/stop cycles during the $3.2(10^9)$ revolutions.

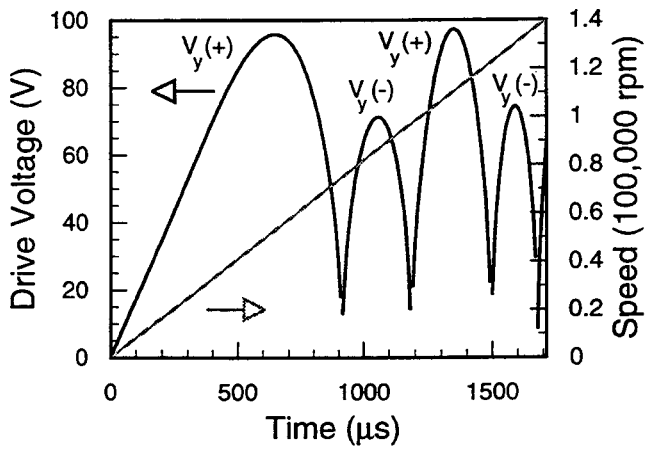


Figure 11. Drive signals to accelerate the 50-μm-diameter drive gear to a speed of 140,000 rpm in two revolutions by linearly increasing the speed, while at the same time driving a 1600-μm-diameter load gear.

CONCLUSION

In conclusion, this surface-micromachined actuation system can overcome the common start-up and in-use operation difficulties associated with stiction. We are now considering using the microengine as a driver in applications which have requirements from reliable start-up after long dormancy to continuous high-speed operation. The scope of the potential applications have broadened considerably after successful demonstration of improved gear tooth design, a gear-speed-reduction unit, and higher force-density comb drives.

ACKNOWLEDGMENT

We would like to acknowledge the entire staff of the Microelectronics Development Laboratory at Sandia National Laboratories for their efforts in the fabrication of these devices. We also thank Michael Callahan at Sandia National Laboratories for his excellent technical direction regarding the application of this technology. This work, performed at Sandia National Laboratories, was supported by the U. S. Department of Energy under contract DE-AC04-94AL85000.

REFERENCES

1. J. J. Sniegowski and E. J. Garcia, "Microfabricated Actuators and Their Application to Optics", *Proc. SPIE Miniaturized Systems with Micro-Optics and Micromechanics, 2383*, 2/7-9/95, San Jose, CA, (1995) pp. 46-64.

2. J. J. Sniegowski and E. J. Garcia, "Surface Micromachined Gear Trains Driven by an On-Chip Electrostatic Microengine", accepted for publication in IEEE Electron Device Letters.

3. R. Legtenberg, E. Berenschot, M.Elwenspoek and J. Fluitman, "Electrostatic Microactuators with Integrated Gear Linkages for Mechanical Power Transmission", *Proc. IEEE Micro Electro Mechanical Systems (MEMS),* 2/11-15/96, San Diego, CA, (1996) pp. 204-209.

4. R. L. Alley, G. J. Cuan, R. T. Howe, and K. Komvopoulos, "The Effect of Release-Etch Processing on Surface Microstructure Stiction", *Technical Digest of the 1992 IEEE Solid-State Sensor and Actuator Workshop,* Hilton Head Isl., SC, 6/22-25/92, (1996), pp. 202-207.

5. L. A. Field, R. M. White, and A. P. Pisano, "Fluid Powered Rotary Gears and Micro-Flow Channels", *Proc. of the 6th International Conf. on Solid-State Sensors and Actuators (Transducers '91),* 6/24-27/91, San Francisco, CA (1991) pp. 1033-1036.

6. E. J. Garcia and J. J. Sniegowski, "Surface Micromachined Microengine", *Sensors and Actuators A, 48*, 203 (1995).

7. E. J. Garcia and J. J. Sniegowski, "Surface Micromachined Microengine as the Driver for Micromechanical Gears", *Proc. of the 8th International Conf. on Solid-State Sensors and Actuators and Eurosensors IX,* 6/25-29/95, Stockholm, Sweden 1 (1995) pp. 365-368.

8. W. C. Tang, T-C. H. Nguyen, and R. T. Howe, "Laterally Driven Polysilicon Resonant Microstructures," *Sensors and Actuators, 20*, 25 (1989).

9. H. E. Merritt, *Gear Engineering*, Pitman Publishing, 1971, Chps. 4 and 6.

10. D. W. Dudley, *Handbook of Practical Gear Design,* McGraw-Hill Book Company, 1984, pp. 3.6.

11. R. D. Nasby, J. J. Sniegowski, J. H. Smith, S. Montague, C. C. Barron, W. P. Eaton, P. J. McWhorter, D. L. Hetherington, C. A. Apblett, and J. G. Fleming, "Application of Chemical-Mechanical Polishing to Planarization of Surface Micromachined Devices", *Technical Digest of the 1996 Solid-State Sensor and Actuator Workshop,* Hilton Head Isl., SC, 6/2-6/96, Transducer Research Foundation (1996), pp. these Proceedings.

Electromagnetic MEMS Scanning Mirrors For Holographic Data Storage

Raanan A. Miller, Geoffrey W. Burr, Yu-Chong Tai, and Demetri Psaltis
Electrical Engineering, MS 136-93, California Institute of Technology,
Pasadena, CA 91125

Chih-Ming Ho
MANE, UCLA
Los Angeles, CA 90024

Romney. R. Katti
Jet Propulsion laboratory
Pasadena, CA 91109

ABSTRACT

A mm-sized analog scanning MEMS mirror capable of deflection angles exceeding 60 degrees has been designed and fabricated. Development of this mirror is driven by the need for simple, compact and inexpensive 3-D volume holographic storage systems. Deflection is achieved through the control of an external magnetic field which interacts with a permalloy layer and the current in a 30-turn copper coil. Modeling of the device is also provided here and it agrees with experimental measurements. The use of the mirror is demonstrated in a holographic data storage system and hundreds of holograms have been successfully stored and retrieved.

INTRODUCTION

Volume holographic data storage is achieved by using two laser beams (one "reference" and one "signal" beam) which interfere with each other and form a 3-D grating in a light-sensitive crystal (such as Fe-doped $LiNbO_3$). In this process, the interfering beams change the optical properties of the crystal and the crystal stores this 3-D interference grating. To read the data, only the reference beam is needed to shine through the crystal and the pre-stored 3-D grating inside the crystal reproduces the signal beam. Many multiplexing methods (such as angular, spatial, and wavelength multiplexing) can be used to store a high density of 3-D gratings i.e., holograms, in a limited crystal volume. However, angular multiplexing is the most common and successful technique, of which storage of up to 10,000 holograms in a 1 cm^3 crystal has been demonstrated [1]. Holographic data storage is attractive because it has the combined potential of large data storage capacity (hundreds of billions of bytes in a volume roughly that of a compact disk), high bit rate (a billion bits per second) and small random access time (< 100 μs) [2].

Nevertheless, current volume holographic storage systems are large in size and need a bulky and typically expensive reference beam scanner. Future applications definitely require smaller systems. While most components of current systems such as lenses and lasers can be readily reduced in size, until now there is no commercial beam-steering devices that has the combined merits of low cost, small size and large angle of deflection. Here we then report our work on a MEMS laser-scanning mirror including its behavior as an electromagnetic actuator, and the use of it in a holographic data storage system.

MIRROR DESIGN AND FABRICATION

Our MEMS scanning mirror (Fig. 1) consists of a plate (4x4x0.04 mm^3 in size to accommodate the size of the reference beam) tethered to the bulk silicon substrate by two serpentine springs. Its bottom mirror-like surface (single crystal silicon) serves as the reflector. On the top side of the plate are the electromagnetic components which consist of a 30-turn copper coil and a layer of thin-film permalloy ($Ni_{93}Fe_7$). These two electromagnetic components are used to provide actuation forces.

The fabrication steps (Fig. 2) start with the formation of a 40 μm thick epitaxial silicon membrane by anisotropicaly etching the bulk silicon down to a boron etch-stop layer. The boron layer is then removed and an oxide layer is grown for insulation. A Ti/Cu seedlayer is next evaporated on the top side of the plate. A photoresist mold is formed and a 30-turn planar copper coil is electroplated (9μm thick). Next an insulating photoresist layer is coated, patterned and hard-baked. A Ti/Cu seedlayer is again evaporated, a photoresist mold is formed and an 11 μm thick permalloy layer electroplated. Here, a small piece of permalloy (as shown in Fig. 1) serves as a current return path when current is provided to the coil. Finally a thick layer of photoresist is coated and the spring and plate structures are photo-lithographically defined. The exposed oxide layer is removed in BHF and the silicon underneath is Reactive Ion Etched, to free the mirror.

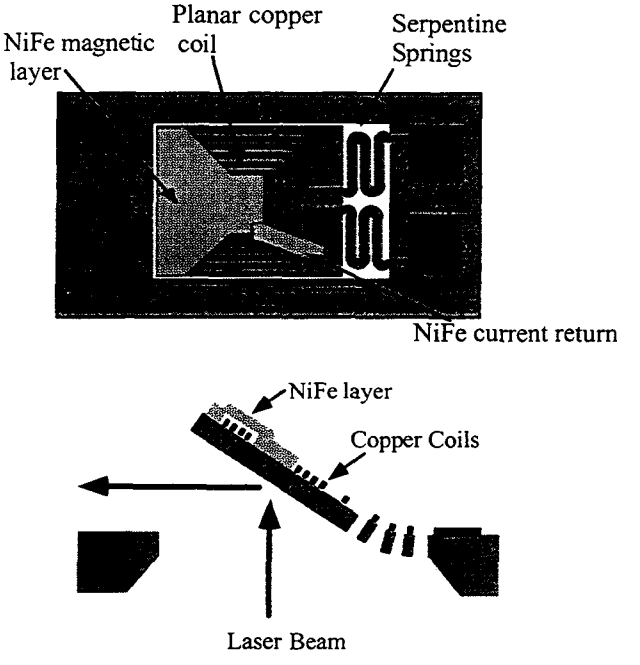


Figure 1. Schematic top view and cross-sectional view (deflected) of a MEMS mirror. The bottom of the plate is used as the mirror surface.

0-9640024-1-8 $3.00 © 1996 TRF

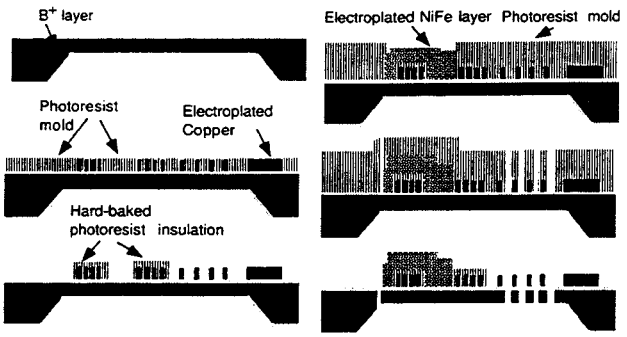


Figure 2. Schematic mirror fabrication steps.

MAGNETIC ACTUATION: THEORY AND EXPERIMENTAL RESULTS

The MEMS mirror can be operated in two modes. First, with no current supplied to the coils, mirror deflections can be induced by the interaction between an external field and the permalloy layer. The deflections can be controlled by changing the strength and angle of the magnetic field. The second is to fix an external field and use the coil current to deflect the mirror. The external magnetic field will first cause the flap to deflect to a fixed "bias" angle. Deflection about this position can then be induced by changing the magnitude and direction of the current supplied to the coils. If both modes are active, the external field can be conveniently used for the coarse deflection control while the fine positional control is done by the current in the coils. These two actuation control methods are discussed in the following.

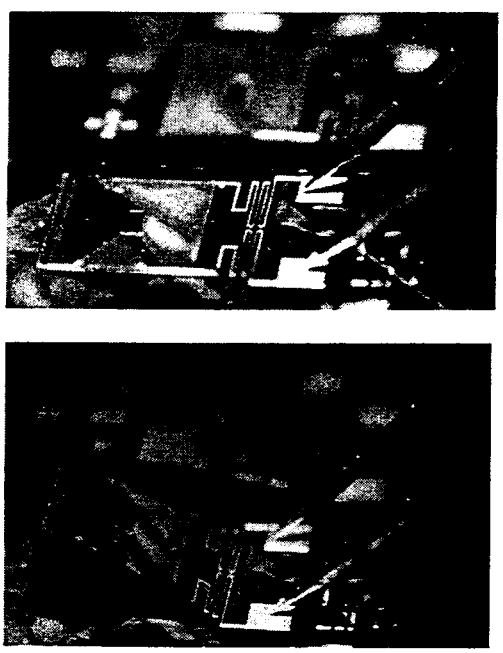


Figure 3. (a) Picture of MEMS mirror in the absence of applied magnetic field and coil current. (b) Actuated mirror in an external field (H_{ext}= 23,880 A/m (300 Oe) vertical to the substrate) with a coil current of around 30 mA. The deflection is caused by both the permalloy and coil current.

PERMALLOY ACTUATION

The mirror actuation due to the permalloy/field interaction was investigated by placing the mirror in a uniform magnetic field H_{ext} (eliminating forces due to field gradients). No current was supplied to the coils. The mirror deflection was then measured for different externally applied field strengths and is shown in Fig. 4. It clearly demonstrates the large deflection (>60°) capability of the mirror.

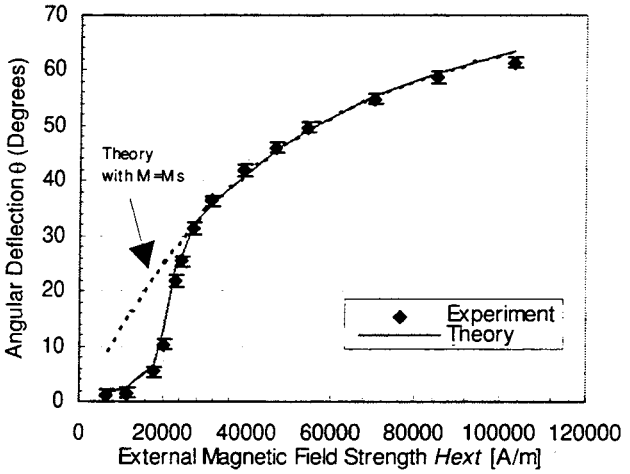


Figure 4. Deflection angle of a MEMS mirror vs. External magnetic field.

The deflection of mirror is modeled using the torque balance (Fig. 5) between a permalloy torque (T_{NiFe}) and, the restoring torque of the spring (T_{spring}). The permalloy torque T_{NiFe} (also described in literature [3,4]) is produced by the interaction between $\boldsymbol{B}_{ext}$ and a magnetic moment $\boldsymbol{m}_{NiFe}$ in the permalloy

$$T_{NiFe} = | \boldsymbol{m}_{NiFe} \times \boldsymbol{B}_{ext} | = V_{NiFe} M_{NiFe} H_{ext} \sin(90 - \theta) \quad (1)$$

Here V_{NiFe}= 7.26x10^{-11} m^3 is the volume of the permalloy, H_{ext} is given in A/m, θ is the deflection angle of the flap relative to the substrate and M_{NiFe} is the magnetization of the permalloy induced by $H_{ext} \cos(90 - \theta)$, the component of H_{ext} parallel to it. The torque (Nm) is then counter-balanced by the restoring torque of the spring as

$$T_{spring} = K \theta \quad (2)$$

Where $K = 4.2X10^{-8}$ Nm/deg is the torsional spring constant, and θ is the deflection angle of the mirror. Eqs. 1 and 2 are then the governing equations used to model the data in Fig. 4, where the theoretical fit (solid line) was generated by equating Eqs. 1 and 2 and solving for θ under different H_{ext} values. Here $M_{NiFe}(H_{ext})$ was determined experimentally by BH-loop (Fig. 5) and Vibrating Sample Magnetometer (VSM) measurements performed on the permalloy layer of this particular mirror. The B-H loop gives the shape of $M_{NiFe}(H_{ext})$, while the VSM measurement determined the saturated magnetization (M_s) of the permalloy to be 0.78 Tesla. An additional curve (dashed line) in Fig. 4 was also plotted with $M_{NiFe} = M_s$. From the point at which this curve and the experimental data diverge, the field H_{sat} required to saturate the permalloy can be determined. H_{sat} is simply the component of H_{ext} in the plane of the permalloy film at this point. It is given by

H_{sat} =27064 *sin* 30° = 13,532 A/m (170 Oe). This value is consistent with the value directly read from the BH-loop (Fig. 5).

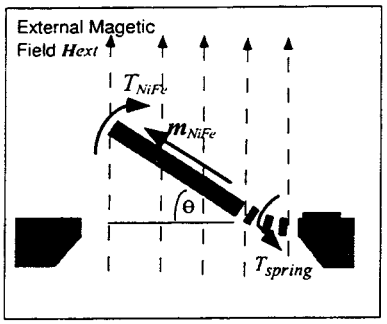


Figure 4. *Torque balance of the mirror with permalloy only in an external magnetic field.*

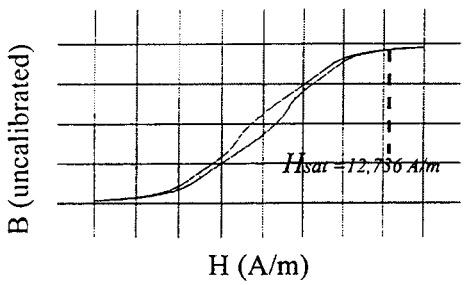


Figure 5. *BH-loop measurement of NiFe layer using an shb-109 BH-looper. Horizontal axis has 3,980 A/m per div. (50 Oe /div.).*

COIL ACTUATION

Here, the effect of varying the coil current in a fixed, uniform externally applied field H_{ext} is discussed. Fig. 6 shows the measured change in deflection angle of the MEMS mirror with an external bias field set to 79,122 A/m (994 Oe), which caused the mirror to deflect to θ =58° due to the permalloy/H_{ext} interaction described previously. This high field value was used to generate large current-induced deflections and to simplify the analysis by ensuring that the permalloy was saturated.

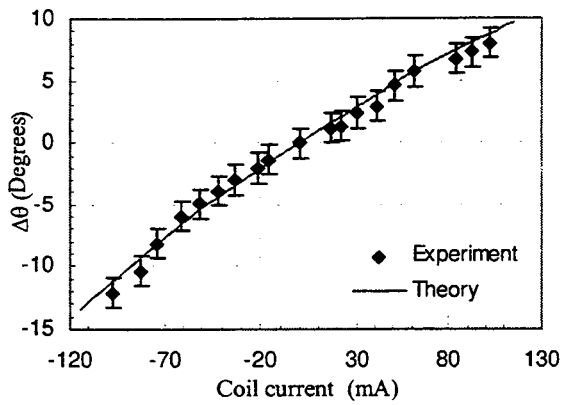


Figure 8. *Change in deflection angle from "bias" position at variable coil current with external field of 79,122 A/m (994 Oe).*

To model Fig. 8, we have to consider two more terms T_{coil} and $T_{thermal}$ in the torque balance equation. First the current in the coils induces an effective magnetic moment $\boldsymbol{m}_{coil}$ (where $m_{coil} = I A_{effective}$). This moment attempts to align with the external field and hence produces a torque on the mirror by[5]

$$T_{coil} = |\boldsymbol{m}_{coil} \times \boldsymbol{B}_{ext}| = m_{coil} B_{ext} \sin\theta \qquad (3)$$

Here I is the current supplied to the coils in amps, $A_{effective}$=*0.000132* m^2 is the sum of 30-turn areas, and $B_{ext} = \mu_o H_{ext}$ is the magnetic flux density in Tesla. Next, the thermal torque is induced by heat generated by this current flowing in the coils. This is modeled by

$$T_{thermal} = -C I^2 \qquad (4)$$

In this case $C = 25.2\text{x}10^{-6}$ is a fitting constant while I is the applied coil current in amps. The negative sign in Eq. 4 means that the thermal torque always bends the mirror down. Now, when a current is applied to the coils with a constant external magnetic field, the torque balance then becomes

$$V_{NiFe} M_s H_{ext} \cos\theta + m_{coil} B_{ext} \sin\theta - C I^2 = K\theta \qquad (5)$$

Eq. 5 has only two unknowns (θ and I), which can be solved and used to fit the data in Fig. 8. The result is shown in the solid line and a good fit is again obtained. It is worth noting that the current control can produce a ± 10° deflection which is more than adequate for the scanning mirror application.

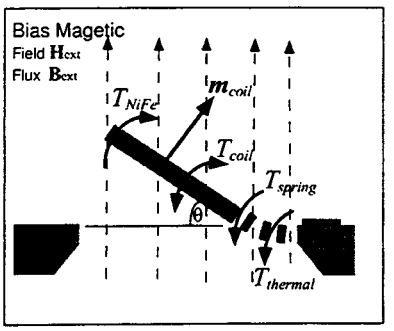


Figure 7. *Torque balance of the mirror with both the permalloy and coil current in an external magnetic field.*

HOLOGRAPHIC APPLICATION

In the holographic data storage application the mirror is biased to approximately 45° in a constant but non-uniform external field (~200 Oe at mirror), provided by a permanent magnet located 2mm above the surface of the mirror. Deflections about this angle (fine control), are induced by varying the coil current.

Two multiplexing schemes are explored. The mirror is first used in an angular multiplexing holographic storage system (Fig. 9a). In this system a single laser beam (Argon laser) is split into a reference beam and a signal beam path. The latter contains an information presentation device (Liquid Crystal Display), a storage crystal, and a detector (CCD camera). The reference beam, is deflected by the MEMS mirror and then passes through a lens system. The lenses keep the beam focused on the same location on the crystal but with different incident angles corresponding the mirrors deflection angle. Fig. 10 shows some typical holograms stored and then retrieved using the MEMS scanning mirror. In addition to the angular multiplexing scheme, the mirror is also used in a combined spatial/angular multiplexing configuration (Fig. 9b). Here, not only is the incident angle of the reference

beam changed but so is its position on the crystal. The primary advantage of this configuration is that two lenses can be eliminated and the mirror placed right at the crystal surface. Once again for demonstration purposes we stored 100 holograms, all were of good quality.

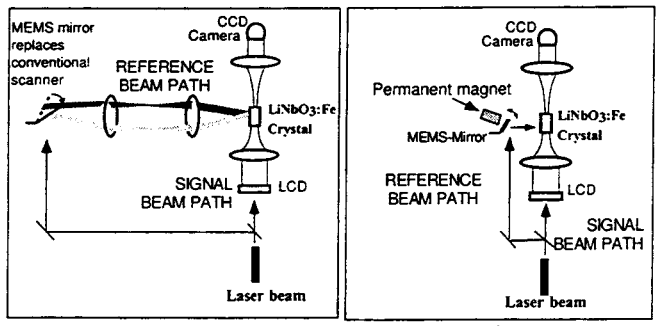


Figure 9. (a) Conventional holographic storage system using commercial beam deflectors. (b) Compact storage system with MEMS mirror.

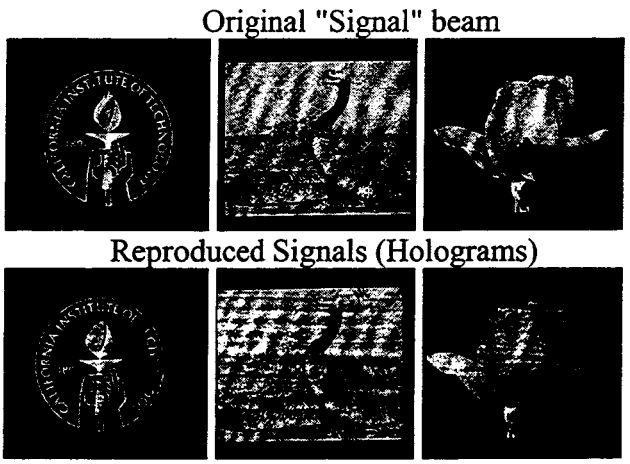


Figure 10. Holograms stored and retrieved, showing original image prior to storage and retrieved image from crystal.

DISCUSSION

Table 1 shows a comparison of the MEMS mirror with some commercial scanning devices. Although we have been able to successfully demonstrate the storage of hundreds of holograms several aspects of the mirror still need improvement.

First, the 0.03° spacing between two holograms is only preliminary data. Experimental improvement needs to be done to further reduce this number. Then, the deflection range of 15° can be reduced while incorporating the course control. Such an approach can actually reduce necessary current and decrease the undesirable thermal effect. Next, the time response of the mirror is relatively slow. Experimentally, it is found that the mirror has a first mode resonant frequency of about 65Hz and a small-angle step response time of 25 ms. It is highly desirable to further reduce this number and future work should increase its damping. Finally, it is apparent that using the MEMS mirror does reduce the overall volume of the demonstrated holographic storage system by more than an order of magnitude. To improve the capacity of the storage then, it is possible to use the saved space to design a multi-mirror array storage system.

Device	MEMS mirror	Commercial Motorized stage	Acousto-optic Deflector
Minimum hologram spacing	0.03°	0.002°	0.0015°
Deflection range	15° *	35°	1.5°
Access time for 0.1° deflection	25 msec	130 msec	70 μsec
Reference beam path volume**	1-10 cm^3 ***	686 cm^3	1372 cm^3
Device volume	1 cm^3	60 cm^3	30 cm^3
Cost	Inexpensive	Expensive	Expensive

* With current control only.
** Not including laser.
*** For compact configuration as in Fig. 9b.

Table 1. Comparison of MEMS mirror with several commercial scanning devices.

CONCLUSION

A large angle of deflection scanning MEMS mirror has been successfully developed. Its operation with either an external field or a coil current has been studied and modeled. The feasibility of using such a prototype scanning mirror for holographic data storage has been demonstrated.

ACKNOWLEDGMENTS

This project was partially supported by AFOSR and DARPA under AFOSR Grant No. F49620-94-1-0008. The authors would like to thank Sang Lee and Gordon Hughes of Seagate technology, and Stanley Bacon for their help with magnetic measurements.

REFERNCES

1. G. W Burr , F. H. Mok and D. Psaltis "Large scale volume holographic storage in the Long Interaction Length Architecture", SPIE proceedings, paper 2297-47, (1994).

2. D. Psaltis and F. H. Mok, "Holographic Memories", Scientific American, pp. 70-75, November (1995).

3. J. W. Judy, R. S. Muller, "Magnetic microactaution of torsional polysilicon structures", Transducers 95, Stockholm, Sweden, 6/25-6/29, pp. 332 - 335, (1995).

4. C. Liu, T. Tsao, Y.C. Tai, J. Leu, et. al, "Out-of plane permalloy magnetic actuators for delta-wing control", Proceedings MEMS 95, Amsterdam, The Netherlands, 1/29-2/2, pp. 328-331, (1995).

5. J. D. Kraus, "ELECTROMAGNETICS", fourth Edition, McGraw-Hill, 1988, pp. 304 - 307.

Batch-Fabricated, Addressable, Magnetically Actuated Microstructures

Jack W. Judy and Richard S. Muller

Berkeley Sensor & Actuator Center (BSAC)
Department of EECS, University of California
Berkeley, California 94720-1770

ABSTRACT

Surface-micromachined, batch-fabricated structures that combine plated-nickel films with polysilicon mechanical flexures to produce individually addressable, magnetically activated devices have been fabricated and tested. Individual microactuator control has been achieved two ways: (1) by actuating devices using the magnetic field generated by coils integrated around each device, (2) by using electrostatic forces to clamp selected devices to an insulated ground plane while unclamped devices are freely moved through large out-of-plane excursions by an off-chip magnetic field. The present application for these structures is as micromirrors for microphotonic systems where they can be used either for selection from an array of mirrors or else individually for switching among fiber paths.

INTRODUCTION

Magnetic actuation of polysilicon flexural MEMS elements provides several desirable features. Techniques to batch-fabricate these elements using electrodeposited NiFe films together with details of their performance were presented at this workshop in 1994 [1-2]. Based upon this earlier work, the following features were demonstrated.

(1) Large deflections (> 1 mm) are achieved using magnetic forces to actuate compliant microflexures.

(2) These large deflections can be achieved both parallel and perpendicular to the plane of the wafer.

(3) Actuation can be achieved using magnetic fields generated by either on- or off-chip sources (enabling remote device control)-- or by applying both at the same time.

(4) The actuating force can be applied in a conducting environment such as a saline fluid.

Torsion-beam suspensions of these magnetic elements were described in later work [3], that demonstrated their applicability for optical scanners, displays, and switches. We report here an enhancement of the static actuation model, a dynamic actuation model, and two methods for controlling individual flexure-based ferromagnetic microactuators. The first method uses coils integrated around each device to produce the magnetic field required for individual microactuator motion (Figure 1). The second method uses electrostatic forces to clamp selected devices to the insulated ground plane, while the off-chip magnetic field actuates unclamped devices (Fig. 2).

STATIC ACTUATION

A static mechanical and magnetic analysis of these torsional ferromagnetic microactuators has been given [3]. The accuracy of the model in reference [3] can be enhanced, particularly at low magnetic fields (where agreement with experimental results is least accurate), by accounting for the coercivity H_c of the magnetic material. If H_c is considered, the magnetization M in the film is expressed as

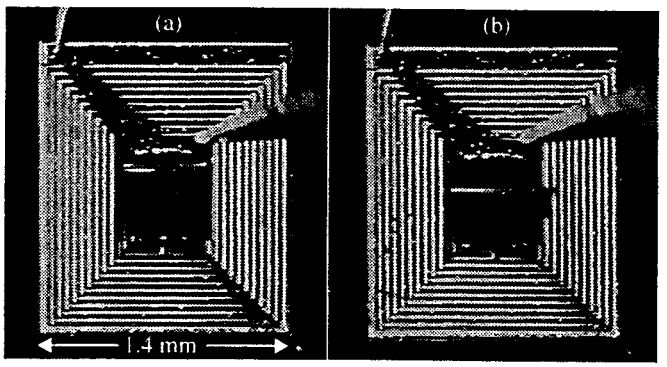


Figure 1. Top views of a microactuated torsion-beam element (0.45 mm square) surrounded by a 10-turn coil; a) no current applied to the coil, b) rotated 45° by H ≈ 5 kA/m when 500 mA is applied to the coil.

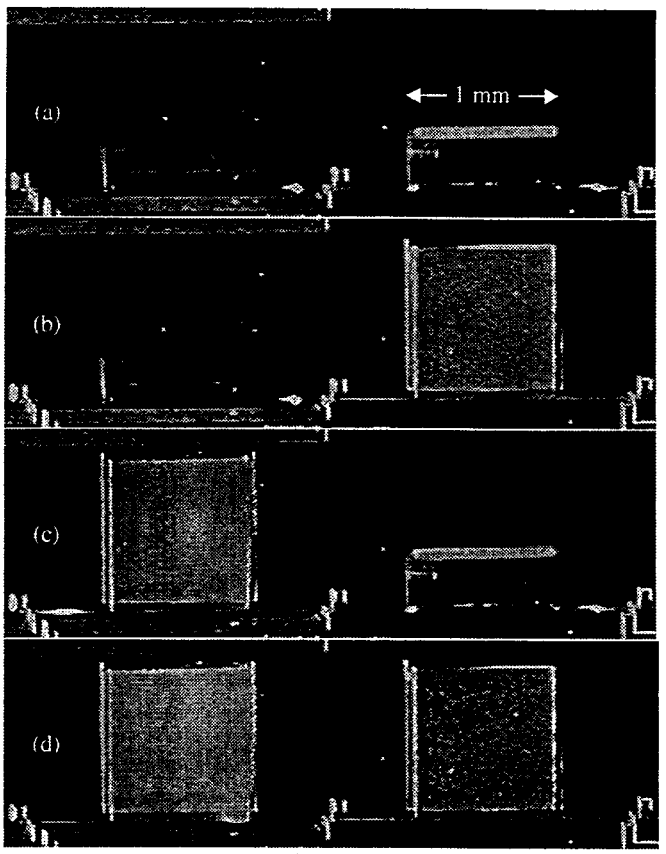


Figure 2. Top views of electrostatically addressed torsion-beam elements (1 mm square); (a) both devices actuated by an off-chip magnetic field H ≈ 8 kA/m, (b) right device electrostatically clamped down and left device actuated, (c) left device clamped down and right device actuated, (d) both devices clamped down ($V_c \approx 5$ V).

0-9640024-1-8 $3.00 © 1996 TRF

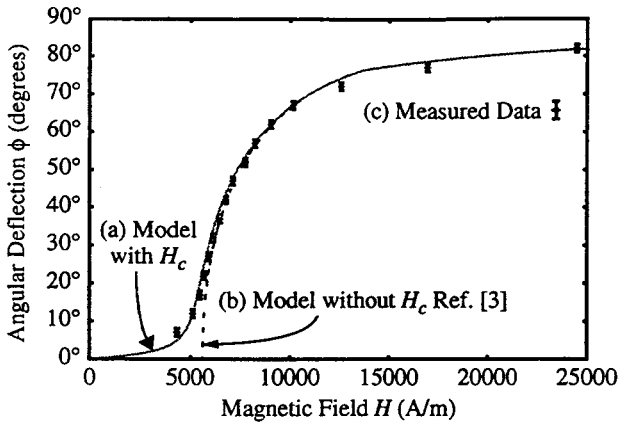


Figure 3. Comparison of calculated and measured dc-angular deflection as a function of magnetic field.

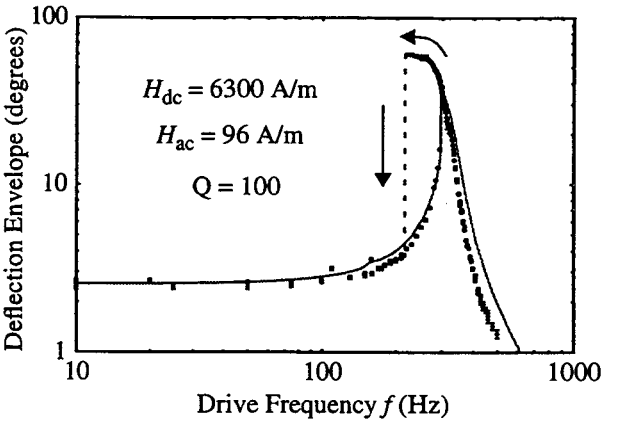


Figure 4. Comparison of calculated and measured mechanical frequency response.

$$M \approx \min\left(\frac{\mu_o (H_c + H_{ext} \cos(\gamma - \theta - \phi))}{\sqrt{N_l^2 \cos^2\theta + N_t^2 \sin^2\theta}}, \; M_s \right) \quad (1)$$

with the permeability of free space μ_o, the angle between the external magnetic field and the substrate γ, the angular deflection of the magnetization from its preferred axis (the easy axis) θ, the angular mechanical deflection from the substrate ϕ, and the magnetic-shape coefficients along the length N_l and thickness N_t of the ferromagnetic element. Figure 3 shows the angular deflection predicted (a) by Eq. (1) (with H_c = 250 A/m), (b) by Eq. (12) of reference [3], and (c) the measured results. Equation (1) is seen to provide excellent agreement even at low magnetic fields.

DYNAMIC ACTUATION

To predict actuation under ac conditions, we formulate a dynamic torsional model:

$$J\ddot{\phi} + C_\phi \dot{\phi} + k_\phi \phi = T_m \quad (2)$$

with angular moment-of-inertia J, angular damping coefficient C_ϕ, angular stiffness of the torsional beam k_ϕ, and driving magnetic torque T_m. The angular damping coefficient can be expressed as $C_\phi = \sqrt{Jk_\phi}/Q$ with quality factor Q. Equation (2) is not a simple linear differential equation because the torque T_m is a function of the angle between the ferromagnetic element and the substrate ϕ. If we assume that the magnetization is along the easy axis ($\theta = 0$), and that the field is perpendicular to the substrate ($\gamma = 90°$), the torque T_m is

$$T_m = V_{mag} M (H_{dc} + H_{ac} \sin 2\pi f t) \cos(\phi) \quad (3)$$

with magnet volume V_{mag}, dc field H_{dc}, ac field H_{ac}, and drive frequency f. When Eq. (3) is substituted into Eq. (2), the result can be solved numerically. A plot comparing calculations for the mechanical frequency response with experimental results, is given in Fig. 4 for a device consisting of a 430×130×15 μm^3 NiFe plate (J = 4.57×10^{-16} kg-m^2 and V_{mag} = 8.4×10^{-13} m^3) attached to a pair of 400×2.2×2.2 μm^3 polysilicon torsional beams (k_ϕ = 2 nN-m/rad) with fitting parameter Q = 100. Two features of the plot in Fig. 4 are readily apparent; first, the model agrees quite well with measurements, second, the peak of the resonance curve leans to the left, indicating that the system exhibits a nonlinear, hysteretic frequency response.

ADDRESSING

A single off-chip field source can be used to actuate large arrays of microelements simultaneously, which can be useful for microtransport and microassembly applications [4], but actuation only of selected elements would permit a number of other uses. We have pursued two methods to achieve addressable actuation in an array of devices. The first method employs magnetic fields generated by coils integrated around each element to achieve a localized field and actuation (Fig. 1) [5]. An option for this scheme is to supplement this coil-driven field with an integrated ferromagnetic core. Although integrated coils with a ferromagnetic core can concentrate magnetic flux, thereby achieving a larger magnetic field, their fabrication process is considerably more complicated [6-7]. For our experiments, planar coils, which spiraled around each microactuator, were fabricated without any magnetic core. The second method for addressable actuation uses electrostatic forces to clamp selected devices to the insulated ground plane, so that the off-chip magnetic field does not affect the clamped devices (Fig. 2).

INTEGRATED COIL

Figure 1 shows a coil-driven device consisting of a 450 μm-square nickel pattern plated on polysilicon, which is attached to the insulated substrate via two torsional springs 200 μm long, 2 μm wide, and 2 μm thick. The electroplated 5 μm-thick nickel is also used as the conducting layer for the 10-turn coil (25 Ω) surrounding the microactuator. The entire device, consisting of the coil and the moving structure, requires a square area 1.4 mm on a side. With no current in the coil, the device lays flat as in Fig 1a; when 500 mA is applied a field of ~5 kA/m is generated that rotates the device more than 45° out of the plane of the wafer. On-chip coil-actuated devices can achieve continuous positional control, however, high power (6.25 W) and large chip area (2.25 mm^2) are required for this design.

Using copper instead of nickel for the coil would reduce the resistance by a factor of ~4 and also prevent the coil from shielding the device from any off-chip magnetic field, so that a small permanent magnet could be incorporated underneath the substrate to provide the dc field necessary to rotate the plate out of the plane of the wafer. Thus, the copper integrated coil would need to provide only a much smaller ac field. Driven at the resonant frequency, such a device should exhibit a large scanning envelope with significantly less power.

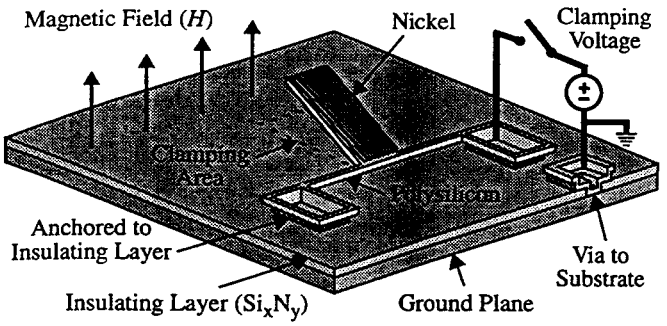


Figure 5. Schematic representation of the electrostatic clamping structure.

ELECTROSTATIC CLAMP

Individual selection of activated elements in an array, that is actuated by a uniform off-chip source, is possible using electrostatic forces to address individual microactuators, as is shown schematically in Fig. 5. Selection is accomplished by applying a voltage to the microactuator, which is anchored to an insulating dielectric layer, while the substrate is grounded. The resultant electrostatic force clamping the element to the surface, must be sufficient to keep selected devices from moving in response to the magnetic field applied to actuate unselected devices.

The clamping force F_c is approximately that felt by the plates of a simple capacitor

$$F_c = \frac{\varepsilon_r \varepsilon_o A V^2}{2d^2} \tag{4}$$

with relative permittivity of the dielectric layer ε_r, permittivity of free space ε_o, clamping area $A = wl$, voltage V, and electrode separation d. In Eq. 4 it is assumed that the clamping voltage is positive, putting the surface of the n-type silicon substrate into accumulation, and that trapped charge in the nitride and at the substrate-nitride interface is negligible. The clamping action can be formulated in terms of the clamping torque T_c, given by

$$T_c = \int_0^l F_c dl' = \frac{\varepsilon_r \varepsilon_o w l^2 V^2}{4d^2} . \tag{5}$$

The clamping torque T_c must exceed the drive torque T_m for the clamp to be effective. If we assume that the external magnetic field is applied perpendicular to the substrate ($\gamma = 90°$) and that the magnetization remains along the easy axis ($\theta = 0$), when the device is clamped ($\phi = 0$) T_m is

$$T_m(\phi = 0) = V_{mag} MH \tag{6}$$

with magnet volume $V_{mag} = lwt$, magnetization M, and magnetic field H. The minimum clamping voltage V_c derived by equating Eqs. (5) and (6) is

$$V_c > \sqrt{\frac{4MHtd^2}{\varepsilon_r \varepsilon_o l}} . \tag{7}$$

The electrostatic clamp has been experimentally investigated using the structures shown in Fig. 2. These devices consist of nickel-plated polysilicon plates (1 mm long, 1 mm wide, and 10 μm thick) that rotate about a pair of torsional beams (400 μm long, 15 μm (left device) to 20 μm wide (right device), and 2 μm thick) that are anchored to a 2.0 μm-thick silicon-nitride layer deposited on the n-type silicon substrate. For these plate structures, using ($\varepsilon_r = 7.5$ and $M = 0.7$ T), a voltage of $V = 4.1$ volts is calculated as sufficient to clamp the device in a field $H = 10$ kA/m. Experimental results

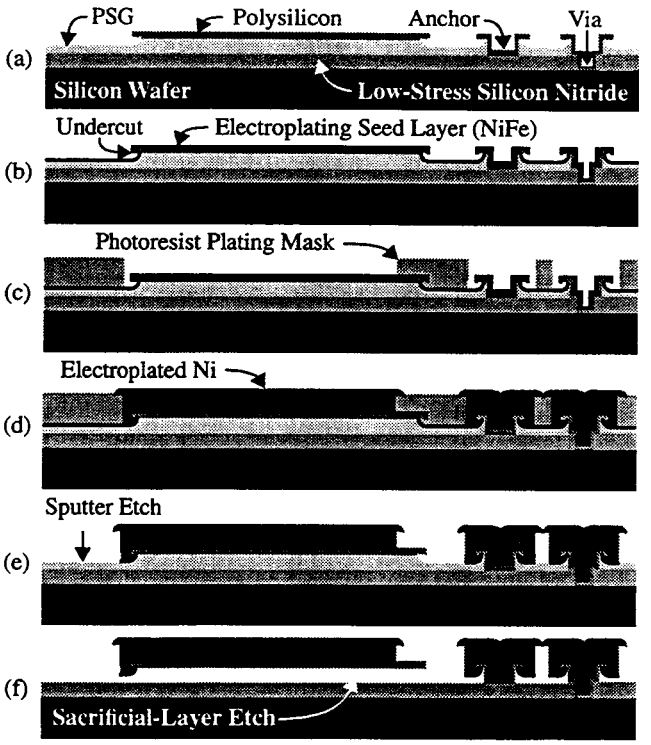


Figure 6. Fabrication process.

have confirmed that ~5 volts can clamp these structures in a 10 kA/m field.

The sequence of images shown in Fig. 2 illustrate individual device selection using electrostatic clamping. In Fig. 2(a), both microactuators are rotated up out of the plane of the wafer. The device on the right in Fig. 2a does not rotate as far as the device on the left because the beams supporting it are wider (20 μm), and hence stiffer, than the beams supporting the device on the left (15 μm). After taking the photograph in Fig. 2(a), the off-chip magnetic field has been turned off so that the devices return to the substrate after which a clamping voltage is applied to the device on the right. Figure 2(b), taken after the magnetic field is turned back on, shows the device on the right clamped down while the device on the left is rotated up by the magnetic field. In Fig. 2(c), the situation is reversed; that is, the device on the left is clamped down and the device on the right is rotated up by the field. Figure 2(d), shows the case of both devices clamped with the magnetic field applied.

Advantages of using electrostatic forces (instead of using integrated coils) to achieve individual microactuator control include the following:

(1) Arrays of elements can be readily addressed using well-known digital-memory address techniques.

(2) The clamping scheme is easily incorporated in a batch-fabrication process.

(3) Clamping is accomplished with very little increase in the area of an array in contrast to that needed for on-chip coils.

(4) No static power is needed for the clamped devices.

FABRICATION

Both the coil-driven and the electrostatically addressed microactuators are constructed in a batch-fabrication process, as shown in Fig. 6, that combines ferromagnetic materials, nickel and nickel iron, and conventional silicon-process materials, single crystal-silicon, silicon nitride, phosphorus-doped silicon dioxide (PSG),

and polycrystalline silicon (polysilicon). The ferromagnetic materials are introduced as a last step following nearly all the conventional IC processing.

The fabrication process begins with an n-type silicon wafer (ρ = 10 Ω-cm) which will be used as the ground plane. Although not done in this process, we could have, at this point, formed heavily-doped regions on the surface of the wafer to act as clamping counter electrodes. A 2.0 μm-thick layer of LPCVD silicon-rich, low-stress silicon nitride is deposited on the wafer surface to form the insulating layer. The clamping voltage can be reduced by thinning the nitride layer, with the limit that it must be thick enough to survive the prolonged hydrofluoric acid (HF) release etch which slowly attacks silicon nitride. A 2.0 μm-thick layer of LPCVD PSG is deposited to form the sacrificial layer.

The first photolithography step defines holes through the PSG layer which will serve as insulated anchor positions for the microactuator. After stripping the first photoresist mask, a second lithography step defines holes through both the PSG and the silicon nitride layer so that electrical contact can be made to the substrate. After stripping the resist, a brief etch in HF is used to remove any native oxide and a 2.0 μm-thick layer of LPCVD polysilicon (605 °C) is deposited to be used as the structural layer. After a 0.5 μm-thick layer of LPCVD PSG is deposited on top of the polysilicon, the wafer is annealed in nitrogen at 1000 C for one hour to dope the polysilicon layer with phosphorus from the PSG and to reduce residual stress in the polysilicon film. A third lithography step defines the top PSG layer with the pattern for the mechanical structure. After the resist is stripped, the top layer of patterned PSG acts as a mask during a reactive-ion etch that selectively etches through the polysilicon layer until it stops on the PSG sacrificial layer. The top PSG layer is then removed in a wet HF etch that also attacks the revealed sacrificial layer and undercuts the polysilicon microstructure by about 0.5 μm (Fig. 6a).

A 0.1 μm-thick layer of NiFe is sputter deposited to form the electroplating seed layer. Since sputtering deposits material in many directions, some NiFe is deposited underneath the undercut polysilicon layer (Fig. 6b). The fourth and final lithography step defines a 4.0 μm-thick resist plating mask, which must be thick enough to cover the step from the sacrificial-PSG layer to the structural-polysilicon layer (Fig. c). If the mask is too thin, spurious hemispherical electrodeposits will form where the electric field breaks down the resist, particularly at the corners of the polysilicon structure where the electric field is concentrated. The plating mask is defined 50 μm beyond the edge of the polysilicon plate on three edges (the far end and 90% of the way back along the sides toward the torsional bars). A significant overexposure is required to develop and remove the resist fully from underneath the polysilicon overhang.

Nickel is next electroplated onto the exposed areas of the seed layer, covering the top, the side, and underneath the polysilicon overhang (Fig. 6d). Thus, the ferromagnetic material is formed in a shape that essentially grasps the polysilicon layer and thereby achieves a good mechanical bond, which is helpful because nickel does not adhere well to polysilicon. We have found that adhesion layers, such as chrome and titanium, do not perform satisfactorily when exposed to a prolonged HF release etch. After the resist plating mask is stripped, the excess seed layer is removed by sputter etching (Fig. 6e), and the wafer is placed in a prolonged HF etch that removes the PSG sacrificial layer (Fig. 6f) (30 minutes for the devices shown in Fig. 2). A supercritical carbon dioxide drying technique is used to prevent the microstructures from sticking to the substrate after the release step [8].

CONCLUSIONS

We have shown that ferromagnetic microactuators with torsional-polysilicon flexures, capable of very large out-of-plane displacements (> 1 mm), can be individually controlled with integrated coils or individually addressed by clamping with electrostatic forces. Individual prototype-torsional devices are shown to be deflected 45° out of the plane of the wafer, by the magnetic field generated when a current of 500 mA flows through a 10-turn coil integrated around each device. Individual prototype-torsional devices, fabricated on a dielectric coated ground plane, can be electrostatically clamped in place using 5 V address potentials when an off-chip magnetic field of 10 kA/m is applied. This field torques unaddressed structures roughly into a perpendicular position. We have also shown that ferromagnetic films can be compatibly integrated with silicon-based materials and processes into a batch-fabrication process with good adhesion results using an under-etch technique to hold the films to the polysilicon layers. By accounting for the coercivity of the ferromagnetic material, the agreement of the static-deflection model with experimental results is improved, particularly at low magnetic fields. A model predicting the dynamic mechanical behavior of these devices has been developed which agrees quite well with experimentally measured frequency responses. An important application area for individually controllable ferromagnetic microactuators with polycrystalline silicon torsion-bar flexures is to microphotonic applications, such as optical scanners, displays, and switches.

REFERENCES

1. J. W. Judy, R. S. Muller, and H. H. Zappe, "Magnetic microactuation of polysilicon flexure structures," Tech. Dig. Solid-State Sensor and Actuator Workshop (Hilton Head `94), Hilton Head Island, SC (June 13-16, 1994), pp. 43-48.
http://www-bsac.eecs.berkeley.edu/archive/conference/hh1994/jjudy/

2. J. W. Judy, R. S. Muller, and H. H. Zappe, "Magnetic microactuation of polysilicon flexure structures," IEEE Journal of Microelectromechanical Systems, v. 4, no. 4, pp. 162-169, 1995.
http://www-bsac.eecs.berkeley.edu/archive/journal/jmems/jjudy/

3. J. W. Judy and R. S. Muller, "Magnetic microactuation of torsional polysilicon structures," 8th International Conference on Solid-State Sensors and Actuators Digest of Technical Papers (Transducers `95), Stockholm, Sweden, (June 25-29, 1995), v. 1, pp. 332-335.
http://www-bsac.eecs.berkeley.edu/archive/conference/trans1995/jjudy/

4. C. Liu, T. Tsao, Y.-C. Tai, W. Liu, P. Will, and C.-M. Ho, "A micromachined permalloy magnetic actuator array for micro robotics assembly systems," 8th International Conference on Solid-State Sensors and Actuators Digest of Technical Papers (Transducers `95), Stockholm, Sweden, (June 25-29, 1995), v. 1, pp. 328-331.

5. B. Wagner, W. Benecke, G. Engelmann, and J. Simon, "Microactuators with moving magnets for linear, torsional, or multiaxial motion," Sensors and Actuators A (Physical), v. A32, no. 1-3, pp. 598-603, 1992.

6. C. H. Ahn, Y. J. Kim, and M. G. Allen, "A planar variable reluctance magnetic micromotor with fully integrated stator and wrapped coils," Proc. IEEE Micro Electro Mechanical Syst. (MEMS `93), Fort Lauderdale, Florida, (February 7-10, 1993), pp. 1-6.

7. H. Guckel, T. R. Christenson, H. J. Skrobis, T. S. Jung, J. Klein, K. V. Hartojo, and I. Widjaja, "A first functional current excited planar rotational magnetic micromotor," Proc. IEEE Micro Electro Mechanical Syst. (MEMS `93), Fort Lauderdale, Florida, (February 7-10, 1993), pp. 7-11.

8. G. T. Mulhern, D. S. Soane, and R. T. Howe, "Supercritical Carbon Dioxide Drying of Microstructures," 7th International Conference on Solid-State Sensors and Actuators Digest of Technical Papers (Transducers `93), Yokohama, Japan, (June 7-10, 1993), pp. 296-299.

PIEZOELECTRICALLY ACTUATED MICROCANTILEVER FOR ACTUATED MIRROR ARRAY APPLICATION

Yongli Huang, Hong Zhang, and Eun Sok Kim

Department of Electrical Engineering
University of Hawaii at Manoa
Honolulu, Hawaii 96822

Sang Gook Kim and Yong Bae Jeon
Daewoo Electronics Co.
Seoul, Korea 100-714

ABSTRACT

For a projection display application, we have developed a piezoelectrically actuated array of cantilevers whose facets can accurately be controlled for a desired gray level. A 40x40 cantilever array with a pixel size of 100x100 μm^2 has been fabricated, and a piezoelectric ZnO thin film (in spite of its relatively low d_{31}) is successfully used to produce 0.116°/V vertical deflection of the cantilever. In the fabrication processing, we have discovered that many materials have their unique critical HF concentration below which there is no etching by vapor HF, and successfully used vapor HF to release cantilevers with a very high yield and a processing simplicity.

INTRODUCTION

In the digital mirror array [1] and the grating light valves [2] for a projection display application, individual mirror elements are deformed or tilted electrostatically. The control of mirror positions in those devices is basically bistable in that the mirror elements are tilted or deformed into either one of two stable positions. Consequently, the gray scale control inherently requires high speed digital circuitry. Moreover, since the mirror elements are constantly banged into the substrate during the operations, long-term reliability may suffer due to in-operation stiction.

On the other hand, a piezoelectric actuation of mirror elements allows us to have linear control of the mirror tilting, and the gray scale control becomes very simple. Also, the mirror elements formed by surface micromachining never have to touch the substrate during operation, and the in-operation stiction is not a concern. We have fabricated an array of piezolectrically actuated cantilevers for a projection display system illustrated in Fig. 1. In the system, the piezoelectrically actuated cantilevers are individually tilted for an image formation through a light modulation technique shown in Fig. 2.

With bulk piezo-ceramics (PLZT), a 200 x 400 array of mirror elements was demonstrated to show high light throughput efficiency [3].

A ferroelectric PZT film has large piezoelectric constant (e.g., d_{31}), and can advantageously be used to actuate microstructures. However, PZT thin film lacks a track record up to now. While sputter deposition of PZT requires very tight process control for repeatable quality of PZT films, sol-gel PZT typically has large residual stress. In contrast, a piezoelectric ZnO thin film has been extensively used in commercial SAW filters for TV's IF filters (tens of million devices per year), and has been proven to be reliable and reproducible (i.e., the ZnO properties are well controlled).

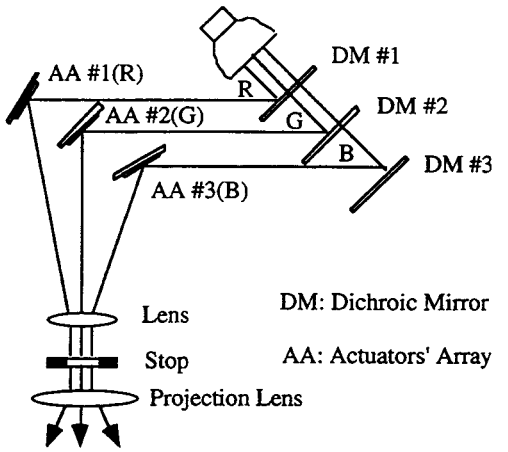
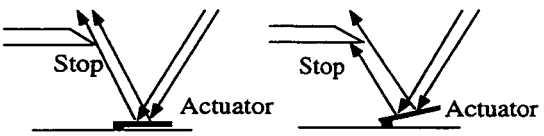


Figure 1. Conceptual picture of the projection display system using the piezoelectrically actuated mirror array.



a) No Tilting (100% passing) b) Tilting (50% passing)

Figure 2. The light modulation technique with a knife-edge filter.

Moreover, ZnO film (deposited by a rf magnetron sputtering) has been demonstrated to be compatible with a CMOS processing [4]. Though ZnO has relatively low piezoelectric constants, there are many actuator applications whose specifications can be met by ZnO film. After a careful calculation, we concluded that it is feasible to meet the specifications for an actuated mirror array application with ZnO film, and have successfully fabricated working arrays (40 x 40) of ZnO-actuated micro-cantilevers.

THEORY AND DESIGN

For a multi-layer cantilever actuated by a piezoelectric film, the neutral plane is typically designed to be at the bottom plane of the piezoelectric film in order to induce maximum bending stress in the piezoelectric film. For a large deflection, the cantilever should be designed as thin as

0-9640024-1-8 $3.00 © 1996 TRF

possible. Consequently, we first set the thickness of each layer to its minimum thickness limited by a processing. Then, based on the thickness of each layer, we calculate the thickness of ZnO to make the neutral plane of the cantilever coincide with the bottom surface of the ZnO layer. The cross section of our cantilever viewed from the free end is shown in Fig. 3. The requirement that the net axial force on the neutral plane be zero gives the following equation:

$$P = \sum \Delta P_i = \sum \frac{E_i y_i \Delta A_i}{R} = 0 \qquad (1)$$

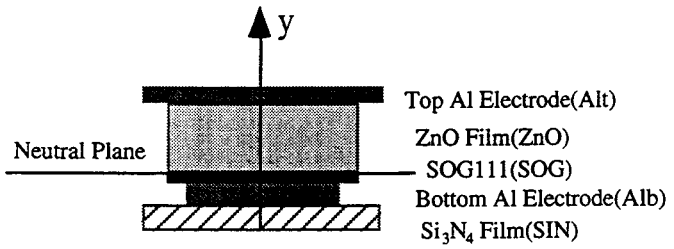


Figure 3. The cross-sectional view of a typical cantilever employed in our mirror array.

where E_i, R, and ΔA_i are the elastic modulus of a layer i, the bending curvature and the cross section area of a layer i, respectively. For the multi-layer beam in Fig. 3, we can derive the following equation for the ZnO thickness (t_{ZnO}):

$$a t_{ZnO}^2 + b t_{ZnO} + c = 0, \qquad (2)$$

where

$$\begin{aligned} a &= E_{ZnO} W_{ZnO} \\ b &= 2 E_{Al} W_{Alt} t_{Alt} \\ c &= E_{Al} W_{Alt} t_{Alt}^2 - E_{SOG} W_{SOG} t_{SOG}^2 \\ &\quad - E_{Al} W_{Alb} [(t_{Alb} + t_{SOG})^2 - t_{SOG}^2] - E_{SiN} W_{SiN} \\ &\quad \times [(t_{Alb} + t_{SOG} + t_{SiN})^2 - (t_{Alb} + t_{SOG})^2] \end{aligned} \qquad (3)$$

where W_i, and t_i are the width and the thickness of a layer i, respectively.

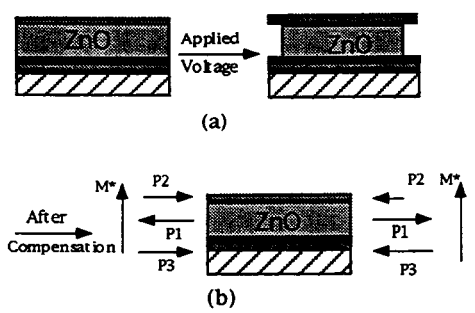


Figure 4. (a) A segment of the cantilever along the length direction is shown to produce a piezoelectric strain under an applied voltage. The shown strain is for the case where the layers are not attached to each other. (b) The piezoelectric strain for the real case (where all the layers are attached to each other) can be equivalently modeled using forces applied to the layers as shown.

As illustrated in Fig. 4, calculation of the deflection and tip displacement of ZnO-actuated cantilever can be treated as a pure bending with the momentum M^* and moment of inertia I^* [5]. The force densities (P in N/m^2) must satisfy the following equations:

$$\begin{aligned} P_1 &= \frac{d_{31} E E_{SiN} (S_2^* + S_3^*)}{S_1^* + S_2^* + S_3^*} \\ P_2 = P_3 &= \frac{d_{31} E E_{SiN} S_1^*}{S_1^* + S_2^* + S_3^*} \end{aligned} \qquad (4)$$

where E is an applied electric field across ZnO film; and S_1^*, S_2^* and S_3^* are the effective cross-section areas defined by:

$$\begin{cases} S_1^* = n_1 W_{Alt} t_{Alt} \\ S_2^* = n_2 W_{ZnO} t_{ZnO} \\ S_3^* = n_1 W_{Alb} t_{Alb} + n_3 W_{SOG} t_{SOG} + W_{SiN} t_{SiN} \end{cases} \qquad (5)$$

with $n_1 = \frac{E_{Al}}{E_{SiN}}$, $n_2 = \frac{E_{ZnO}}{E_{SiN}}$ and $n_2 = \frac{E_{SOG}}{E_{SiN}}$.

Thus, M* and I* can be obtained by following equations:

$$\begin{aligned} M^* &= \frac{P_1}{2} n_2 W_{ZnO} t_{ZnO}^2 + \frac{P_2}{2} \{ n_3 W_{SOG} t_{SOG}^2 \\ &\quad + n_1 W_{Alb} [(t_{Alb} + t_{SOG})^2 - t_{SOG}^2] \\ &\quad + W_{SiN} [(t_{Alb} + t_{SiN} + t_{SOG})^2 - (t_{Alb} + t_{SOG})^2] \\ &\quad - n_1 W_{Alt} [(t_{Alt} + t_{ZnO})^2 - t_{ZnO}^2] \} \end{aligned} \qquad (6)$$

$$\begin{aligned} I^* &= \frac{1}{3} \{ n_2 W_{ZnO} t_{ZnO}^3 + n_1 W_{Alt} [(t_{Alt} + t_{ZnO})^3 - t_{ZnO}^3] \\ &\quad + n_3 W_{SOG} t_{SOG}^3 + n_1 W_{Alb} [(t_{SOG} + t_{Alb})^3 - t_{SOG}^3] \\ &\quad + W_{SiN} [(t_{Alb} + t_{SiN} + t_{SOG})^3 - (t_{SOG} + t_{Alb})^3] \} \end{aligned} \qquad (7)$$

But the deflection curvature of the cantilever bending ($1/R$) is

$$\frac{1}{R} = \frac{M^*}{E_{SiN} I^*} \qquad (8)$$

For a cantilever array shown in Fig. 8, the cantilever deflection angle θ and the tip displacement Δy are:

$$\begin{cases} \theta = \frac{L_1}{R} \\ \Delta y = R(1 - \cos\theta) + (L - L_1)\sin\theta \end{cases} \qquad (9)$$

where L and L_1 are the whole cantilever length and the portion of L that is covered by the ZnO, respectively.

For a cantilever with L = 150 μm and L_1 = 70 μm, the deflection θ and the tip displacement Δy are calculated to be 0.06 °/V and 0.12 μm/V, respectively.

The fundamental resonant frequency of the cantilever can be estimated by the following equation:

$$f_r = \frac{1.875^2}{2\pi l^2}\sqrt{\left(\frac{EI}{\rho A}\right)_{eff}} \qquad (10)$$

where l, E, I, ρ and A are the length, the elastic modulus, the moment of inertia, the density and the cross-section area of composite cantilever, respectively. We typically design f_r to be about 150 kHz, much higher than the operation frequency.

DEVICE FABRICATION

A 40x40 array of ZnO actuated cantilevers (shown in Figs. 6 and 7) is fabricated on a silicon wafer with the fabrication steps illustrated in Fig. 5. The following is a brief description of the 7-mask process. After depositing Si_3N_4 for an electrical isolation, a doped LPCVD polysilicon is deposited and patterned for connecting bottom electrodes. A sacrificial layer of Spin-On-Glass (SOG) is then spin-coated and cured in two steps (at 400 °C for 20 min. and then 850 °C for one hour). Two types of SOG (SOG311 and SOG511 from Accuglass) are used as sacrificial layers for the thicknesses of 0.5 μm and 1 μm. After patterning the SOG for the anchor area, a cantilever supporting layer of 1300 Å thick low stress LPCVD Si_xN_y is deposited. After patterning the silicon nitride with an RIE for an electrical feedthrough window, a 1000 Å thick Al film is deposited by evaporation for the bottom electrode. Additional layer of 3000 Å thick Al is deposited for a step coverage. After spin coating a 1800 Å thick SOG (SOG111) for an electrical isolation, 4500 Å thick ZnO film is sputter-deposited. In a certain HF concentration, the lateral etch of ZnO by vapor HF is very slow, and the SOG111 can be patterned by vapor HF using the same photolithography that is used to pattern the ZnO so that we save one photolithographic step. Finally 800 Å thick Al is evaporated, and used as the top electrode, the reflective surface and the etch mask during the etch channel formation and removal of the sacrificial layer. The perspective view of the cantilever element is shown in Fig. 8.

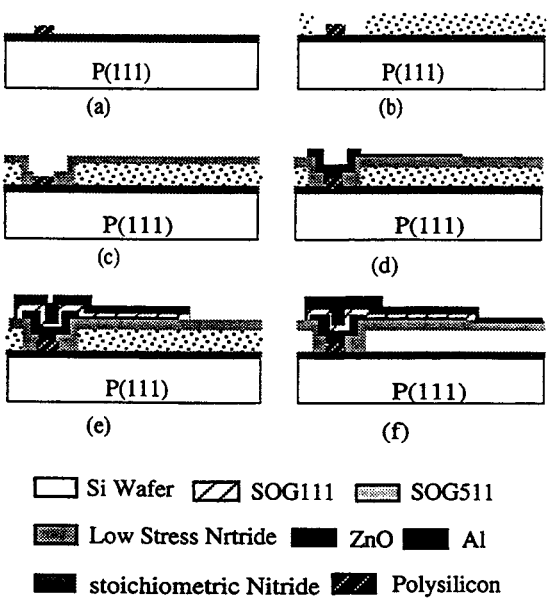


Figure 5. Fabrication process of the cantilever array (cross-sectional views across A-A' in Fig.6).

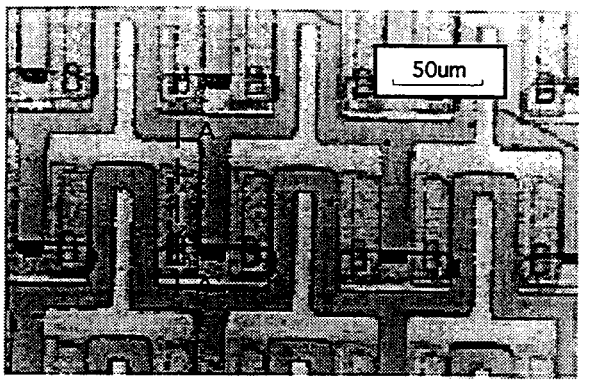


Figure 6. Photograph of close-up top view of the piezoelectrically-actuated cantilever array for a projection display. Each pixel occupies 100 x 100 μm² area.

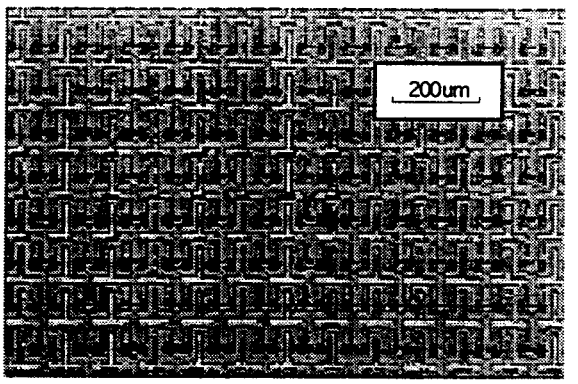


Figure 7. Photograph of the part of the fabricated 40 x 40 array of the piezoelectric cantilevers.

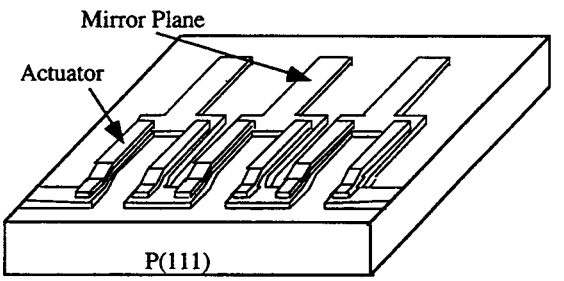


Figure 8. Perspective view of the cantilever array.

Vapor HF is used to release the cantilever. We have discovered that many materials have their unique critical HF concentration below which there is no etching by vapor HF. The critical HF concentration varies depending on a type of a material. As we can see in Fig. 9, Al is rarely etched at a sufficiently diluted vapor HF (i.e., 75 ml/min : 6 ml/min of N2 : HF), while SOG is still highly etched. Thus, with that HF concentration, bare 800Å thick Al can be used as an etch mask during etching SOG to release cantilevers. This discovery has led us to a simpler processing with a higher yield. Specifically the delineation of the top electrode and the etch channel can be performed with one mask: this allows us to save one troublesome step of photoresist strip after the cantilever release. Also we do not have to reserve some margin for protecting Al during cantilever release, and the surface coverage of the reflective surface can be increased.

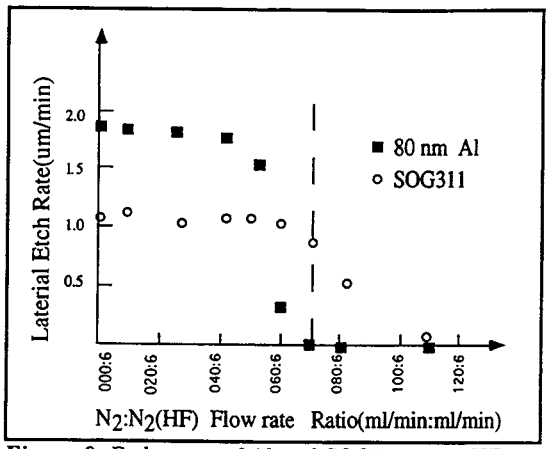


Figure 9. Etch rates of Al and SOG in vapor HF.

EXPERIMENTAL RESULTS AND DISCUSSION

The tip displacement of cantilever by an electrical input is measured with a laser interferometer. In order to distinguish piezoelectric response from thermal response, an electrical input of an unbiased square wave is applied to the cantilever, and a displacement response of a square-wave form is obtained from 0.1 Hz to as high as 25 kHz. The cantilever tips are observed to deflect about 0.19 μm/V vertically in piezoelectric response to a $10V_{p\text{-}to\text{-}p}$ square wave of 10 Hz. Due to the relative high resistance of each cantilever element (above 100 MΩ), the thermal deflection is about 2 or 3 orders of magnitude smaller than the piezoelectrical response according to our theoretical calculations and experimental measurements. Therefore there is no thermal-enhanced deflection in our cantilevers, unlike those reported earlier [6]. A piezoelectric response of cantilever to a $10V_{zero\text{-}to\text{-}peak}$ square wave input at 50Hz is shown in Figure 10.

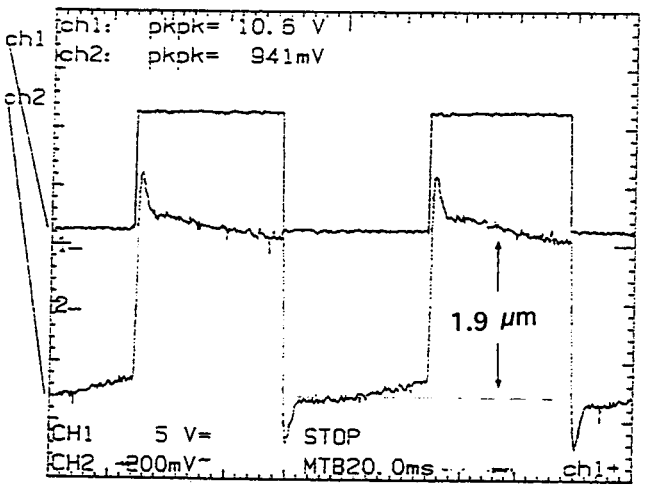


Figure 10. The cantilever-tip displacement (CH2) for a 50 Hz, 10 $V_{zero\text{-}to\text{-}peak}$ square wave input (CH1).

Due to the relative narrow air gap underneath the cantilever, the air damping can reduce the vibrational amplitude of the cantilever significantly. For a cantilever with about 1 μm air gap, the vibrational amplitude in the upward direction is observed to be almost twice larger than that in the downward direction. Indeed the frequency response of the cantilevers is dominated by the air gap distance which is determined by the thickness of the sacrificial layer. Typical frequency responses to a square wave with a 5V amplitude are shown in Figs. 11 and 12 for the cantilevers with the air gap of about 0.5 μm and 1 μm, respectively. If the cantilevers are initially deflected about 30°, the roll-off frequency can be as high as 10 kHz as shown in Fig. 13. Thus, the thickness of the sacrificial layer can be a design parameter to control the response time of the cantilever deflection.

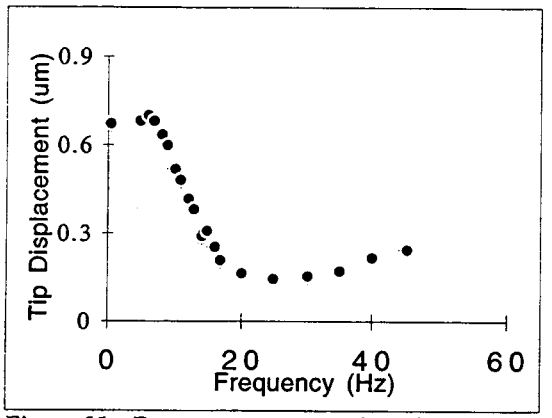


Figure 11. Frequency response of tip displacement of the cantilever with a 0.5 μm high air gap.

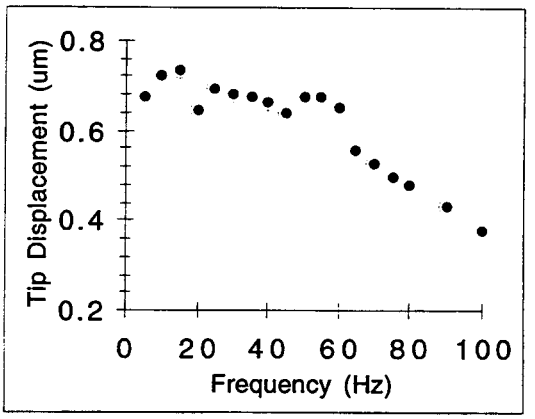


Figure 12. Frequency response of tip displacement of the cantilever with a 1 μm high air gap.

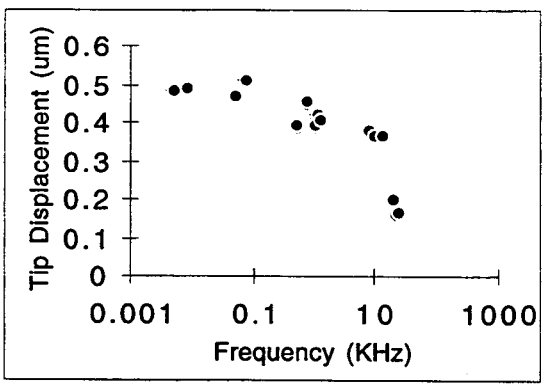


Figure 13. Frequency response of tip displacement of the cantilever that is initially bent up about 30°.

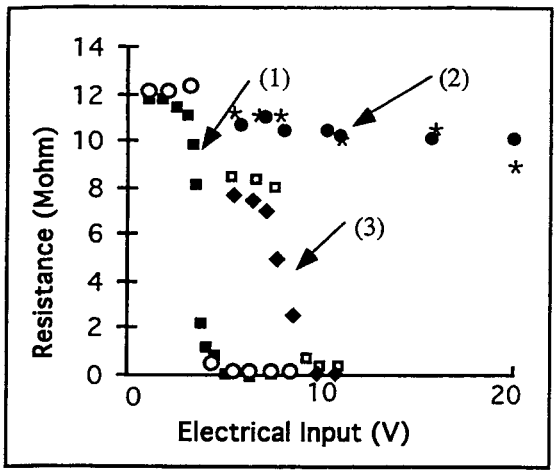


Figure 14. The resistance of ZnO and thin SOG on a row of 40 cantilevers as a function of the magnitude of a 8 Hz square wave.

The cantilever has been observed to deflect linearly in the range of 1 - 10 V applied voltage.

The resistivity of ZnO film in the cantilever has been observed to change due to moisture as confirmed by the data shown in Fig. 14. The data marked with (1) in Fig. 14 are for the devices that have gone through the vapor HF release step, and show sharp drops of the ZnO resistance when the magnitude of the electrical input (8 Hz square wave) is higher than 3 $V_{peak\text{-}to\text{-}peak}$. With a post-processing baking of the devices at 80 °C for 2 hours, the resistance drop is non existing even with more than 20 $V_{p\text{-}to\text{-}p}$ input, as can be seen in the data marked (2) in Fig. 14. The devices exposed to ambient moisture for about one week shows their ZnO resistance dropping sharply with 8 $V_{p\text{-}to\text{-}p}$ input (see (3) in Fig. 15). Unless the ZnO is permanently damaged by too much current, the resistance drops due to moisture can be reversed with baking at 80 °C.

CONCLUSION

We have demonstrated feasibility of piezoelectric ZnO-actuated cantilevers for a projection display application. A complete fabrication processing for the cantilever arrays has successfully been developed, and used to fabricate working arrays of cantilevers (each array consisting of 40 x 40 cantilevers). Relatively large piezoelectric responses with good linearity have been observed.

ACKNOWLEDGMENT

The authors wish to thank You Kwang Kim for profiler measurements made at Daewoo Electronics and numerous helps. They also wish to thank Drs. Yun Woo Nam and Jeong Beom Ji of Daewoo Electronics, and Yonghua Song of UH for helpful discussion.

REFERENCE

1. J.M. Younse, "Mirrors on a chip," *IEEE Spectrum*, pp. 27 - 31, November 1993.
2. R.B. Apte, F.S.A. Sandejas, W.C. Banyai and D.M. Bloom, "Deformable Grating Light Valves for High Resolution Displays," *Technical Digest of the 1994 Solid-State Sensor and Actuator Workshop*, Hilton Head Island, SC, June 13 - 16, 1994, Transducer Research Foundation, Cleveland (1994), pp. 1 - 6.
3. G. Um, D. Foley, A. Szilagyi, J. B. Ji, Y. B. Jeon and Y. K. Kim, "S7-6 Recent Advances in Actuated Mirror Array (AMA) Project Development," *ASIA DISPLAY* (1995), pp. 95 - 98
4. R.P. Ried, E.S. Kim, D.M. Hong and R.S. Muller, "Piezoelectric Microphone with On-Chip CMOS Circuits," *IEEE/ASME Journal of Microelectromechanical Systems*, vol. 2, pp. 111-120, September 1993.
5. F.R. Shanley, *Strength of Materials*, pp. 312-316. New York: McGraw-Hill, 1957.
6. J. G. Smits and W. S. Choi, "Very large Deflection with Quadratic Voltage Dependence of ZnO on Si3N4 Bimorph," *IEEE Transaction on Ultrasonics, Ferroelectrics, Frequency-control*, vol. 39, no. 2, pp. 302-304, March 1992.

TUNABLE IR FILTERS WITH INTEGRAL ELECTROMAGNETIC ACTUATORS

T.R. Ohnstein, J.D. Zook, and H.B. French
Honeywell Inc., Honeywell Technology Center
Plymouth, Minnesota 55441

H. Guckel, T. Earles, J. Klein, and P. Mangat
University of Wisconsin Center for Applied Microelectronics
Madison, Wisconsin 53706

ABSTRACT

Tunable IR filters with integral electromagnetic drive actuators have been fabricated using the MEMS LIGA process at the University of Wisconsin. Two types of actuators with IR filter loads have been demonstrated: a single-pole solenoid-type actuator and a three-phase electromagnetic stepper motor drive. The IR filters were designed for a tunable range from 8 to 32 μm. The filter and actuator structures are fabricated from electroplated permalloy. Measurements of the magnetic properties of the permalloy show it to be a soft magnetic material with low coercive force (0.3 Oe) and high permeability (~65,000). An IR filter structure with a large displacement stepper motor has been demonstrated. The bi-directional stepper motor had a total travel of 1.7 mm with an average force per step of 2.4 mN (basic step size of 25 μm) at a drive current of 45 mA per phase. This is the first demonstration of a MEMS linear electromagnetic stepping motor.

INTRODUCTION

Honeywell and the University of Wisconsin are developing microelectromechanical tunable IR filters for IR spectral analysis. The filters and integral actuators are fabricated by the MEMS LIGA process for making high-aspect-ratio metallic microstructures [1]. These devices will be the key elements in compact gas analyzers for applications in industrial process control and environmental monitoring. The tunable IR filter (Figure 1) is a mechanically deformable structure consisting of an array of parallel metal plates joined by spring flexures and driven by an integral linear actuator [2]. The plates form an array of identical waveguides whose transmittance properties can be varied by changing the spacing between them (Figures 2 and 3). The filter concept has been verified theoretically and demonstrated with IR filters fabricated with fixed dimension filter plate spacings [2,3]. The present paper describes the design, fabrication, and performance of the integral actuators used for tuning the filters, including the first working MEMS electromagnetic stepping motors (Figure 4).

ACTUATOR DESIGN

Tunable IR filters with two types of integral electromagnetic actuators have now been fabricated and demonstrated with tunable filter loads. Solenoid-type linear actuators, which have been used to drive spring-loaded microstructures and reported previously [1,2], have now been integrated with a tunable filter load. An electromagnetic stepper motor drive with an integral tunable filter has also been demonstrated with greater travel for filter tuning range and improved position control over the single-pole, solenoid-type actuator. The overall size of the filter and actuator structures ranges from 6 x 6 mm to 6 x 12 mm (Figure 4). The active area of the filters ranges from 1 x 1 mm to 2.5 x 2.5 mm (Figure 4). Filters have been fabricated with a wavelength cutoff range from

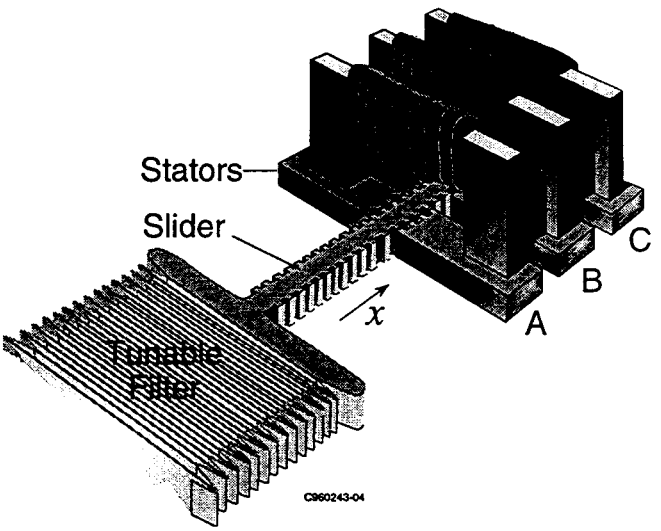


Figure 1. Tunable IR filter with integral stepping motor. The slider and tunable filter are electroplated from the same LIGA mask and assembled onto the stator and mounting posts fixed to the substrate. The electromagnets are assembled into the stators after the coils are wound.

8 to 32 μm. The filter structures with integral slider and spring supports are electro-plated onto a sacrificial layer, released, and assembled onto a substrate having magnetic pole faces and attachment posts (Figure 4). The mandrels for the electromagnets are plated separately, coils wound with copper magnet wire, and mounted with teeth that fit precisely into slots in the stators as shown in Figure 1. All the filter and actuator structures are fabricated using the MEMS LIGA process and are made of electroplated permalloy (approximately 78% Ni, 22% Fe).

The MEMS stepping motors are based on rotary switched-reluctance (SR) stepping motors that have well-known design principles [4,5,6]. In a linear motor, the rotor is replaced by a toothed slider that is coplanar with three identically toothed stators. In Figure 1, the teeth of stator A are aligned with the slider teeth, and stators B and C are displaced by one-third of a tooth period. The aligned position is the stable equilibrium position, meaning there is no force when A is energized. This is also the maximum inductance position for coil A. When the coil B current is also turned on, stator B pulls the slider to the right. When current A is then turned off the slider teeth align with stator B. Thus sequential currents in coils A, B, and C drive the slider to the right by one tooth period for each drive cycle. Maximum slider displacement occurs when the magnetic holding force of one stator equals the restoring force of the support springs and filter load.

0-9640024-1-8 $3.00 © 1996 TRF

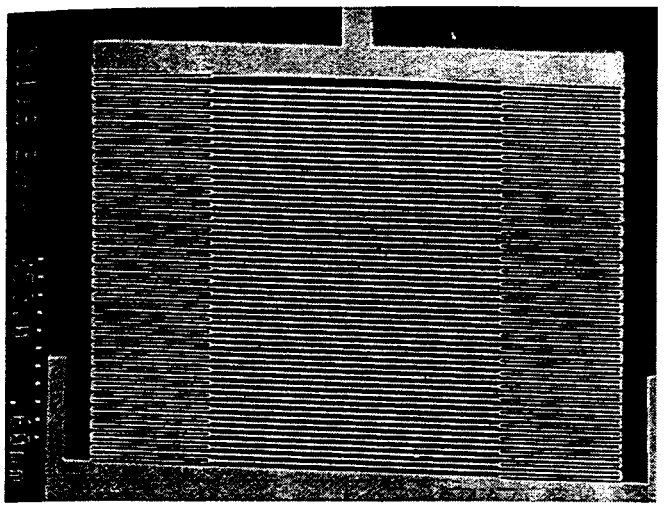


Figure 2. SEM photograph of tunable IR filter. The active filter area is in the center section with supporting spring flexures at each side. The active area is 1 by 1mm (with filter in fully compressed position).

The stepper motor designs have 30-μm teeth and 45-μm spaces, with 10 teeth per stator and a basic step size of 25 μm per phase. By energizing pairs of coils, the step size can be reduced to 12.5 μm.

The chosen tooth shape is based on large machines whose design is well established on the basis of modeling as well as empirical design rules [7]. The gaps between the slider and stator were nominally 2 μm, also scaled down from the large machines. Magnetic finite element modeling (FEM) was used to verify the baseline design and provide a basis of comparison with experiment. The two-dimensional model included two teeth from the stator, four from the slider, and a flux closure path. Magnetic fields were calculated for 16 different positions of the slider with respect to the stator, from the completely aligned case to the completely misaligned case. The force is equal to the change in stored magnetic field energy between positions, divided by the displacement.

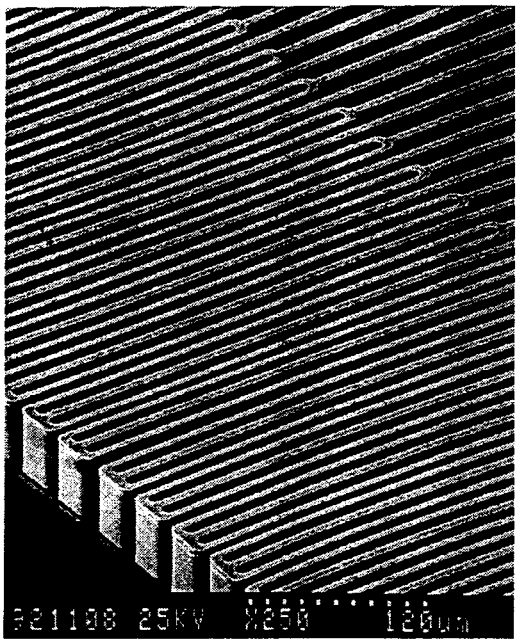


Figure 3. Enlarged view of filter spring flexures that join the filter plates at the sides of the active filter area.

In the baseline model for the stator and slider the nonlinear B-H relationship used was based on preliminary measurements of the permalloy properties. The B-H curve had a relative permeability of 2700 and a saturation magnetization B_S of 1T.

Figure 5 shows the calculated force for six current (I) drive levels, from 1 to 32 ampere-turns in binary steps, for an electroplated permalloy thickness of 100 μm. The force is zero from symmetry considerations at both the aligned (stable) and misaligned (unstable) positions. For a linear material, each curve would have the same shape and an amplitude that scales as I^2.

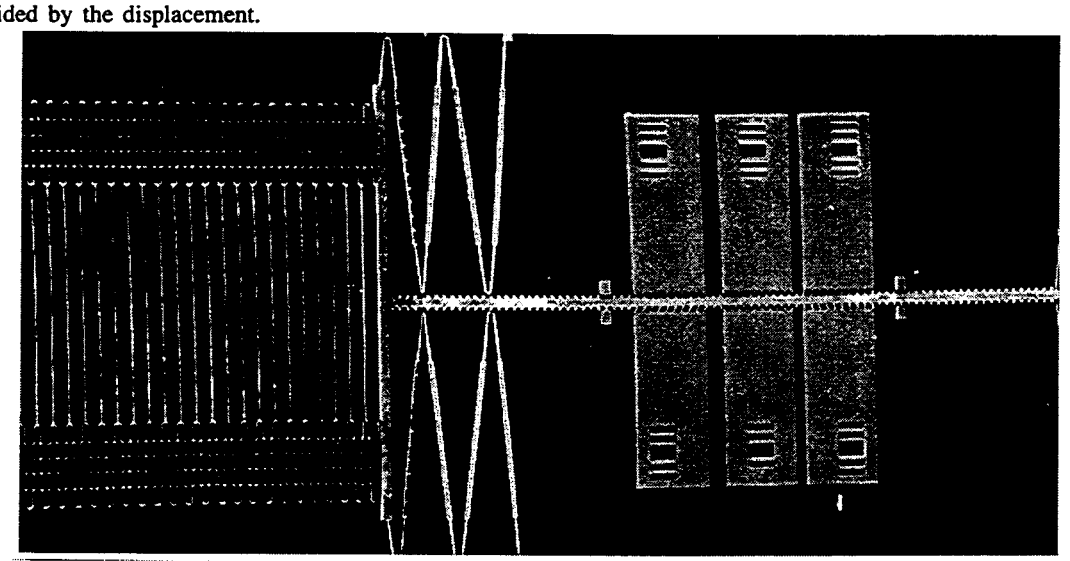


Figure 4. SEM photograph of an assembled IR tunable filter with a three-phase electromagnetic stepper motor drive. The filter, slider, and suspension spring structure is fabricated separately from the fixed substrate structures, mounting posts, and electromagnetic drive pole pieces. The drive coils (not shown) assembled in this photograph are mounted onto the pole pieces.

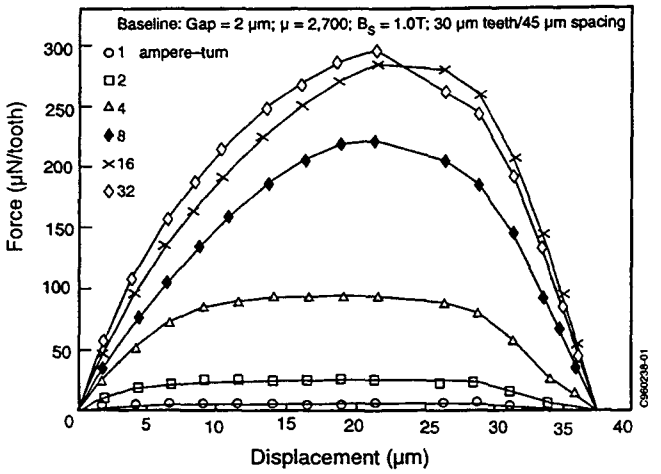


Figure 5. FEM calculation of the magnetic force as a function of slider position for six different current levels. The force is calculated from the change in magnetic field energy between two adjacent positions.

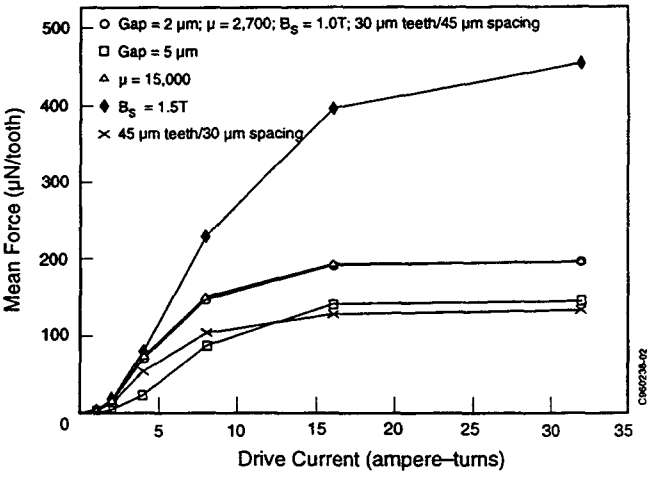


Figure 6. Calculated average force as a function of current for various changes in the assumptions of the model.

The actual curves change shape and bunch together with increasing current because of magnetic saturation. The maximum force that can be produced with this design and material parameters is about 300 μN/tooth. The mean force at high drive levels is about 200 μN/tooth or 4 mN/stator (for 20 teeth per stator).

Figure 6 illustrates the effects of changes in material properties and dimensions. The curves show the result of (1) increasing the gap to 5 μm (to reduce material saturation effects); (2) increasing the material permeability to 15000 (either by annealing or composition changes); (3) increasing B_S to 1.5 Teslas (by composition changes); and (4) using 45-μm-wide teeth separated by 30 μm (for easier assembly). Figure 6 shows the calculated mean force per tooth between the aligned and misaligned positions plotted against the current drive level. Increasing the gap or the tooth width gives less force than the baseline case, while increasing the permeability has virtually no effect except at low drive levels (before material saturation sets in), and increasing B_S increases the limiting force as B_S^2. B_S would increase with iron content in the alloy [8] at the risk of a tradeoff in mechanical properties of the structures with a change in the alloy.

MATERIAL CHARACTERIZATION

To measure the magnetic properties of the electroplated permalloy, toroid test structures (1.7 mm ID., 2.5 mm OD.) were fabricated along with the filter structures. The toroids were wound with a primary winding of 15 turns and a secondary winding of 30 turns, using 43-AWG wire. A known, time-varying current (proportional to H) is applied to the primary, while measuring the voltage (proportional to dB/dt) across the secondary.

At present the primary is driven with a 10-Hz triangular current waveform that extends into saturation. The very small cross-sectional area of the toroid produces a very small secondary voltage at this frequency, requiring a high-gain, low-noise, low-offset instrumentation amplifier. Integration of the voltage vs. time is done in software. The ascending and descending ramp waveforms are averaged before integrating, and the offset voltage is chosen to make the slope of the B-H curve zero in saturation. The currents and voltages are converted to H and B using the physical dimensions of the toroid, the number of windings, the gain in the electronics, and the drive frequency. Using this procedure the B-H curve shown in Figure 7 was obtained for the permalloy toroid. A maximum permeability of about 65,000, a saturation magnetization of about 9 kG (0.9 Teslas), and a coercivity of about 0.3 Oe were derived from this curve. These values are all of the same magnitude as, but somewhat poorer than, those expected for 78 Permalloy (100,000, 10.8 kG and 0.05 Oe) [8].

The effective Young's modulus and mechanical fatigue properties of permalloy are being measured using the test structure shown in Figure 8. It consists of a horseshoe electromagnet and a cantilever arm that is deflected toward the pole face when the drive coil is energized. The resonant frequency is easily measured with a gain/phase analyzer and is quite consistent from sample to sample at 20,100 Hz, indicating a Young's modulus of 1.5×10^5 MPa, a little lower than the bulk value for permalloy. A circuit is being designed to drive the cantilever continuously to test the fatigue properties of the material as a function of strain amplitude. Indications to date are that electroplated permalloy is an excellent spring material [2].

ACTUATOR PERFORMANCE

Tunable IR filters with two types of integral electromagnetic actuators have now been fabricated: single-pole electromagnetic drives and three-phase electromagnetic stepper motors. The filter design in Figures 2 and 3 has been driven in resonance with a single-pole actuator with a total displacement of 250 μm. The three-phase stepper shown in Figure 4 has been driven the full amplitude (1 mm) in one direction. Displacement in the other direction has been limited so far to 0.7 mm, apparently because of twisting of the flexures that cause them to hang up on the fixed parts. The twisting may be due to small variations in thickness of the permalloy springs.

Preliminary electrical measurements of the average actuator force during a step have been performed following the procedure used for large stepping and SR motors [9]. The inductance of an individual drive coil is measured as a function of dc drive current while the slider is held in a fixed position.

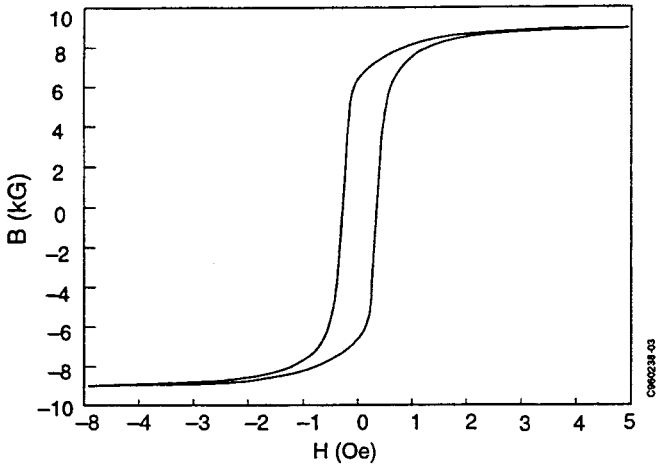


Figure 7. Measured B-H characteristics for an electroplated permalloy toroid. The average of many loops was taken and the voltage and current data were adjusted to make the curve symmetric.

The integral of the inductance with respect to current gives the flux linkage as a function of current. Integrating again with respect to current gives the energy stored in the magnetic field.

The measurements and integrations are performed for the slider in the aligned and unaligned positions for each individual drive coil. The difference in the two energies then gives the average drive force as a function of current (as shown in Figure 6 above) for the model. The procedure gave different force curves for each of the three coils, but the average force per step at the maximum test current of 45 mA was 2.4 mN. This value is in close agreement with the FEM analysis when the number of teeth and material thickness are considered. As fabrication imperfections in the stepping motor are eliminated, its output force will approach the design value of 4 mN/step and be more than adequate to drive the tunable filters, which are designed to have force constants of 1 mN/mm or less.

CONCLUSIONS

Tunable IR filter structures with integral electromagnetic drive actuators have been fabricated and demonstrated. A single pole, solenoid-type actuator and a three-phase electromagnetic stepper motor drive have been demonstrated with tunable IR filter loads. The IR filters were designed for a tunable range from 8 to 32 μm. The filter and actuator structures are fabricated with the MEMS LIGA process using electroplated permalloy. Measurements of the magnetic properties of the permalloy show it to be a soft magnetic material with low coercive force (0.3 Oe) and high permeability (~65,000). An IR filter with a large travel electromagnetic stepper motor drive was demonstrated. The bi-directional stepper motor had a total travel range of 1.7 mm with an average force per step of 2.4 mN at a drive current of 45mA. A characterization of the transmission properties of the tunable IR filters has started as the next step in the development of the filters.

ACKNOWLEDGMENTS

This work was partially supported by ARPA Contract DABT63-93-C-0066. The motor drive electronics and demonstration were done by Ernie Satren. Material characterizations were done by Nick Demma. Additional processing of the structures was performed at Honeywell by Dawn Murphy.

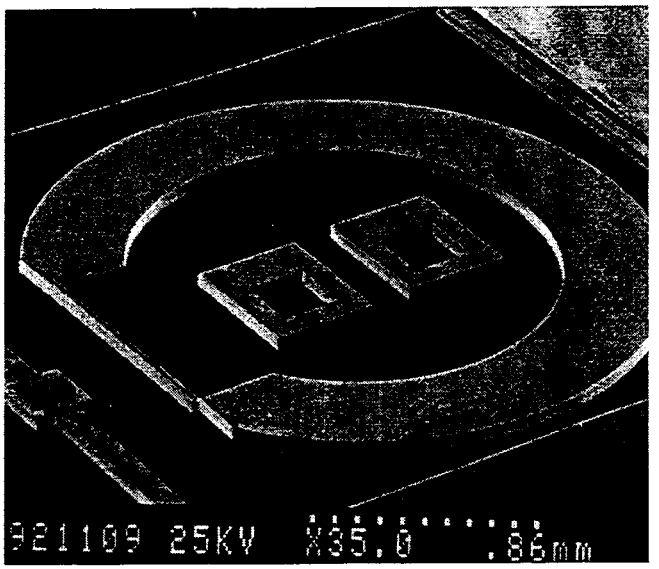


Figure 8. Photograph of unreleased electroplated parts including a structure for material characterization. The horseshoe magnet is wound with drive and pickup coils. The cantilever beam is driven in resonance. The beam is 1-mm long by 30-μm wide by 80-μm tall. There is a 2-μm gap between the end of the beam and the magnet.

REFERENCES

1. H. Guckel, et. al., "Electromagnetic Linear Actuators With Inductive Position Sensing For Micro Relay, Microvalve And Precision Positioning Applications," *Technical Digest of Transducers '95, The 8th International Conference on Solid-State Sensors and Actuators*, Stockholm, Sweden, 6/25-29/95, pp. 324-327.

2. T. R. Ohnstein, et. al., "Tunable IR Filters Using Flexible Metallic Microstructures," *Proc. 1995 Micro-Electro-Mechanical Systems*, Amsterdam, The Netherlands, 1/29-2/2/95, pp. 170-174.

3. J. A. Cox, et. al., "Optical Performance of High-Aspect LIGA Gratings," *Proc. Micro-optics and Micromechanics Conf.*, SPIE Vol. 2383, San Jose. CA (Feb, 1995), pp. 17.

4. "Theory and Application of Step Motors," Ed. by B.C. Kuo, West Publishing Co., St. Paul (1974).

5. "Proc. Int. Conf. on Stepping Motors and Systems 19-20 Sept, 1979," Ed. by A. Hughes, University of Leeds (1980).

6. "Switched Reluctance Motor Drives—A reference book of collected papers," Ed. by T.J.E. Miller, Intertec Communications, Ventura, CA (1988).

7. T.J.E. Miller, "Switched Reluctance Drives," [6] above, pp.3-69, also M.R. Harris, et.al., "Limitations of Reluctance Torque in Doubly-Salient Structures," [6] above, pp 160-170.

8. D. Jiles, "Introduction to Magnetism and Magnetic Materials," Chapman and Hall, New York (1991).

9. G. Singh, "Mathematical modeling of step motors," [4] above, pp. 33-75.

A CHRONOLOGY OF THERMAL INK-JET STRUCTURES

Christopher C. Beatty

Hewlett-Packard Ink-jet Supplies Business Unit

Corvallis, Oregon 97330

ABSTRACT

Over the past dozen years, thermal inkjet (TIJ) pens have become one of the largest businesses based on micromachining technology. The basic functioning of an inkjet pen is discussed and four generations of mircostructures from the leading manufacturer are examined. Future trends are analyzed including nozzle count and drop volume.

INTRODUCTION

Before discussing the evolution of the TIJ microstructures, a discussion of the basics of the firing mechanism common to all of them is in order. A two-dimensional view of the firing sequence can be seen in Figure 1. These plots come from computational fluid dynamics simulations of the original TIJ architecture.

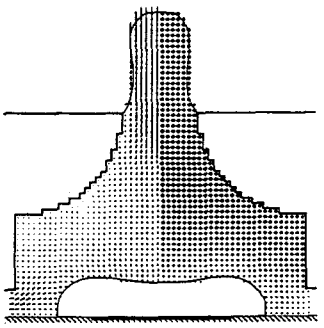


Figure 1a. TIJ firing sequence at five microseconds

Firing begins with an electrical pulse sent to a thin-film heater at the base of a chamber filled with ink. The pulse is typically only a few microseconds long, but superheats a thin boundary layer of ink to approximately 250°C.

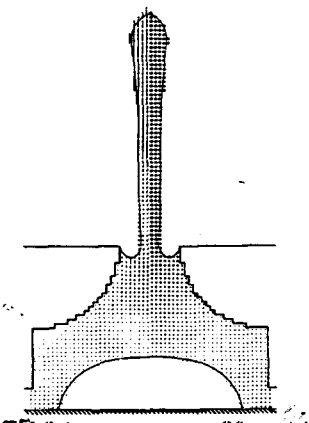


Figure 1b. TIJ firing sequence at fifteen microseconds

This causes homogeneous or nearly homogeneous boiling across the surface of the heater. This bubble is referred to as a "drive bubble" since it pushes the ink out of the chamber. After five microseconds, the pulse is complete and the bubble has partially formed and has begun to expel ink out of the orifice. Some ink is also pushed out of the refill channels, but geometries are constructed to minimize this. At fifteen microseconds (figure 1b), the bubble has fully formed and the droplet is being ejected. The tail of the drop has not yet separated. Figure 1c shows the collapse of the bubble and the beginning of the refill process at 24 microseconds. The collapse of the bubble cause a tremendous pressure pulse on the thin films due to the implosion or cavitation of the bubble. Depending on the geometry of the chamber, an additional 25-50 microseconds are necessary to let the chamber refill and the orifice meniscus settle into a stable position.

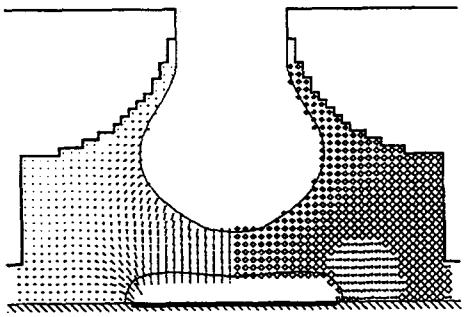


Figure 1c. TIJ firing sequence at twenty-four microseconds.

Figure 2 shows the effect of changing the firing frequency on an inkjet pen. For low frequencies, the firing is completed and the ink has settled back to the equilibrium position for the meniscus. As the frequency is raised, the drop volume decreases because the meniscus is below the equilibrium position.

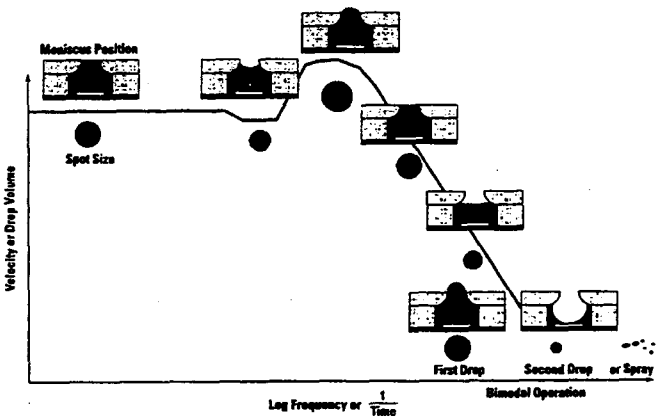


Figure 2. Drop weight versus frequency curve

0-9640024-1-8 $3.00 © 1996 TRF

Solid-State Sensor and Actuator Workshop
Hilton Head, South Carolina, June 2-6, 1996

Further increase in the frequency actually raises the drop volume because the firing is occurring while the meniscus is bulging outward. Frequencies above this value decrease the drop volume as there is more encroachment into the refill process.

TIJ 1.0

The first drop-on-demand thermal inkjet pen was successfully commercialized in 1984. It required the merging of several technologies including thin film processing, electroforming and fluidic interconnection.

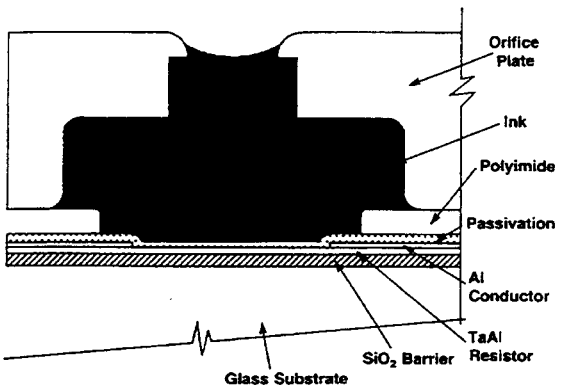


Figure 3. TIJ 1.0 thin-film cross-section.

A cross-section of the TIJ 1.0 firing chamber can be seen in figure 3. The glass substrate was chosen because it was easily available from another LED fabrication group at the same site and it was thought that silicon would be too expensive. In retrospect, the wide use of silicon substrates and the optimization of equipment for them outweighed the initial cost difference. Despite an insulating substrate, a silicon dioxide layer was sputtered on the glass plate to avoid diffusion of sodium impurities into the metal films which reduced reliability. The resistor heater which is pulsed to fire the pen is made from a sputtered tantalum-aluminum alloy (TaAl). This binary mixture has the unique property of having a nearly zero temperature coefficient of resistance (TCR). This is an important feature to ensure uniform energy delivery during the rapid heating to several hundred degrees Centigrade during a firing event. The resistors are connected to bonding pads with a conventional aluminum alloy. The bonding pads have a thin nickel diffusion barrier and a gold bonding surface.

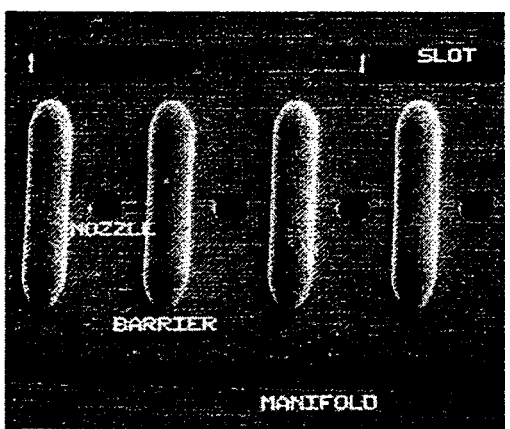


Figure 4. SEM photograph of a TIJ 1.0 orifice plate

The dielectric passivation layer over the heater area protects the metals from the corrosive mixture of ink and high temperature. Silicon carbide provides excellent chemical resistance in this environment, but as a semiconductor it provides an electrical leakage path at higher temperatures. The solution was a dual film dielectric of silicon nitride for electrical isolation and silicon carbide for chemical inertness. The films are deposited sequentially in a plasma-enhanced chemical vapor deposition system (PECVD).

Once the thin film structure is completed, the electroformed nickel orifice plate is attached with a polyimide layer. The orifice plate not only defines the bore for the ejected drop, it also creates barriers between adjacent nozzles to reduce hydraulic crosstalk. A SEM photograph of the TIJ 1.0 orifice plate can be seen in figure 4.

A typical alphanumeric character produced by the 96 dot per inch (dpi) ThinkJet printer can be seen in figure 5.

Figure 5. 96 dpi TIJ 1.0 character.

TIJ 2.0

The first entry into the color inkjet market came in 1987. An enhanced monochrome pen was also introduced. This new family was called TIJ 2.0 and several innovations were present.

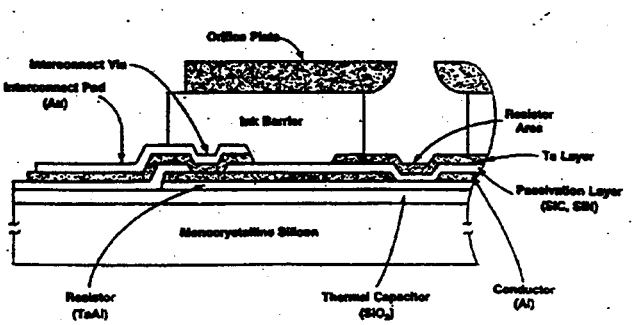


Figure 6. TIJ 2.0 thin-film cross section

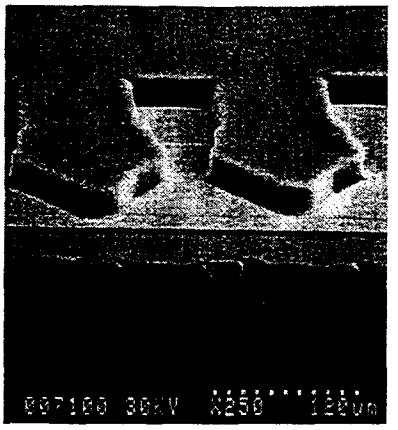


Figure 7. TIJ 2.0 firing chamber without orifice plate.

In the thin film arena, the substrate is a silicon wafer to take advantage of the equipment set and processing knowledge available. A layer of tantalum is present on top of the passivation to withstand the impact of the cavitation when the drive bubble collapses. The toughness of the tantalum provided more than an order of magnitude improvement in the number of cycles the printhead can withstand. Firing chambers can now be shaped into more sophisticated geometries through the use of a photo-imageable polymer layer (figure 7). The nickel orifice plate is still used, but it defines only the bore instead of the entire chamber.

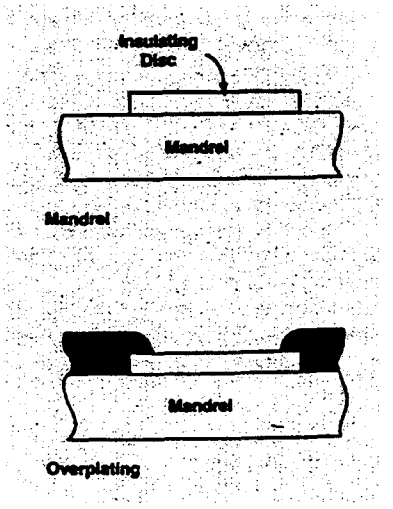


Figure 8. Orifice plate fabrication.

The process of creating the orifice plate is also unique in that it involves electroforming nickel on a stainless steel substrate called a mandrel that has patterned PECVD silicon nitride "buttons". The converging nozzle shape is created by the isotropic growth of the nickel film while the exit bore is determined by the amount of overplating onto the top of the nitride disc (figure 8).

As in the TIJ 1.0 process, ink feeds through the chip via a slot created by an abrasive sand drill. The rough edge created by the drilling process can be seen in the foreground of figure 7. Print quality improves dramatically when 300 dpi architectures are employed (figure 9).

Figure 9. 300 dpi TIJ 2.0 character

TIJ 2.5

As the number and density of nozzles rose, it became increasingly difficult to interconnect the pen to the printer. Previous generations required a separate pad for each firing resistor. When this was no longer practical, a three mask NMOS process was integrated into the printhead (figure 10 on next page).A row and column addressing scheme is used to select the resistor to be fired. This reduces the pad count by a factor of one over the square root of the number of pads needed in the direct drive scenario (figure 11).

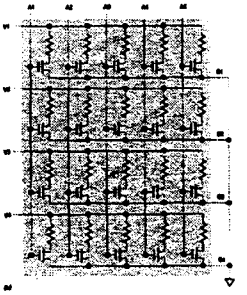


Figure 11. Row and column matrix addressing for TIJ 2.5

While much of the thin film set is unchanged, evolution of the firing chamber architecture is taking place. Smaller, better defined drops emerge from TIJ 2.5 pens than their predecessors.

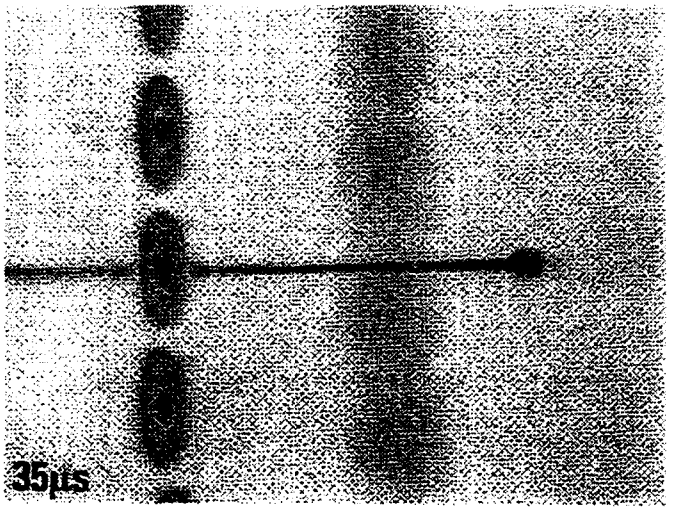


Figure 12. TIJ 2.5 drop ejection at thirty-five microseconds.

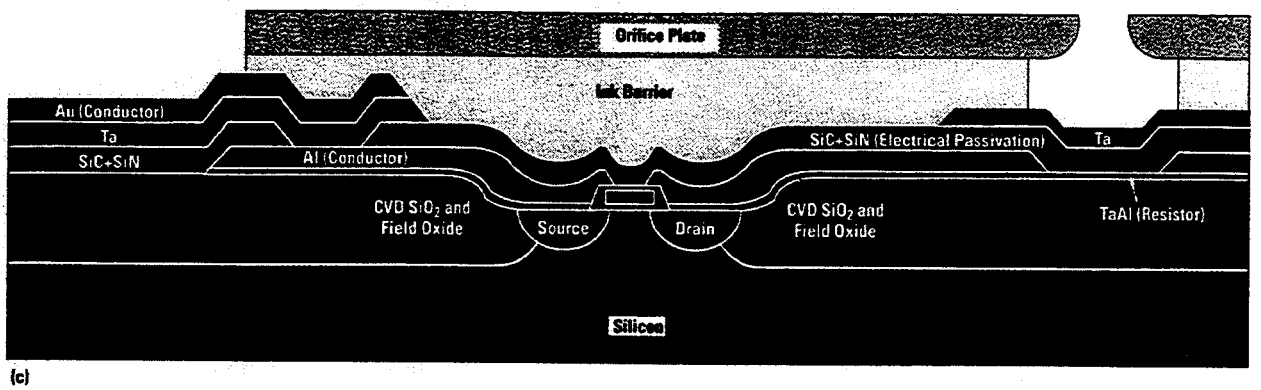


Figure 10. TIJ 2.5 thin-film cross section.

TIJ 3.0

The current generation of TIJ printheads utilizes TIJ 2.5 silicon but substitutes a laser ablated orifice in the flex circuit for the electroformed orifice. Laser ablation enables arrays of high density, high aspect ratio converging nozzles. For the first time, ink feeds over the sawed edge of the chip rather than through a drilled slot. This conserves expensive silicon die area.

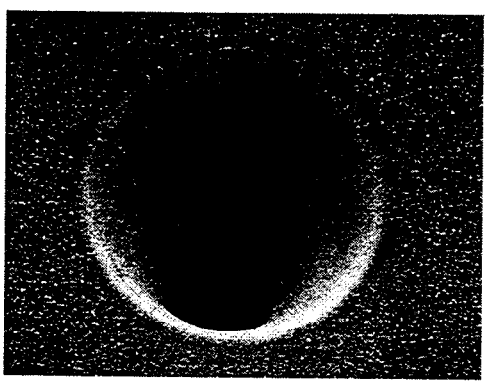


Figure 13. Laser ablated TIJ 3.0 orifice.

FUTURE TRENDS

The thermal inkjet market continues to grow at an astounding rate as innovations are brought into manufacturing. It is interesting to project the path into the future for key metrics of inkjet performance. Two notable parameters are the drop weight and the total nozzles per pen (figures 14 and 15). Drop weight or volume is important because the focus of development is increasingly on image quality as opposed to print quality. Inkjet technology is beginning to move beyond text applications into photographic quality images. The threshold line on the graph represents the minimum drop that the human eye can perceive. While this line varies depending interactions of the ink, media, and printer, it is clear that inkjet is rapidly approaching this regime. For drop volumes below this line, photographic quality images are possible since the observer will not perceive individual dots. Nozzle density goes hand-in-hand with advancements in volume since the print engine must be able to place the smaller drops at the appropriate intervals. This metric is also on course in the advancement of the technology rising from 96 dpi in TIJ 1.0 to 600 dpi in the TIJ 3.0 family.

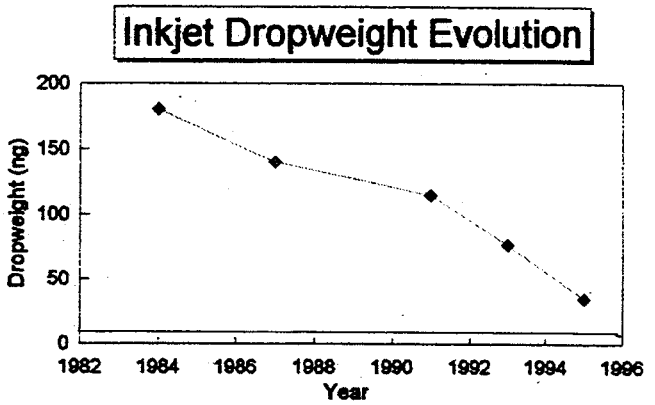


Figure 14. Inkjet drop weight versus time.

Total nozzle count has a pragmatic justification – as the drop volume decreases, the number and/or rate of firings must be increased to maintain the throughput of the printer (pages/minute). The rise in both nozzle count and firing frequency has enabled higher throughput despite the decreasing drop volume.

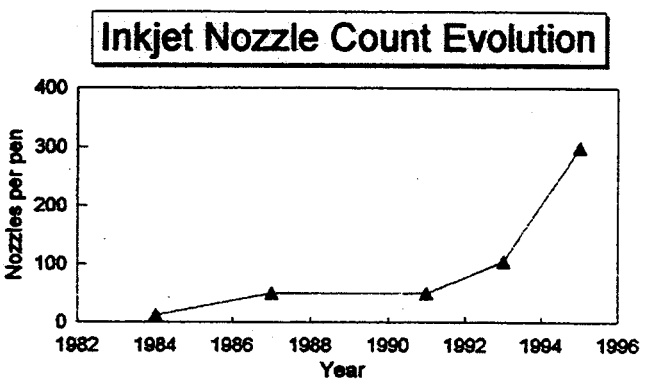


Figure 15. Inkjet nozzle count versus time.

CONCLUSIONS

While impressive progress has been made in thermal inkjet technology, the future is brighter still (figure 16). The competitive marketplace and wide acceptance of inkjet printing is driving innovation on many fronts. Ever higher levels of functionality, performance and speed are the expectations of a growing customer base.

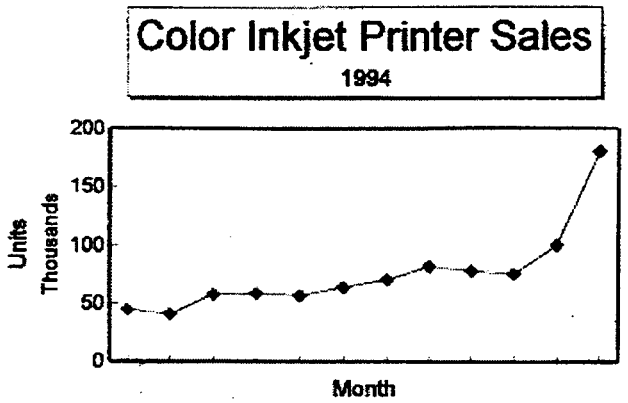


Figure 16. Color Inkjet sales for 1994.

REFERENCES

1) Allen, Ross R., Meyer, John D., Knight, William R., "Thermodynamics and Hydrodynamics of Thermal Ink Jets", Hewlett-Packard Journal, May 1985, pages 21- 27

2) Bhaskar, Eldurkar V., Aden, J. Stephen, "Development of the Thin-Film Structure of the ThinkJet Printhead", Hewlett-Packard Journal, May 1985, pages 27-33

3) Siewell, Gary L., Boucher, William R., McClelland, Paul H., "The ThinkJet Orifice Plate: A Part With Many Functions", Hewlett-Packard Journal, May 1985, pages 33-37

4) Askeland, Ronald A., Childers, Winthrop D., Sperry, William R., "The Second-Generation Thermal InkJet Structure", Hewlett-Packard Journal, August 1988, pages 28-31

5) Buskirk, William A., Hackleman, David E., Hall, Stanley T., Kanarek, Paula H., Low, Robert N., Treuba, Kenneth E.., Van de Poll, Richard R., "Development of a High-Resolution Thermal Inkjet Printhead", Hewlett-Packard Journal, October 1988, pages 55-61.

6) Bohórquez, Jaime H., Canfield, Brian P., Courian, Kenneth J., Drogo, F., Hall, Corrina A.E., Holstun, Clayton L., Scandalls, Aneesa R., Shepard, Michele E., " Laser-Compatible Inkjet Text Printing", Hewlett- Packard Journal, February 1994, pages 9-17

7) Aden, J. Stephen, Bohórquez, Jaime H., Collins, Douglas M., Crook, M. Douglas, Garcia, André, Hess, Ulrich E., "The Third Generation HP Thermal Inkjet Printhead", Hewlett-Packard Journal, February 1994, pages 41-45.

WIRELESS INTEGRATED MICROSENSORS*

K. Bult, A. Burstein, D. Chang, M. Dong, M. Fielding, E. Kruglick, J. Ho, F. Lin, T. H. Lin, W. J. Kaiser, R. Mukai, P. Nelson, F. L. Newburg, K. S. J. Pister, G. Pottie, H. Sanchez, O. M Stafsudd, K. B. Tan, C. M. Ward, G. Yung, S. Xue
Electrical Engineering Department, University of California, Los Angeles
Los Angeles, CA 90095-1594

H. Marcy and J. Yao
Rockwell Science Center
Thousand Oaks, California

ABSTRACT

Distributed measurement and control capability has been a primary goal for military and civil transportation, manufacturing, biomedical, environmental management, safety and security systems. Microsensor networks, coupled with the high performance of compact computing systems, can provide powerful monitoring and control capability. However, the cost, volume, physical limitations, and installation requirements for cable interfaces has seriously hindered the development of distributed sensor networks with conventional technology. Specifically, while the cost of microsensors may be low, the infrastructure complexity for field deployment or incorporation of sensors dominates the cost of a distributed sensor system. A wireless microsensor network offers the capability of rapid, low cost deployment and installation. Advances in *integration* are required to combine microsensors, analog mixed-signal systems, processors, and wireless communication. This paper presents new methods that provide diverse sensor capability, with new low power solutions for wireless sensor networks.

INTRODUCTION

New product opportunities and new system capabilities are enabled by the development of a low cost distributed Microelectromechanical Systems (MEMS) technology. The development and deployment of distributed monitoring and controls has been hindered in the past by the requirements of complex installation and communication network requirements. Conventional distributed sensors have required cable interface, and therefore, extensive modification to structures and capitol equipment for their installation. In contrast, a wireless microsensor network may be distributed rapidly and without modification to large structures and systems. The wireless network architecture allows LWIM nodes to be rapidly deployed in a broad spectrum of military applications ranging from battlefield perimeter security and shoreline reconnaissance to personnel health monitoring. Wireless sensors may also be applied in areas where volume and mass constraints limit the application of conventional wireline interface sensors. In particular, LWIM nodes may be applied to rotating machinery for condition based maintenance without the complex slip-ring systems that would normally be required for a sensor electrical interface.

A set of unique requirements exist for distributed wireless microsensor networks. The individual low cost sensor nodes must be 1) reconfigurable by their base station, 2) autonomous to permit local control of operation and power management, 3) self-monitoring for reliability, 4) power efficient for long term operation, and 5) must incorporate diverse sensor capability with highly capable microelectronics.

LWIM intelligent node technology, based on commercial, low cost CMOS fabrication and bulk micromachining, has demonstrated capability for multiple sensors, electronic interfaces, control, and communication on a single device. LWIM nodes are fabricated by the new CMOS Integrated MicroSystems (CIMS) process. CIMS provides high sensitivity devices for vibration, acoustic signals, infrared radiation and other diverse signal sources. The central challenges for low cost, manufacturable, LWIM devices are the requirements for micropower operation and the complete integration of a CMOS RF transceiver. The CIMS process, micropower measurement, and micropower RF communication systems are described below.

WIRELESS MICROSENSOR NETWORKS

The wireless microsensor network architecture described here depends on micropower nodes operating with a single, microprocessor base station, (supplied by conventional power sources), and numerous distributed wireless microsensors. Network architecture and communication protocols are developed to exploit the asymmetry of distributed sensor communication. Specifically, most information flow is from the sensor nodes to the base station with substantially less flow in the form of commands to the sensor nodes from the base station.

Typical applications may be optimally serviced by sensor networks having local signal processing by sensor nodes. Thus, individual nodes may propagate measurements of battlefield environment, machine condition, or patient condition, periodically to the base station at low duty cycle. In particular, only upon an alarm condition will continuous data transmission be required. This method permits a base station to service a much larger network than would be possible for simple continuous communication with sensor node. In addition, low duty cycle operation, combined with proper power management, allows low power operation. Periodic updates of the network base station, by distributed network sensor nodes permits detection of changes in environmental or system operation. For example, individual sensor nodes may provide continuous *measurement* of a vibration spectrum, while only *transmitting* the observation of a *change* in this spectrum. By exploiting the low duty cycle requirements for sensor communication, large efficiencies may be obtained in sensor node and base station operation. Remote programmability (enabled by including transmit and receive capability at each node) permits high data rate operation at any time that a request is made by the base station.

Completely independent LWIM nodes must operate at microampere current levels and low voltage. This allows long operating

0-9640024-1-8 $3.00 © 1996 TRF

Solid-State Sensor and Actuator Workshop
Hilton Head, South Carolina, June 2-6, 1996

life from compact battery systems. Alternatively, for some condition based maintenance applications, with nodes mounted directly on a motor or drivetrain shaft, LWIM nodes may receive power by continuous or periodic reception of RF energy from a nearby power source via an inductive coupling. Typical low duty cycle, low data rate (10kbps) and short range (10 - 30m) communication permit 30 μA average current for an LWIM node operating at 3V. A conventional Li coin cell provides this current level for greater than a three-year unattended operating life.

WIRELESS MICROSENSOR FABRICATION

Rapid reductions in the fabrication cost of CMOS digital circuit technology, along with improvements in performance, provide motivation for the development of microsensor structures and measurement circuits in this technology. CMOS technology now conveniently provides the embedded control and micropower digital systems needed for LWIM nodes. Challenges remain for low noise, micropower analog measurement and RF communication systems. Demonstrated micropower CMOS measurement systems [1] for MEMS devices will be described below. CIMS provides devices ranging from inertial to infrared,[2] and acoustic sensors in the same process (Figures 1, 2 and 3). CIMS combines commercial CMOS (post-processed after foundry-fabrication by XeF_2 micromachining)[3] with high performance bulk micromachined sensor and actuator structures (Figure 2 and 3) by flip chip bonding. Finally, micropower CMOS RF circuits (Figure 4) are being developed for the wireless microsensor network.

The CIMS process offers several advances over previous techniques. First, by separating the CMOS and bulk micromachining processes, conventional low cost CMOS technology may be directly applied. This offers the system developer flexibility to update the circuit technology rapidly to exploit the most optimum processes that become available. In addition, the separation of CMOS and sensor element fabrication permits the introduction of novel materials, for example pyroelectric systems,[2] without disturbing critical CMOS processing.

High performance accelerometers are implemented in surface and bulk micromachined structures. Polysilicon, surface micromachined accelerometer proof mass and flexure structures are readily combined with integrated circuit technology. However, the limitations of surface micromachining constrain proof mass thickness and area. Thus, typical surface micromachined structures display high mechanical resonance frequencies and, therefore, low responsivity to acceleration signals. In contrast, bulk micromachined structures provide proof mass values that maybe two orders-of-magnitude larger than those of surface micromachined structures. However, while these structures provide adequate mechanical responsivity, conventional bulk micromachined structures are not conveniently integrated with microelectronics.

Accelerometer structures for high performance applications present an additional unique requirement. High sensitivity acceleration measurement relies on high resolution capacitive detection of proof mass position. This is accomplished by measurement of the capacitance between parallel, planar electrodes mounted on the proof mass surface and on the accelerometer structure. The presence of parasitic capacitance between the measurement electrode and its supporting ground plane, may drastically degrade measurement sensitivity. Thus, Si-on-glass processes have appeared for this application.

The CIMS process offers performance advantages for microsensors. Specifically, the use of bulk micromachined structures enables large proof mass values for inertial and vibration sensors. Single crystal flexures provide stability. Finally, the long standing problem of parasitic capacitance is directly addressed by a unique suspended electrode structure (Figures 1 and 2). The suspended electrode eliminates the need for Si-on-glass for many applications. Also, the proper design of the suspended electrode permits control of squeeze-film damping.

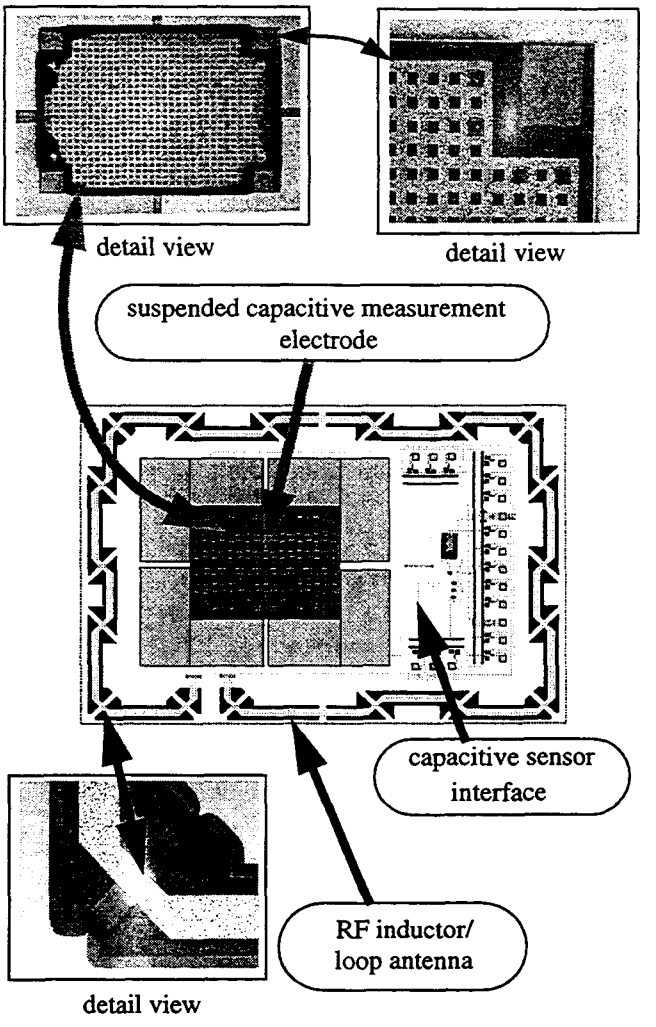


Figure 1. Wireless Microsensor Interface Die. The Interface Die provides sensor measurement, sensor interface, control, and RF communication functions in a single integrated component. The Interface Die, fabricated in commercial 2.0μ CMOS, may be integrated with high performance sensor elements by flip chip bonding. These micrographs display the low parasitic capacitance, suspended measurement electrodes, and suspended inductor / loop antenna. The CMOS die was processed after fabrication by a maskless etch in XeF2

CMOS INTEGRATED MICROSYSTEMS ACCELEROMETER

The acceleration and infrared sensors, implemented by the CIMS process, are shown in Figures 1, 2, and 3. Figure 1 displays the CMOS interface die component of the CIMS process. The interface die may support low parasitic capacitance measurement, data conversion and RF communication capability. The CIMS accelerometer relies on a single crystal flexure supporting a proof mass with a resonance frequency that may be adjusted (by design) from 100Hz to 10kHz. With the established low noise performance of CMOS low power amplifiers ($100nV/(Hz)^{1/2}$) and with capacitive bridge excitation amplitudes of less than 3V, sensor capacitor gap values of 20μ, the 100Hz accelerometer suspension may provide a system sensitivity of less than $1\mu g/(Hz)^{1/2}$.

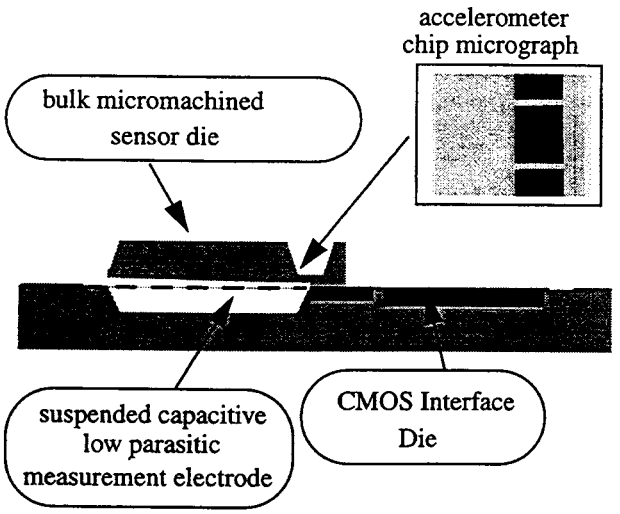


Figure 2. CIMS Accelerometer. The CIMS accelerometer employs capacitive sensing with a suspended measurement electrode. A micromachined cavity drastically reduces substrate parasitics that limit conventional devices. Membrane perforation provide control of squeeze film damping without requirements to perforate the proof mass.

CMOS INTEGRATED MICROSYSTEMS: INFRARED SENSOR

Infrared sensor systems are required in applications ranging from military surveillance, manufacturing process control pyrometry, spectroscopic and gas sensing applications. Wireless integrated infrared sensors address these applications with low cost, rapidly deployed systems. Uncooled, broad band, thermal infrared MEMS sensors have been implemented with bolometric, thermopile, and pyroelectric methods. Pyroelectric structures are of particular interest due to their large responsivity. However, the incompatibility of pyroelectric materials with conventional CMOS processing has hindered the development of this technology.

Single crystals of $LiTaO_3$, bulk ceramics of $PbTiO_3$ (and related compounds), KTN, [4] organic polymers and Mylar have all been used in pyroelectric IR detectors. Ferroelectric $PbTiO_3$ (PTO) and $Pb_{1-x}Ca_xTiO_3$ (PTCX) are attractive materials for this application because of their large pyroelectric coefficient **P**, low dielectric constant value and low specific heat value. This results in optimal responsivity and frequency response characteristics. Now, additional capability is offered by the tunable (composition dependent) Curie temperature of $Pb_{1-x}Ca_xTiO_3$. Curie temperature may be varied continuously from 490° to 0° C as x increases from 0 to 0.5. Thus the proper choice of x provides a pyroelectric film with a Curie temperature slightly above room temperature, placing the maximum in the temperature dependent value of **P** at the desired device operating temperature.

In addition to a large pyroelectric response, a high-sensitivity pyroelectric detector requires both small thermal capacitance and small thermal conductance to maximize the temperature change and signal induced by incident IR radiation. Thin film geometries provide optimal response for pyroelectric sensors by minimizing thermal capacitance. In addition, thermal conductance must be minimized by suspending the thin film structure to provide thermal isolation.

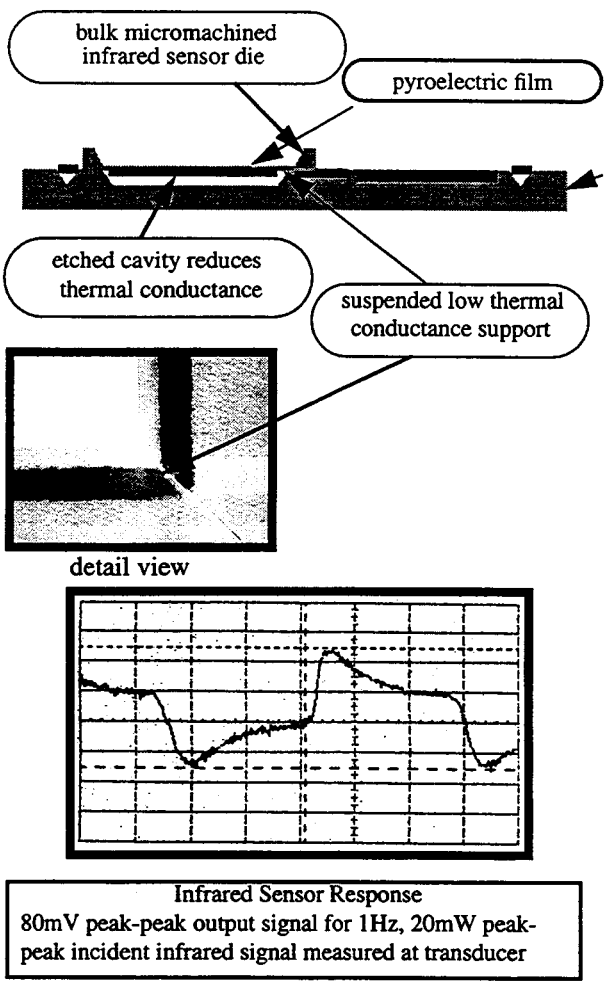


***Figure 3. CMOS Integrated Microsystems Infrared Sensor** The wireless infrared sensor employs a low thermal conductance structure supporting a sol-gel thin film pyroelectric infrared sensor. Flip chip bonding of sensor to interface die permits separation of the high temperature, ideal infrared material processing, from CMOS fabrication.*

The CIMS process offers a unique solution to pyroelectric materials integration. The CIMS infrared sensor relies on a new silicon-compatible $PbTiO_3$ (PTO) sol-gel material system that is applied to the CIMS sensor die in a modular process separated from CMOS fabrication. The CIMS structure provides ideal, thermally isolated, structures for support of the pyroelectric film detector. Flip-chip bonding integrates the optimized detector with a CMOS signal processing system, ideally suited for pyroelectric sensor measurement.

The sol-gel synthesis of PTO and PCTx is based on the process first developed by J.B. Blum and S.R. Gurkovich[5], but with significant modifications. The alkoxide sol-gel precursor of PTO is prepared as follows: $Pb(C_2H_3O_2)_2{\cdot}3H_2O)$, lead acetate trihydrate, is first dissolved in either 2-methoxyethanol (2-Me) or in methanol at a concentration of 1.0 molar. Acetic acid anhydride, $(CH_3CO)_2O$, (AAA) is then added in a molar ratio of 3:1 with respect to lead, removing the water of hydration by producing acetic acid. A 10% excess of AAA is used to ensure that any water present in the solvent or ambient air is also immediately transformed into acetic acid, thus keeping the precursor anhydrous. Titanium isopropoxide, $Ti(OC_3H_7)_4$, is then added into solution. The double metal alkoxide $(C_2H_3O_2)$-Pb-O-Ti-$(OC_3H_7)_3$ is formed along with the reaction by-product, isopropyl acetate. Synthesis is conveniently completed in ambient air without interference from atmospheric water, since any water absorbed by the solution will react with excess AAA. Also, storage of the precursor solution requires only hermetic container sealing. In addition, acetic acid generated in the reaction provides the low pH environment necessary for slow hydrolyzation. Precursors appeared to have shelf-life longer than 6 months without gelation or formation of particulates. This represents an improvement over previous processes which require all procedures to be carried out in a dry nitrogen atmosphere, as well as extensive refluxing of the solution above 100 °C to remove water and the addition of nitric acid to prevent premature gelation. The PCTX precursor is produced using calcium acetate monohydrate, $Ca(C_2H_3O_2)_2{\cdot}H_2O$, as the source of Ca. Methanol is used as solvent for the PCTX precursors due to the low solubility of calcium acetate in other alcohols. Precursors were spin-coated onto the silicon substrates with a conventional photoresist spinner to form a gel layer about 0.3 μ thick. The samples are then dried at 400 °C for 1 hour to remove most of the organics. Multiple spin-dry cycles are used to obtain the desired thickness. Final heat treatment at 650 to 750 °C for 1 to 2 hours removed all organics, leaving smooth, crack-free films. All procedures are performed in air.

The CIMS thermal infrared sensor structure, with a thin membrane sensor support structure, was fabricated by conventional bulk micromachining. A p^+ etch-stop epitaxial layer, of 4μ thickness is deposited on an n-type substrate to form the sensor substrate structure. EDP etching releases the suspended membrane shown in Figure 3. Spin coating methods deposit PTO or PTCX followed by the heat treatment steps. A buffered oxide etch (BOE) provides patterning for the thin film and p^+ supporting structures. An Al electrode completes the sensor structure. The pyroelectric signal is detected by measurement of the potential developed between the highly conductive p^+ substrate and the deposited Al film.

Detailed thermomechanical response for the thermal infrared detector model was implemented using a SPICE simulation method. The suspended PTO film, Si supporting membrane, and Si low thermal conductance supports (see Figure 3), were treated as thermal resistance mesh networks. SPICE simulation of detector frequency response is in excellent agreement with experimental

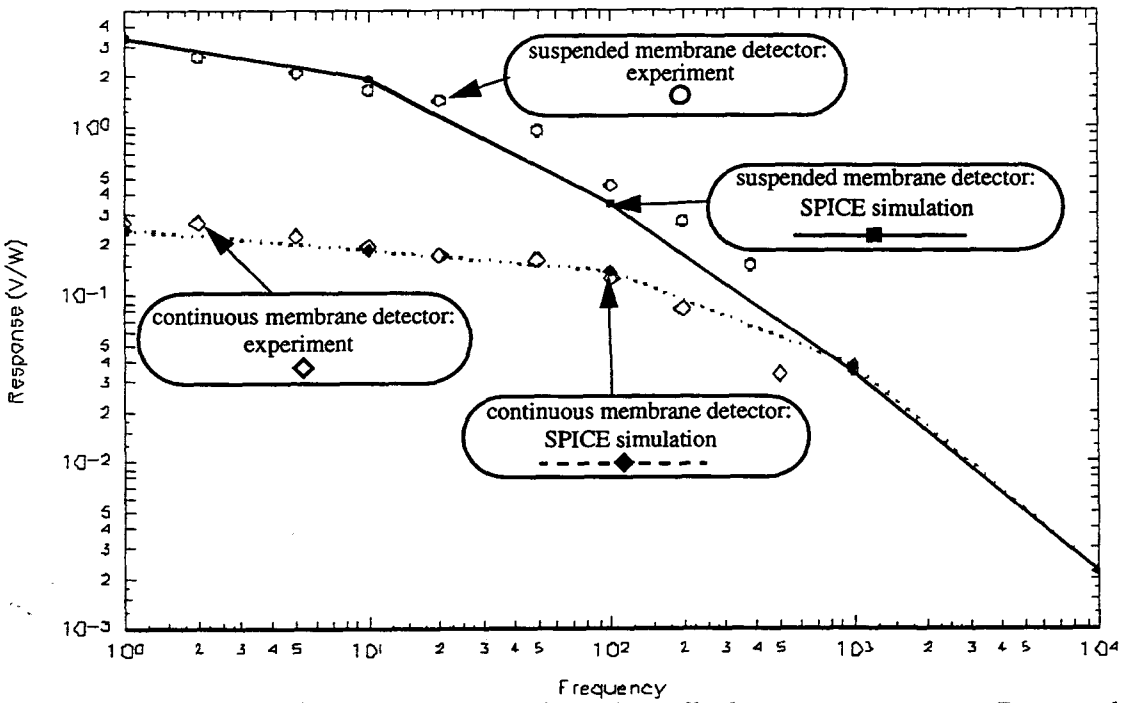


Figure 4. Response of the CIMS infrared sensors for two PTO pyroelectric film detector support structures. Frequency dependence of responsivity is shown for continuous p^+ Si membranes, and for suspended p^+ Si membranes supported by narrow mechanical bridge elements (see Figure 3). Responsivity enhancement due to the low thermal conductance suspension is seen clearly.

results.[2]. Direct measurement using a chopped CO_2 laser source yields a responsivity of 2.5V/Watt at $\lambda = 10.6\mu$ for this pyroelectric sensor (see Figure 3 for an example of detector transient response). This prototype device demonstrates excellent performance without an infrared absorber layer. Infrared absorbers for optimized detectors will be deposited by conventional metal "black" and continuous metal film absorbers. Modeling of detector thermal response demonstrates that the pyroelectric material response is equal to that of bulk PTO material.

Figure 4 shows measured and calculated frequency dependence of the infrared sensor responsivity. The SPICE thermal response simulation and measurements are shown for PTO pyroelectric films supported by both continuous membrane and suspended membrane (see Figure 3) structures. The large responsivity improvement provided by the low thermal conductance suspension are apparent. In addition, the excellent agreement between SPICE simulation and experiment indicate the capability for accurate thermal infrared sensor design offered by the SPICE simulation method and CIMS fabrication.

Integrated measurement systems, with micropower operational amplifiers and data conversion systems (Figure 5) provide switched-capacitor measurement capability for CIMS infrared sensors.

WIRELESS MICROSENSOR COMMUNICATION

The wireless sensor technology reported here differs from prior work in that the sensor, control, and communication system are *integrated together in a single unit.* In addition, we have addressed the challenge of combining low noise, high performance measurement and data conversion sensor interfaces with micropower operation in CMOS systems. An important technology challenge is the development of micropower, integrated CMOS RF systems. This requires advances in both level of integration, sharp reduction in operating power, and a conversion from conventional bipolar technology to CMOS.

Powerful wireless microsensor networks may only be implemented with nodes capable of bidirectional communication. Remote LWIM nodes must be dynamically reconfigurable by a base station or distributed control system. Thus, each LWIM node must operate with both receiver and transmitter functions. Now, the requirement for operation with compact battery cells presents new constraints on transceiver operation. First, average current requirement (at 3V supply bias) for the LWIM node must be less than 30μA to provide desired operating life. This requirement may be met partially by proper power management through low duty cycle operation of the transceiver system. For example, typical surveillance applications permit duty cycle values less than 10^{-3} for 10kbps data rate. However, in addition, peak operating LWIM node current must be less than 1 - 10mA due to maximum current restrictions for compact Li cell sources. LWIM nodes, therefore, require advances in transceiver technology over conventional systems that require greater than 100mA supply current at 3V. In particular, the fundamental restriction on input-referred amplifier noise (sensitivity) and bias current, limit conventional receiver sensitivity. In addition, high performance radio technology has not been demonstrated in conventional CMOS integrated systems.

The LWIM transceiver system relies on new methods for obtaining sensitivity and selectivity with low power CMOS circuits. First, receiver sensitivity relies on the sensitivity of the RF amplifier front-end. The LWIM receiver system obtains low input-referred noise and low power operation by introduction of high-Q L-C tank circuits that provide passive (noise-free) voltage gain. In addition, low oscillator phase noise, required for high selectivity is obtained by the use of coupled high-Q L-C resonators and surface acoustic wave (SAW) resonators for narrow band stabilization.

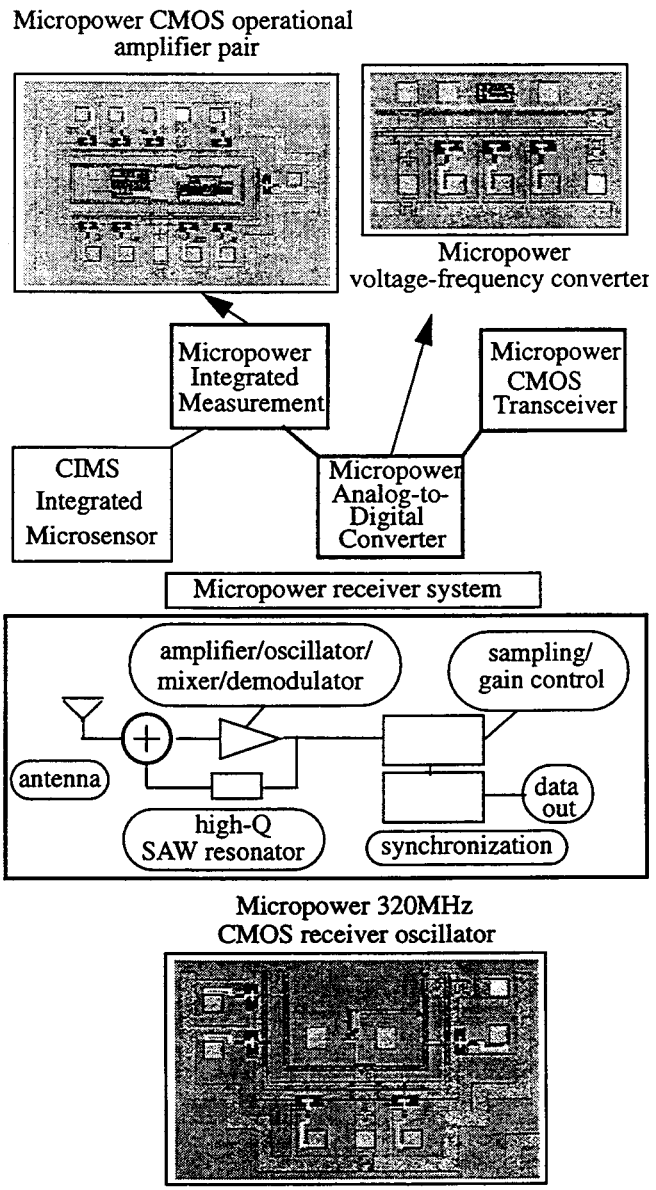


Figure 5. Wireless Integrated Microsensor Interface and Communication System. *Measurement, data conversion, and communication systems have been demonstrated in commercial 2μ CMOS. A micropower operational amplifier (supply current 1μA) is combined with micropower voltage-frequency data converters (supply current 2μA). The transceiver is based on a low power CMOS, SAW stabilized and digitally-sampled regenerative architecture (peak supply current 30μA).*

The low operating power and low phase noise of the LWIM receiver architecture has been demonstrated by a new regenerative receiver system. The high selectivity micropower receiver system operates at 310MHz with detector currents of less than 30 μA at 3V supply bias. This device relies on the regenerative receiver principle. However, it differs from previous systems in that analog sampling of the regenerative wave form is employed to provide demodulation. Stable operation through automatic gain control is now possible. Figure 5 shows a block diagram of the prototype LWIM receiver. A SAW-stabilized Colpitts oscillator employs a single loop inductor functioning as a receiving antenna. RF signal is added to oscillator loop signal at the inductor. Periodic quenching and measurement of amplitude of the 310 MHz loop oscillation is controlled by a digital sampling system. Received RF signal strength modulates oscillator amplitude. Finally, oscillator amplitude is detected through direct measurement of oscillator current. Prototypes of this amplitude-shift keying (ASK) receiver operate with sampling clock rates as high as 1MHz. Thus, transceiver data rate capability greater than 100kbps may be permitted. Prototype LWIM nodes operating with 1mW transmitter power, 1kbps data rate, and 10m range, require single loop antennae of less than $1cm^2$ area. The SAW stabilized, low phase noise oscillator can also provide a local oscillator for advanced, frequency-hopped spread spectrum micropower transceivers.

Low power interface and control circuits have also been implemented for LWIM nodes. Figure 5 shows micropower operational amplifiers (1 μA supply current) and voltage-to-frequency converters for analog to digital conversion (1 μA supply current at 250kHz clock rate). Finally, low power, low clock rate digital control provides sequencing for the LWIM microsensor data conversion and communication system.

CONCLUSIONS

The wide range of distributed monitoring and control applications requires microsensors that integrate network interfaces with diverse measurement capability. A low power integrated wireless microsensor (LWIM) technology has been developed using new processes, circuits, and communication systems. The MEMS microsensor structures incorporate high performance, bulk micromachined transducers with conventional CMOS microelectronics. Low power sensor interface, digital control, and RF communication systems have been demonstrated. Both RF induction and battery powered nodes may be implemented with the micropower system demonstrated here. LWIM nodes are under development for a diverse set of military and commercial distributed MEMS applications.

* Research supported by the Defense Advanced Research Projects Agency (DARPA) under contract J-FBI-95-169

REFERENCES

[1] A. Burstein and W. J. Kaiser, "Mixed Analog-Digital highly sensitive sensor interface circuit for low cost microsensors", *Sensors and Actuators* (in press)

[2] J. G. Ho, P. R. Nelson, F. Lin, D. T. Chang, W.J. Kaiser, and O. M. Stafsudd, "Sol-gel derived Lead and Calcium Lead titanate pyroelectric detectors on Si MEMS structures". *Proc. SPIE* (in press).

[3] F. Chang, R. Yeh, P. Chu, E. Hoffman, E. Kruglick, K. Pister, M. Hecht, "Gas-phase silicon micromachining with xenon difluoride", *Proc. SPIE Microelectronic Structures and Microelectromechanical Devices for Optical Processing and Multimedia Applications*, 117-128 (Oct. 1995).

[4] O.M. Stafsudd, M.Y. Pines, "Characteristic of KTN pyroelectric detectors," *J. Opt. Soc. Am.*, Vol. 62, 10, 1153 (1972)

[5] J.B. Blum, S.R. Gurkovich, "Sol-Gel derived $PbTiO_3$," *J. Mat. Sci.* 20 (1985) 4479.

A SMART AUTOMOTIVE ACCELEROMETER WITH ON-CHIP AIRBAG DEPLOYMENT CIRCUITS

Leland "Chip" Spangler **Christopher J. Kemp**

Ford Microelectronics, Inc.
9965 Federal Drive
Colorado Springs, CO 80921

ABSTRACT

This paper describes a "smart" accelerometer for use in automotive passive restraint systems. This accelerometer performs, in a single integrated component, all of the sensing and signal processing functions required to assess vehicle crash severity and generate a timely airbag deployment signal if and when it is needed.

INTRODUCTION

Automotive passive restraint systems continue to utilize the benefits of microelectronics to provide more sophisticated occupant safety features. Traditional microelectronic advantages such as miniturization, system integration, and part count reduction are being used to provide additional passenger safety through wider applications and more advanced features.

Sensors play a primary role in these smart systems as they provide the information the system uses to make deployment decisions. Occupant detection and occupant position sensors will be used in the future to allow these features to be implemented so that occupant safety in a crash can be improved. Accelerometers are, however, the first high volume electronic sensor to see widespread use in electronic-based passive restraint systems. Already the evolution of electronic frontal airbag systems is placing increasing demands on accelerometers, including greater accuracy, better diagnostics, improved robustness, smaller size, and lower cost. The ever decreasing size of electronic components, with the ability to integrate more functions, will drive the components to incorporate more and more of the system function.

This paper describes a "smart" automotive accelerometer that continuously evaluates vehicle decelerations and generates a timely airbag deployment command if needed. Vehicle and system level requirements have led to the adoption of several key driving factors in designing this version of the accelerometer. These requirements include a high degree of accuracy in accelerometer sensitivity and offset, insensitivity to temperature as well as voltage and process variations, a high level of reliability, low sensitivity to electromagnetic radiation and power supply noise, and low power consumption. The specification of low quiescent current of this fully-integrated system supports key-off operation as well as the low power requirements of electric and hybrid passenger vehicles. In addition, the small size supports the demand for reduced module size which in turn facilitates vehicle-level installation.

These factors lead to a different approach to system partitioning that supports the requirements of next generation frontal and side passive restraint systems. The partitioning reduces demands on the system and microprocessor by performing signal processing and algorithm computation in the accelerometer so that the output of the accelerometer is a signal that is used to activate the passive restraint.

APPLICATION

Passive restraint systems are refered to as "distributed" when mechanical acceleration switches are used at multiple positions near the front of the vehicle to evaluate crash severity and decide "if and when" an airbag should be deployed (Figure 1). A separate wiring harness is used to connect these under-hood sensors with the central diagnostic module. Single-point electronic passive restraint systems (Figure 1) eliminate the under-hood sensors and wiring harness, and instead use a single accelerometer to provide data for a microprocessor-based algorithm that makes deployment decisions. The electronic module that contains the accelerometer and microprocessor is located in the passenger compartment of the vehicle. This simplifies vehicle assembly, lowers cost, and improves reliability while offering system and component self-test capability.

SENSOR DESIGN

The accelerometer is a two-chip device that consists of a differential, capacitive micromachined sense element die assembled in an IC package along with a CMOS interface chip containing EEPROM calibration circuits. It is an adaptation of the accelerometer reported by Cole [1] and an extension of the accelerometer we previously reported [2]. The system block diagram is shown in Figure 2, where the signal path is indicated by the bold arrows. The three modular components of the accelerometer (two chips and the package) can be developed, scaled, or optimized without making changes in the other components of the sensor. This flexibility offers advantages in lowering development costs and time, while allowing application specific designs that meet specific system specifications.

The sense element die produces capacitance variations in response to the acceleration signal. The sense element chip uses a

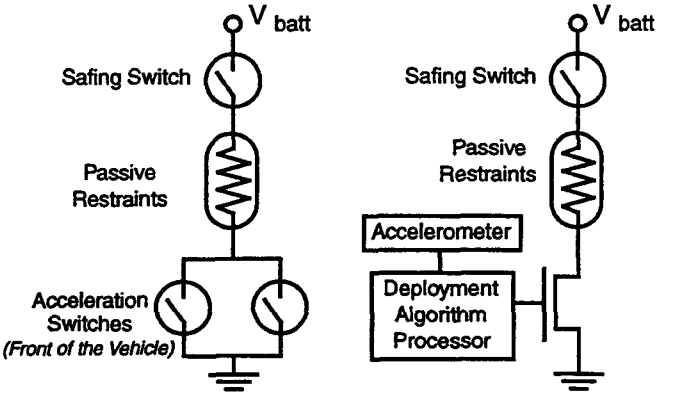


***Figure 1.** Schematics of automotive passive restraint systems. On the left is a distributed sensor system, on the right is a single-point sensing system.*

0-9640024-1-8 $3.00

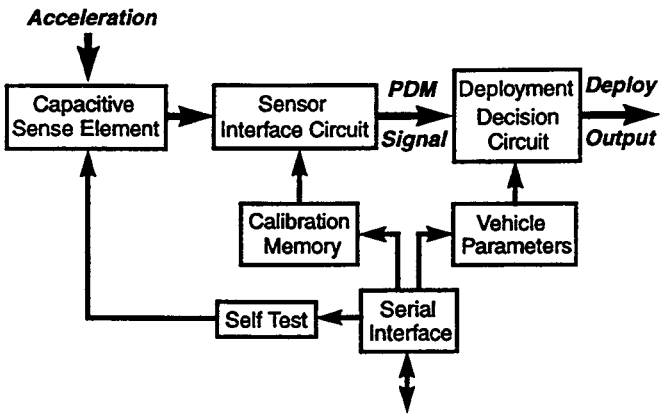


Figure 2. The block diagram of the smart accelerometer.

mirror-symmetric arrangement of two torsional microstructures that generate a fully differential, femtofarad-level signal that is fed to the interface chip via bond wires.

The interface chip uses a charge-mode delta-sigma modulator to convert the sense element signals into a pulse-density-modulated output that is proportional to the applied acceleration[3]. The interface chip has EEPROM and a resistor string DAC that allow the gain and offset of the completed device to be trimmed to the desired specification. A programmable digital algorithm uses circuit blocks that calculate magnitudes, changes in velocity over time intervals, and peak acceleration data to make deployment decisions. The programmable nature of this algorithm allows it to be configured for many vehicle types. A serial interface allows the device to communicate with a microprocessor to facilitate accelerometer calibration and algorithm programming during device manufacture, and to facilitate data transfer and a secure self test activation after the device is installed in a vehicle.

This device is presently packaged in a multilayer ceramic package, but the modular nature of the technology allows other versions of the accelerometer to be packaged in a conventional plastic integrated circuit package [2].

MICROMACHINED SENSE ELEMENT

The sense element contains two torsional microstructures in which movable capacitor plates are suspended, from a central "pedestal" support, above fixed capacitor plates metallized on a glass substrate. A simplified schematic is shown in Figure 3. Typical values for the capacitance at zero acceleration are 150fF while full-scale acceleration results in approximately a 15fF increase to C_A and C_D and a somewhat smaller decrease to C_B and C_C.

The element is fabricated using a single-crystal silicon, dissolved-wafer process [2] on a glass substrate which is mounted in the same package with, and wirebonded to, a CMOS ASIC which contains all of the remaining circuitry. Dimensioning of the element's plate thickness, air gap, torsion beams, and damping holes creates an overdamped, second-order mechanical system with a resonant frequency which is typically 2.4KHz and a dominant pole which is typically at 375Hz. This bandlimiting of the acceleration signal by the sense element is used to provide antialias filtering for the sampled-data interface circuit and signal filtering for accurate algorithm processing. To provide a self test capability, the sense element's glass substrate incorporates additional metallized regions which are used, during a self test operation, to electrostatically deflect the movable plates.

The central pedestal support greatly reduces the susceptibility of the sensor to die stresses since it mechanically isolates the microstructure from the substrate and from the packaging materials. This eases packaging constraints and also improves the manufacturability of the sensor. This mechanical configuration is one factor that allows the accelerometer to be manufactured without temperature compensation circuitry. The use of two torsional microstructures in the accelerometer offers several advantages that are a result of the four capacitors being configured as a full bridge. This electrical and mechanical symmetry of the structure greatly reduces sensitivity to cross axis forces, power supply noise, and electromagnetic radiation, through the use of a fully-differential signal path.

The design of the torsion arms' dimensions offers flexibility in meeting a variety of performance specifications that the applications require. Typical torsion arm lengths are in the range of 250 micrometers long with a cross sectional area that is in the range of 100 square micrometers. The "heavy" side of the microstructure is approximately 3 times more massive than the "light" side, and each suspended microstructure occupies an area of approximately 0.8 square millimeters.

The fixed plates of the sense element are formed from metalized regions on the glass substrate. These fixed plates are larger than the silicon microstructure so that cross axis forces causing the microstructure to twist parallel to the substrate do not cause the nominal capacitance to change. The electrostatic self test is implemented on the "heavy" side of the microstructure by placing additional fixed electrodes under the movable silicon. Self test voltages, provided by the interface chip, are applied to these fixed plates during the self test operation, and the resultant electrostatic force causes the microstructure to rotate in a manner that mimics the microstructure's motion during a vehicle crash.

The initial configuration of the device uses a hermetic package and a sense element die that is not capped. With this configuration, electrical contact to the microstructure is made via a bond pad and wire that are directly applied to the pedestal region. The hermetic seal of the ceramic package results in a fixed gaseous environment that is an important factor in determining the mechanical response of the microstructure to vibrations. Gas viscosity changes as a function of temperature and this changes the damping characteristics of the microstructure. Both gas type and pressure must also be considered as factors that influence the mechanical response of the microstructure. The pressure inside the package is dependent primarily on two factors: the sealing temperature of the lid to the package, and the ambient pressure at sealing. The lid is solder sealed to the ceramic package using an

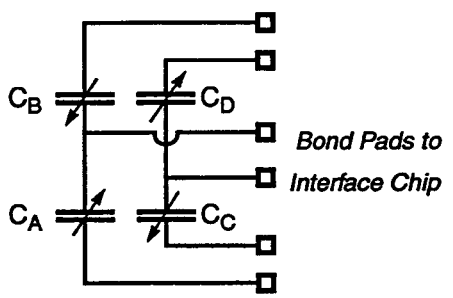


Figure 3. Schematic of the differential capacitive sense element.

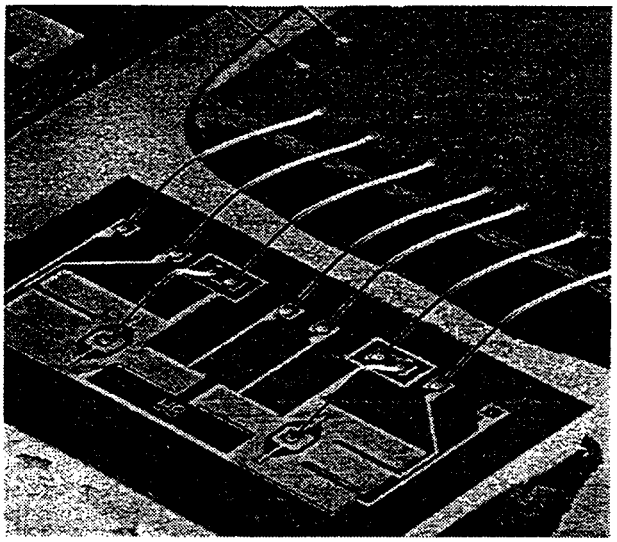


Figure 4. Scanning electron micrograph of the sense element chip wire bonded to the interface chip in a ceramic package

elevated temperature in a conventional furnace. Colorado Springs is located at an altitude of over 6500 feet which results in a local barometric pressure that is about 80% of that at sea level. The result is that the pressure in the cap is less than 0.5 atmospheres.

INTERFACE AND ALGORITHM CIRCUITS

The next modular component of the accelerometer to be discussed is a CMOS interface chip which has two primary functions. First, the interface circuit converts the differential femtofarad-level capacitance signals from the sense element to a pulse-density-modulated output. This circuit contains calibration circuitry to remove offset and sensitivity variation from the fully assembled device. The Deployment Decision Circuit (DDC) is the second major function on the interface chip. The DDC implements a programmable algorithm which, when given the appropriate parameters (based on the vehicle body and sensor location) decides if and when a passive restraint should be activated. In addition, the interface chip implements the self test and reset functions as well as the serial interface to both the system and to production test/calibration equipment. The interface chip, fabricated in a 1.6mm CMOS process, has a die area of approximately 19 mm^2.

The interface circuit is a first-order, charge-mode $\Delta\Sigma$ modulator similar to that reported in [3]. The primary difference is the fully differential nature of the present sensor and signal path. The interface circuit drives the acceleration-dependent sense element capacitors with voltage transitions which pump charge packets either onto, or off of, the modulator's integrator on each cycle of the 250KHz clock. On a given clock cycle, charge is pumped onto the integrator if its output voltage on the previous clock cycle was negative, and off of the integrator if its output voltage on the previous clock cycle was positive. The output pulse density is determined by the ratio of the charge pumped on to the charge pumped off. These charges depend on the sense element capacitances (and thus the acceleration) as well as the voltage transitions produced by the interface circuit. Calibration is performed by precise adjustment of these voltage transitions [3]. The interface circuit is designed to be very sensitive to changes in sense element capacitances; a change in the charge packet sizes of only six electrons produces a one hertz change in output pulse density.

The nature of the pulse density modulated (PDM) output simplifies the implementation of the on-chip deployment decision circuit, whose (simplified) block diagram is shown in Figure 5. Each pulse of PDM output corresponds to a specific velocity change, so for example, with a 50 g calibration and a 4MHz clock, each PDM pulse corresponds to a change velocity of approximately 2×10^{-3} meters/sec. This approach allows the deployment algorithm circuits to be implemented, with a minimum amount of chip real estate, using counters, timers, and pulse integrators.

The "prediscriminator" in the DDC constantly evaluates the output of the accelerometer, and based on the programmed parameters and the acceleration, a "threshold" logic signal is generated. This threshold signal is used as an enable for the other blocks in the "discriminator" portion of the DDC.

The discriminator output is based not only on acceleration thresholds and magnitude calculations but also on total velocity *and* acceleration changes calculated over particular time periods. The DDC allows these calculations to be made simultaneously in real time. The output of these individual blocks are then logically combined to implement the overall deployment decision algorithm. Since the DDC's counters, timers, and pulse integrators are programmable over a wide range, the algorithm parameters can be adjusted to optimize deployment decision performance for a particular vehicle body style and sensor location.

SENSOR PACKAGING

The modular approach to the accelerometer is also evident in the packaging technologies selected for the device. The first samples of this accelerometer use conventional packaging technology in which the two die are mounted in the cavity of a multilayer ceramic DIP. The accelerometer is completed by wirebonding the two die together using conventional gold wire bonds. Wire bonds are also used to connect the interface chip to the package pins. A solder lid is attached over the package cavity as discussed previously.

The accelerometer package and die mounting influence the orientation of the primary axis of sensitivity. The ceramic through-hole package provides orientation flexibility while at the same time the solder sealed lid ensures a controlled ambient for the accelerometer. A SIP style package has also been developed and this allows the axis of sensitivity to be parallel the module printed wiring board.

While the multilayer ceramic technology has been used to provide a controlled and reliable environment for the microstructure, the hermetic packaging requirement can also be transferred to the sense element die through the development of a wafer-level capping technology that is compatible with the dissolved wafer process [2]. Wafer level capping allows the accelerometer to also be assembled in a conventional molded plastic

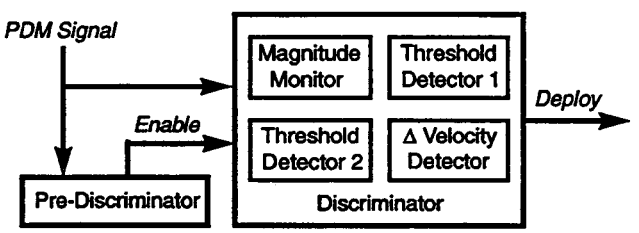


Figure 5. The block diagram of the deployment decision circuit.

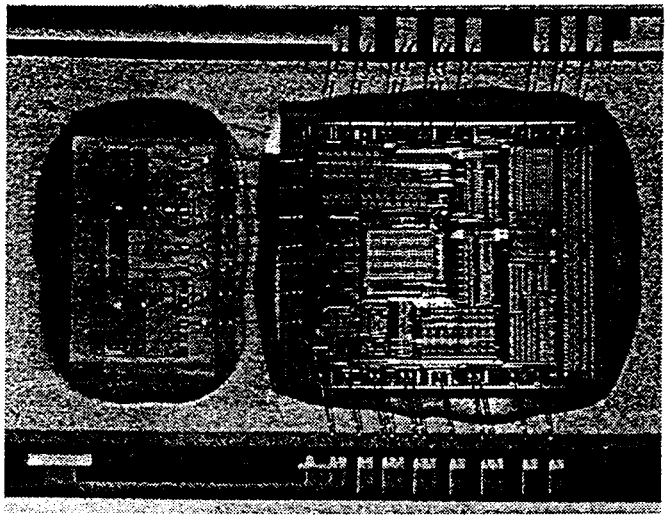


Figure 6. Photograph of the two-chip smart accelerometer.

package which supports many package styles and lead configurations while lowering overall sensor cost.

SUMMARY AND RESULTS

The accelerometer and its modular components have been fabricated with technologies that will continue to evolve to support a variety of sensor configurations, cost reductions, and feature optimizations. The configuration discussed in this paper offers excellent accelerometer performance, reliability, and manufacturability, as well as design features that minimize overall system costs. Figure 6 is a photomicrograph of the accelerometer, showing the sense element die and the signal processing ASIC in the ceramic package with the lid removed.

The sensor is designed to operate from a 5 volt supply and an external 4MHz clock. On-chip dividers generate an interleaved two phase clock at 250KHz which is the sampling frequency for the delta-sigma modulator. Greater clock speeds can be used to increase the resolution of the sensor. The part has been tested to clock frequencies up to 16MHz.

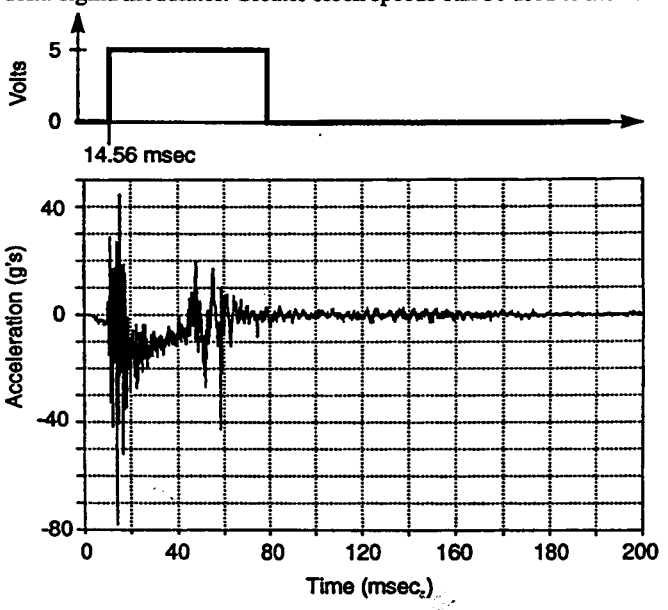


Figure 7. Acceleration data recorded at the crash barrier in a 14 mph impact with a 90° wall (bottom) and the measured output of the smart accelerometer (top).

The accelerometer is typically calibrated under specific acceleration conditions in a custom designed centrifuge or "shaker." These calibration instruments use reference devices to monitor the acceleration applied to each device that is being calibrated. Only two room temperature readpoints (0g's and 15g's) are used for calibration, even though the full scale range of the accelerometer is +/-50 g's. The exact reference value measured at these readpoints is then used by the test software to generate specific calibration coefficients for each device. These parameters are then programmed into the accelerometer's EEPROM to set the offset and sensitivity of the accelerometer. Sensitivity and offset have temperature coefficients of less than 60 ppm/°C at the worst case temperature *and* voltage conditions.

Measured performance of the integrated airbag deployment system, for a particular crash, is depicted in Figure 7. The crash acceleration waveform is shown along with the deploy pulse, which occurred at 14.56 msec after impact (t=0). This differs by only 740 µsec from the simulated deploy pulse which occurred at 15.30 msec after impact. A wide variety of responses is possible from the device based on accelerometer calibration and DDC algorithm coefficients.

The accelerometer has been subjected to standard automotive integrated circuit reliability tests. This includes temperature cycling, vibration endurance, shock and drop testing, and accelerated life testing. Large quantities of parts have been exposed to these tests and the successful results demonstrate the sensor's ability to withstand the automotive environment. Module and vehicle-level testing of the accelerometer at the crash barrier along with real world tests in fleet vehicles has also proven that the design objectives have been met.

ACKNOWLEDGEMENT

The authors would like to acknowledge the staff at Ford Microelectronics, Inc. who have worked to make this project successful. In addition, the support of the Automotive Components Division and Ford Research Laboratories has been essential.

REFERENCES

1. J. Cole, "A new sense element technology for accelerometer subsystems," *Transducers '91*, June 24-27, 1991, pp. 93-96.

2. L. Spangler, C. Kemp, " ISAAC (Integrated Silicon Automotive Accelerometer)", *8th International Conference on Solid-State Sensors and Actuators, Transducers '95*, Stockholm, Sweden, June 25-29, 1995, pg. 585.

3. C. Kemp, L. Spangler, "An Accelerometer Interface Circuit," *1995 Custom IC Conference*, May 1-4, 1995.

A LOW-POWER GENERIC INTERFACE CIRCUIT FOR CAPACITIVE SENSORS

Navid Yazdi, Andrew Mason, Khalil Najafi, and Kensall D. Wise

Center for Integrated Sensors and Circuits
University of Michigan
Ann Arbor, MI 48109-2122

ABSTRACT

This paper presents a low-power generic interface circuit compatible with smart microsystems and capable of interfacing with a large variety of capacitive sensors. The interface chip is highly programmable, can communicate with an external microcontroller over a standard sensor bus, supports sensor self-test, and includes a temperature sensor. The circuit can interface with up to six external sensors and contains three programmable internal reference capacitors in the range of 0.15-8pF. The chip measures 3.2mm x 3.2mm in a standard 3μm single-metal double-poly p-well process, dissipates less than 2.2mW from a single 5V supply and can resolve input capacitance variations of less than 1fF.

INTRODUCTION

Smart microsystems combining a large mix of sensors, actuators and electronic circuitry are of increasing demand in a wide range of applications, including those in distributed sensing and control systems, industrial process automation, instrumentation, and environmental monitoring. Capacitive microsensors are very attractive for these microsystems due to their low power, high sensitivity, and self-test capabilities [1]. However, the base capacitance and sensitivity of these capacitive sensors can vary over a wide range, which must be covered by the interface circuit. Hence, there is a growing need for a low power generic interface for use with capacitive sensors in smart microsystems. Although there have been numerous readout circuits introduced for capacitive sensors [2-3], to date a generic interface has not been reported. Such an interface circuit should be able to read out various capacitive microsensors with a wide range of base capacitance and sensitivity, communicate with the main microsystem controller, support self-test and self-calibration, support multi-ranging for a single sensor, dissipate low power, and occupy very small area with no external components. This paper introduces such an interface circuit.

The interface chip is currently being utilized in the microsystem for environmental monitoring shown in Fig. 1 [1]. The following sections provide a description of the interface chip and its functions. First, the current microsystem which employs the interface chip is briefly described. Then, the interface chip architecture is presented, followed by a description of its major building blocks. Finally, measurement and test results for the fabricated chip are presented.

MICROSYSTEM DESCRIPTION

The microsystem utilizes an open architecture with a standard sensor bus [1], which allows the system to be populated with various sensors and actuators as required by the application. For instance, the currently developed microsystem accommodates sensors for a variety of environmental parameters, including temperature, barometric pressure, humidity, and acceleration [1]. The heart of the microsystem is a microcontroller unit (MCU) that provides the system intelligence and allows for data storage as well as software routines for each sensor to perform in-module sensor calibration, digital compensation, and self-test. The MCU communicates with the front-end transducers via the sensor interface chips and a nine-line sensor bus as illustrated in Fig. 1. The interface chip is a critical component of this overall system because it allows a standard microcontroller to communicate and collect data from a variety of sensors. Once sensor data has been collected and calibrated in-module, it can be stored or transmitted out through either a hardwired serial I/O port or an on-board telemetry device.

The sensor bus is a central part of the microsystem architecture, which provides access to multiple sensors/ actuators while using a minimum number of interconnects without any loss of functionality. The sensor bus consists of nine lines: three for power, two handshaking lines used with a clock and serial data line for synchronous serial communication, a data output line, and a data valid/interrupt signal [1]. The three power leads are the system supply, ground reference, and a switched reference voltage for devices that can be turned off during a low-power system sleep mode. For synchronous communication with the front-end devices, a chip enable and data strobe provide handshaking for use with a programmable-frequency clock and a serial data line. The last

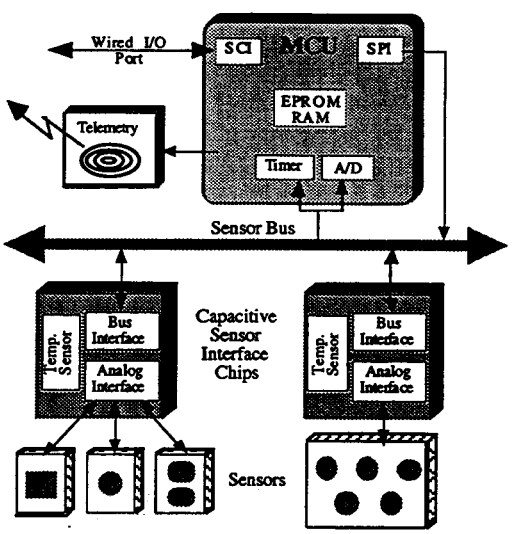


***Figure 1:** Block diagram of a multi-parameter sensing microsystem.*

0-9640024-1-8 $3.00 © 1996 TRF

Solid-State Sensor and Actuator Workshop
Hilton Head, South Carolina, June 2-6, 1996

two lines of the sensor bus are for sensor data output, which can be an analog level or a frequency-encoded bitstream, and data valid which signals the MCU when valid data is present on the data out line. Data valid is also used as an interrupt from the front-end devices that can initiate an event triggered sensor readout.

INTERFACE ARCHITECTURE

Figure 2 shows the block diagram of the interface chip. This chip utilizes a low-noise front-end charge integrator to read out the difference between the sensor capacitance and a reference capacitor [3]. Communication with the MCU is performed over the sensor bus and is handled by the bus interface unit. Serial data transmitted over the bus is received, decoded, and applied to control the other circuit blocks illustrated in Fig. 2. The input multiplexer enables the chip to interface with up to six sensors. Furthermore, the chip can be digitally programmed to operate with one of three external/internal reference capacitors. The on-chip reference capacitors are laser trimmable in a range of 0.15pF to 8pF. Programmability of the reference capacitors allows the chip to interface with capacitive sensors having a wide range of base capacitance and also provides offset control.

The chip analog signal path consists of an input multiplexer, the input charge integrator, a gain stage, an output sample/hold circuit, and an output multiplexer. The gain stage can be programmed on-line to any of its three gain settings. These gains can be used to accommodate input sensors with different sensitivities or for multi-ranging a single sensor and are set by laser trimming the on-chip capacitors. The overall readout sensitivity can be varied from 0.23 mV/fF to 73.5 mV/fF using both digital programming and laser trimming, which gives an effective gain variation from 1 to 312. The output multiplexer provides access to the sensor bus for both the capacitive readout circuitry and the on-chip temperature sensor. The interface chip additionally supports self-test and self-calibration. A 3b on-chip DAC can be utilized to generate a variable-amplitude two-phase clock for driving the sense and reference capacitors in each input charge integration cycle. The variable amplitude clock can also be used to apply a varying electrostatic force to the sensor for self-test and self-calibration.

CIRCUIT BUILDING BLOCKS

The readout circuit is a switched-capacitor circuit consisting of a low-noise charge integrator, a digitally-programmable gain stage and a sample/hold circuit as shown in Fig. 3. Since the

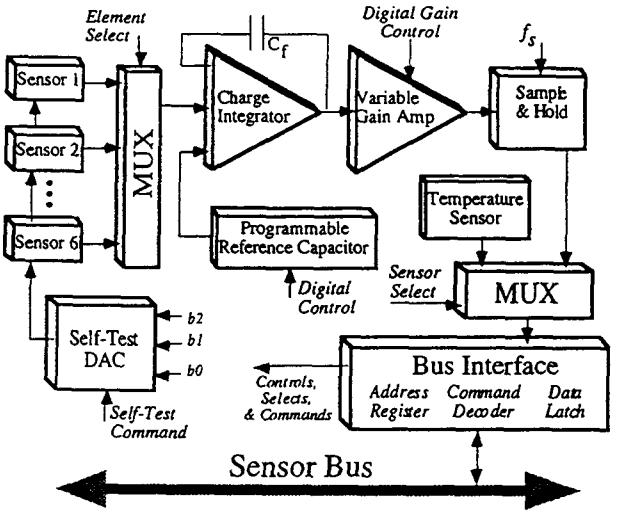


Figure 2: Diagram of the capacitive sensor interface chip showing the major circuit blocks.

circuit operates from a single supply, an internal analog ground reference level, half way between the supply level and ground, is generated on-chip and is shown as Vref in Fig. 3. Clock switching noise is reduced by using sized dummy switches at the high-impedance nodes; a fully differential circuit is not utilized to save power and die area. The control clocks of the dummy switches are slightly delayed so that better charge compensation occurs when the actual switches are turned off completely.

The input switches are controlled so that at each instant only one of the sense capacitors (Cs) and one of the reference capacitors (Cref) are connected to the circuit. During $\phi 1$, the reset switch of the charge integrator is closed and Cs is charged through the charge integrator output. Once $\phi 1$ goes low, a packet of charge proportional to the difference between Cs and Cref is integrated on the feedback capacitor of the first stage (Cf1). Meanwhile, the second stage integrator is in reset mode and either Cg1, Cg2, or both Cg1 and Cg2 charge to the output level of the first stage. The gain of the second stage is determined by the ratio of the total capacitance switched into its input to the feedback capacitance (Cf2). Clock phases $\phi 3$ and $\phi 4$ are slightly delayed $\phi 1$ and $\phi 2$ clocks which become active only when gain 1 or 3 is selected. Similarly, $\phi 5$ and $\phi 6$ are delayed $\phi 1$ and $\phi 2$ clock phases that are active only when

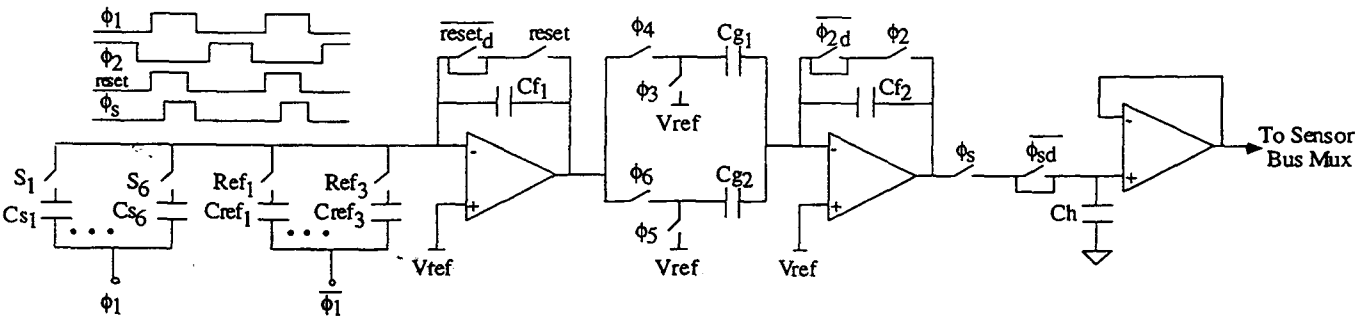


Figure 3: Schematic diagram of the switched-capacitor readout circuit.

digital gain 2 or 3 is set. Therefore, the gain of the second stage is selected by controlling switching of Cg1 and Cg2 at its input. During $\phi 3$ ($\phi 5$) clock phase, the charge stored on Cg1 and Cg2 capacitors is integrated on Cf2. Finally, the output of the second stage is sampled and held at the input of the third stage during ϕs.

Figure 4 shows the schematic of the opamp used in the switched-capacitor circuit. The opamp is designed to provide low noise and good high frequency PSRR with low power dissipation, and to drive capacitive loads over a wide range. Transistors M1-M5 form the input differential stage, with large PMOS transistors (M1-M2) used to reduce noise. The main gain stage is a class AB cascode output stage consisting of M8-M11. The dominant pole is determined by the load capacitance, which alleviates the need for a compensation capacitor and effectively improves the high frequency PSRR. The load capacitance can vary over a wide range without causing stability problems and is driven efficiently by the class AB output stage.

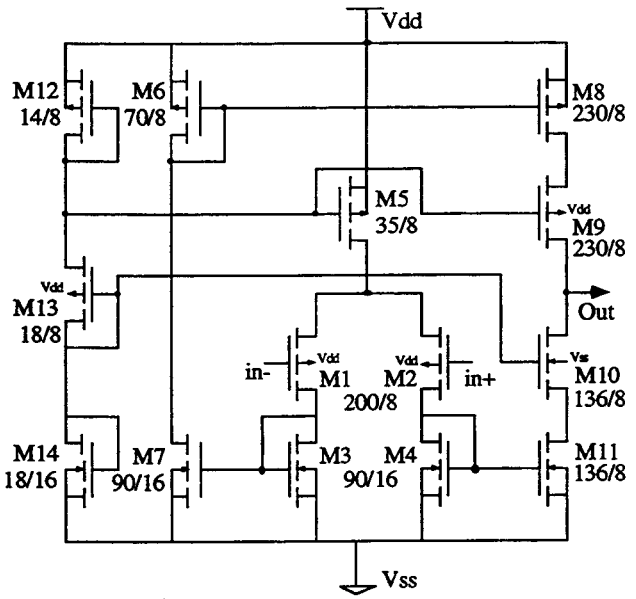


Figure 4: Schematic diagram of the low noise operational amplifier used in the readout circuit front-end.

The self-test DAC consists of three switched current mirrors driving an on-chip polysilicon resistor. The DAC output changes the amplitude of the clock when a control line is set. The temperature sensor is an on-chip oscillator with a temperature-dependent output frequency.

The bus interface circuitry consists of a series of shift registers that load input data, a laser-programmable 4b chip address comparator, command and element address decoders, and a 4b data latch accessed by a specific write command. The sensor bus data format provides a 5b element address to access up to 32 element addresses. On the interface chip three of these element address bits are used to select one of the six capacitive sensors or the on-chip temperature sensor while the remaining two bits are decoded as reference capacitor select lines.

TEST RESULTS

The chip has been fabricated using a standard 3µm, single metal, double poly p-well CMOS process at the University of Michigan. Figure 5 shows a photograph of the chip, which measures 3.2mm x 3.2mm and dissipates less than 2.2mW from a single 5V supply. Approximately 85% of the power dissipation is due to the quiescent currents of the five opamps integrated on the chip. Table 1 summarizes the chip specifications. Test results have shown full circuit functionality with an input capacitance resolution of less than 1fF and a maximum readout clock rate of 50kHz. Although carefully sized dummy switches are used in the switched-capacitor circuit, the resolution is still limited by the clock switching noise rather than kT/C and input amplifier flicker noise. Figure 6 shows the output of the charge integrator with a 6pF reference capacitor as the sensor capacitance is varied from 3pF to 8pF. Figure 7 illustrates the output of the readout circuit as the input capacitance changes in 0.5pF steps. Programmability of the gain is illustrated in Fig. 8, where the measured output is plotted for various digital gain settings.

The temperature dependence of the readout circuit offset and sensitivity has been measured over the range from 0 to 60°C. The circuit sensitivity is not significantly affected by its operating temperature and in all measurements the temperature dependence of the output voltage has been dominated by that of offset. The main source of offset temperature dependence is due to amplification of the first stage opamp offset in the second gain stage. The maximum temperature dependence of offset is equivalent to less than 0.16fF/°C input-referred capacitance. Note that the temperature dependence of the readout circuit does not adversely affect the performance of the overall microsystem since temperature compensation is performed in-module by the system microcontroller of the system.

The microsystem shown in Figure 1 has been implemented with two interface chips used to link a Motorola 68HC11 MCU to sensors for pressure, humidity, and acceleration [1]. The low power and capacitive readout capabilities of the interface chip make it a vital part of this battery-powered microsystem, which provides a prototype for long-life, wireless, distributed multi-parameter sensing systems.

CONCLUSIONS

In this paper a generic interface circuit for capacitive sensors has been presented. The chip can interface with a large variety of capacitive sensors, is highly programmable, supports communication over a standard sensor bus with a microcontroller, and supports self-test. The chip dissipates less than 2.2mW from a single supply, can operate with readout clocks as high as 50kHz and resolves input capacitance variations of less than 1fF.

ACKNOWLEDGMENT

The authors would like to thank Mr. Chingwen Yeh for the CMOS circuit fabrication. Also the help of the staff and students of the Center for Integrated Sensors and Circuits, the University of Michigan, is gratefully acknowledged. This work was supported by the Advanced Research Projects Agency

under contract JFBI 92-149, and in part by NSF-NYI grant #ECS-925 7400.

REFERENCES

1. A. Mason, N. Yazdi, K. Najafi, K.D. Wise, "A Low Power Wireless Microinstrumentation System for Environmental Monitoring", *Proc. Transducers 95*, Stockholm, Sweden, June 1995, pp. 107-110.

2. Y. Park and K.D. Wise, "An MOS Switched-Capacitor Readout Circuit for Capacitive Pressure Sensors", *Proc. IEEE Custom Circuit Conf.*, May 1983, pp. 380-384.

3. K. Watanabe and W.S. Chung,, "A Switched Capacitor Interface for Intelligent Capacitive Transducers", *IEEE Trans. on Instrumentation and Measurement*, Vol. IM35, No. 4, Dec. 1986, pp. 472-476.

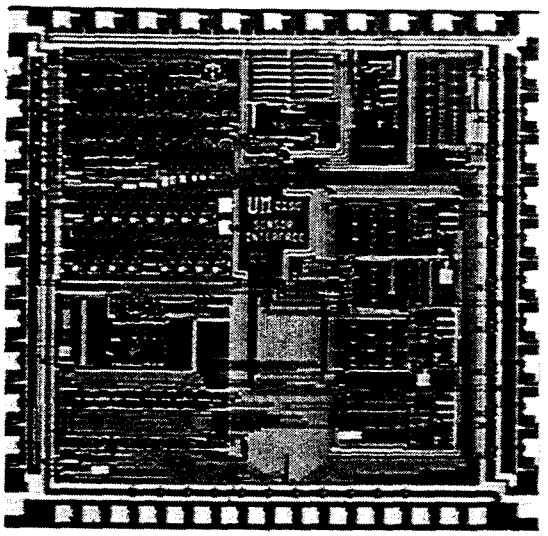


Figure 5: Photograph of the capacitive interface chip.

Table 1: Summary of the interface chip specifications.

Number of sensing channels	6 Capacitive + Temperature
Reference capacitor channels	3 internal/external, Internal capacitor programmable 0.15pF-8pF
Chip address	4 bits
Sensitivity 1 [mV/fF]	0.23-22.3
Sensitivity 2 [mV/fF]	0.23-50.9
Sensitivity 3 [mV/fF]	0.47-73.5
Resolution	< 1fF input capacitance
readout clock	< 50 kHz
DAC	3 bits for self-testing
Supply	Single 5V
Power dissipation	< 2.2 mW
Temperature Dependence of Offset	< 0.16 fF/°C (input referred capacitance variation)

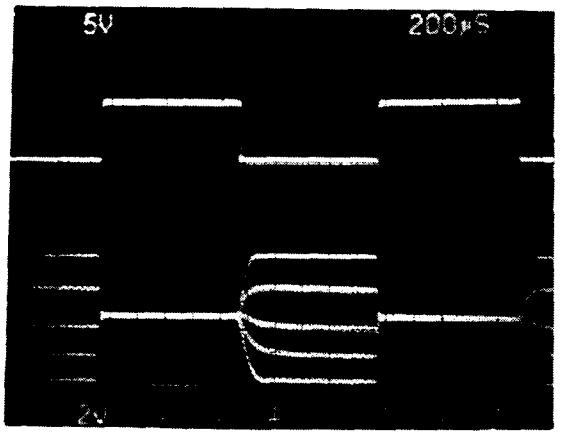


Figure 6: Upper trace: Reset control of the input charge integrator; Lower trace: Output of the charge integrator with a 6pF reference capacitor as the sensor capacitance changes from 3pF to 8pF.

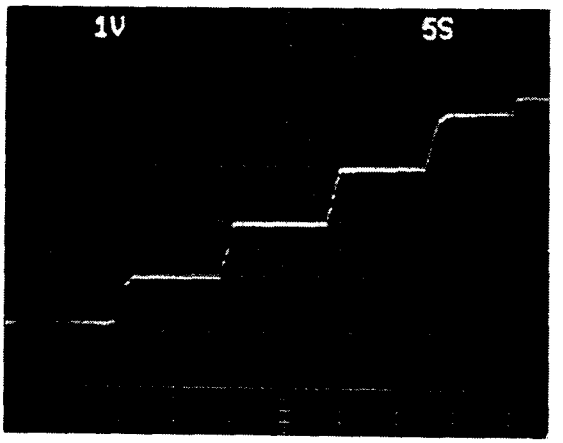


Figure 7: Readout circuit output as the input capacitance changes in 0.5pF steps; No laser trimming for gain adjustment was performed and the lowest second-stage gain is selected.

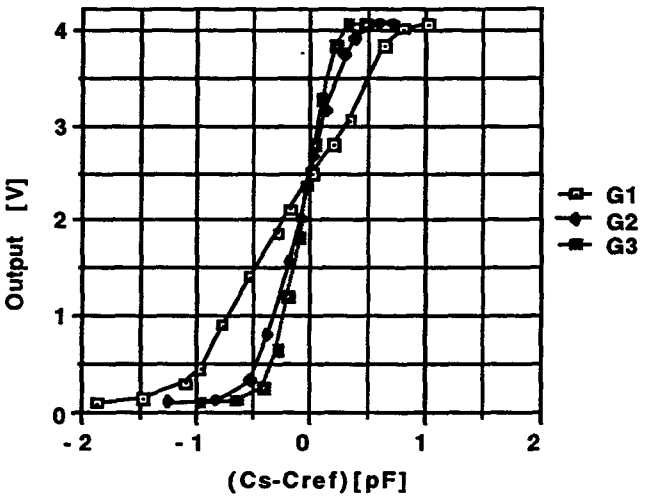


Figure 8: Measured output of the chip as a function of the input capacitance for different second-stage gain settings.

IMPROVED CANTILEVERS FOR AFM THERMOMECHANICAL DATA STORAGE

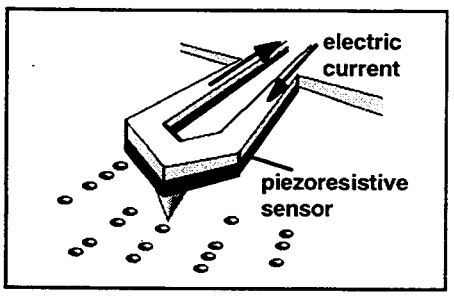
Benjamin W. Chui, Timothy D. Stowe, Yongho S. Ju, Hyongsok T. Soh, Stephen C. Minne, Kenneth E. Goodson, Calvin F. Quate, and Thomas W. Kenny
Departments of Electrical Engineering, Applied Physics, and Mechanical Engineering
Terman 551, Stanford University, California 94305-4021

H. Jonathon Mamin, Bruce D. Terris, Robert P. Ried, and Dan Rugar
IBM Research Division, Almaden Research Center,
650 Harry Road, San Jose, California 95120-6099

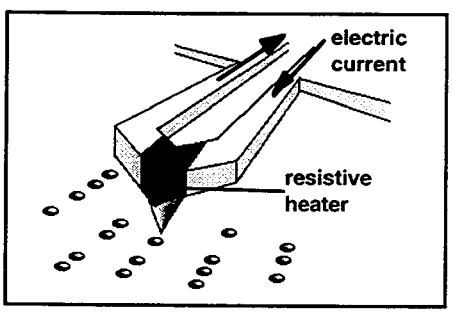
chui@leland.stanford.edu, mamin@almaden.ibm.com, kenny@cdr.stanford.edu

INTRODUCTION

In a previous report by Mamin *et al* [1], a new micro-mechanical approach to high-density data storage was described. In this technique, an atomic force microscope (AFM) tip is placed in contact with a rotating polycarbonate disk. A pulsed laser is used to heat the tip and thereby melt indentations into the disk, which has a glass transition temperature of 120-140 °C. Another laser is used to sense the deflection of the cantilever as the tip passes over the recorded indentations. Bit densities up to 30 Gbit/in^2 (50 times that of CD-ROM) have been demonstrated on a conventional AFM, and data rates up to 1.2 Mbit/s have been achieved on a rotating sample.



(a) Integrated piezoresistive sensor



(b) Integrated resistive heater

Figure 1: Integration of read/write elements onto single-crystal silicon cantilever for thermomechanical data storage.

One of the main problems with the existing read/write scheme is the need for external laser sources, which are bulky and require tedious alignment. As suggested in Figure 1, our ultimate goal is to develop an integrated read/write cantilever and eliminate the use of lasers. For this purpose we have developed single-crystal silicon cantilevers with piezoresistive elements for readback, based on the approach of Tortonese *et al* [2,3]. We have also made separate cantilevers with integrated resistive heaters for writing.

The ideal integrated cantilever will have to satisfy many conditions simultaneously. In order to read marks at densities of 20-50 Gbit/in^2, it must have a tip that is sharp on the 500-1000 Å scale. To achieve adequate signal-to-noise ratio, it needs a sensor capable of detecting a few angstroms of motion in the readback bandwidth. The cantilever must be soft—with a stiffness of 1 N/m or less—to allow for operation at loads below 10^{-7} N, which is necessary to avoid wear of the tip or sample. At the same time, the resonant frequency should be as high as possible, preferably in the MHz range. The combination of low stiffness and high frequency requires the cantilever to have low mass. For writing, an integrated element for heating the tip is required, and the thermal time constant should be as short as possible, preferably on the order of 1 μs.

In this work, we have not yet attempted to satisfy all these constraints simultaneously. Instead, we have focused on two areas. First, we have integrated piezoresistive sensors with cantilevers that are considerably softer than existing devices. With a cantilever of stiffness about 1 N/m, we have measured a deflection sensitivity of 1.6×10^{-2} Å/√Hz and a calculated resonant frequency of 250 kHz. To achieve this combination, we had to make considerably thinner cantilevers through some novel process development. Second, we have fabricated cantilevers with integrated electrical heaters, and have observed thermal time constants of roughly 40 μs. We have implemented both types of cantilevers on a rotating sample for demonstrating reading and writing without the use of lasers.

CANTILEVER FABRICATION

The fabrication process of the micro-cantilevers, shown in Figure 2, is adapted from reference [2]. The starting material is a silicon-on-insulator (SOI) wafer [4] with a 5-micron top silicon layer. In Figure 2a, an SF_6 plasma etch is used to undercut an oxide-resist mask to form a blunt tip, which is then sharpened by low-temperature oxidation (Figure 2b). In Figure 2c the cantilever itself is patterned and a 1000 Å oxide layer is grown over it. After this, boron is implanted through the oxide (Figure 2d) and subject to rapid thermal annealing. (Our process differs

0-9640024-1-8 $3.00 © 1996 TRF

from reference [2] in that the implant comes *after* the oxidation step—the significance of this will be explained more fully in the next section.) In Figure 2e, contact vias are etched through the oxide and aluminum is deposited and patterned to form electrical connections to the cantilever. Finally, a backside TMAH etch releases the cantilever (Figure 2f).

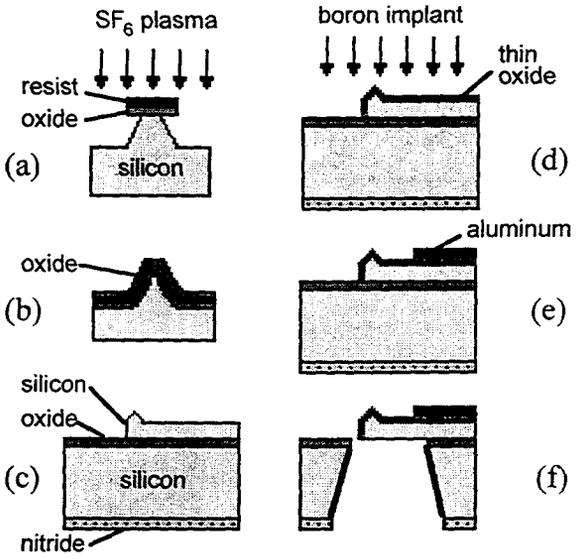
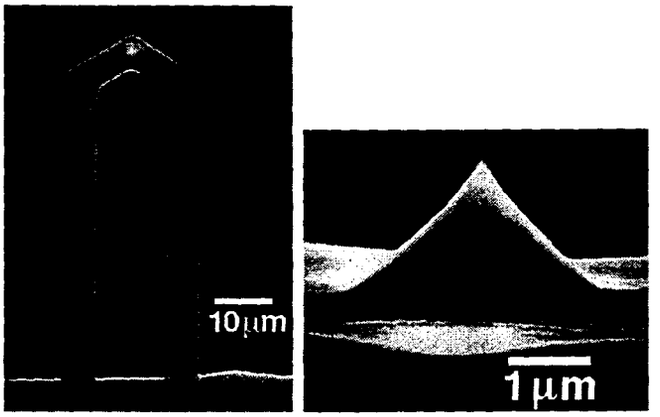


Figure 2: Fabrication process of 1-μm thick cantilevers

Figure 3 below shows an SEM micrograph of a fabricated cantilever and a close-up of the tip. The tip is seen to be very sharp, with a radius of curvature below 300Å. Such a tip is quite suitable for reading and writing marks on a 1000 Å scale.



(a) (b)

Figure 3: SEM images of (a) released cantilever and (b) tip

INTEGRATED PIEZORESISTIVE SENSORS

As explained above, it is necessary for the cantilever to be reasonably soft so that low forces can be maintained on the tip. To achieve low stiffness without sacrificing frequency, it is necessary to reduce the cantilever thickness considerably from the 2-4 μm typically used in the process of Tortonese [2]. We have chosen to make 1-μm thick cantilevers in this case. Implementing a piezoresistive sensor in such a thin cantilever is difficult, however. The reason is that the piezoresistor must be confined to a fraction of the cantilever thickness. Otherwise the compressive stress experienced by the piezoresistor on one surface cancels the tensile stress on the other, reducing the overall sensitivity.

We have developed a process for implantation and annealing of boron that prevents its diffusion throughout the cantilever thickness. The goal is to avoid post-implant high-temperature steps such as oxidation or annealing at 1000°C, which are known to cause significant dopant diffusion. Our process therefore calls for the 1000Å oxide to be grown first at 1000 °C, followed by a 40-keV boron implant of $5\times10^{14}/cm^2$ through the oxide. The implant is activated by a 10-second rapid thermal anneal (RTA) step at 1000 °C and a 40-minute furnace anneal at 800 °C. A spreading resistance measurement of the implant profile [5] reveals a junction depth of roughly 0.4 μm below the cantilever surface.

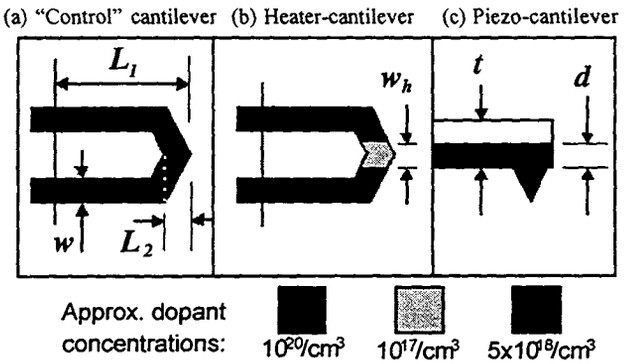


Figure 4: Diagram (not drawn to scale) showing cantilever dimensions, as well as differences between "control" cantilevers, heater-cantilevers, and piezoresistive cantilevers. Typical dimensions for our present design are L_1 = 50-300 μm, L_2= 7 μm, w = 8 μm, w_h = 8 μm, and t = 1 μm.

A detailed calculation of the piezoresistive sensitivity of a cantilever can be found in reference [2]. To simplify the calculations, we have made the following assumptions:

(a) The width of the transverse part of the cantilever (where the tip is located) is small compared with the length of the cantilever (i.e. L_2 is much smaller than L_1). (Relevant cantilever dimensions are illustrated in Figure 4.)

(b) The piezoresistor doping is uniform through part of the cantilever thickness up to a step junction at depth d.

(c) The piezoresistor is uniformly distributed across the entire length of the cantilever. This reduces the sensitivity by a factor of 2 from the case where the resistor is concentrated at the base of the cantilever.

Under these conditions, the sensitivity is given by

$$\frac{\Delta R}{R} = \frac{3\,\pi_L\,E\,t\,\beta\,y}{4\,L_1^2} \qquad (1)$$

where $\Delta R/R$ is the relative change in cantilever resistance,
π_L is the longitudinal coefficient of piezoresistivity of silicon,
E is Young's modulus of silicon,
t is the thickness of the cantilever,
L_1 is the length of the cantilever, and
y is the displacement of the cantilever tip.

Here β is an efficiency factor equal to 1 for the case where the piezoresistor is confined just to one surface of the

cantilever. If assumption (b) is satisfied, the value of β is given by $\beta = (1 - d/t)$, where d is the depth of the piezoresistive layer. A thicker piezoresistive layer implies a lower value of β.

Taking $\pi_L = 5.7\times10^{-10}$ m^2/N (in the [110] direction for boron-doped silicon at 5×10^{18}/cm^3) [2], $E = 1.7\times10^{11}$ N/m^2 [2], $t = 1$ μm, and $d = 0.4$ μm (implying $\beta = 0.6$), we can calculate the sensitivity for different cantilever lengths. The theoretical sensitivity, which varies as the inverse of the square of the length, has been calculated with no fitting parameters and is shown as the dotted line in Figure 5.

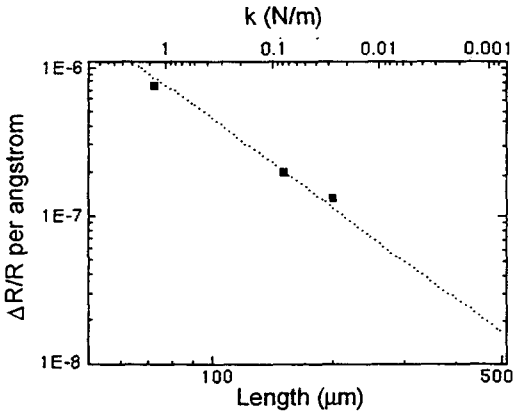


Figure 5: Sensitivity of piezoresistive cantilevers of various lengths. The dotted line represents the theoretical sensitivity calculated with no fitting parameters.

We have tested cantilevers of various lengths to determine their sensitivity. This was done by placing the cantilever on a piezoelectric actuator (which is part of a atomic force microscope) and modulating the cantilever support with the tip in contact with a fixed surface. Typical oscillation amplitudes were on the order of 1000 Å. The response was measured with a simple full bridge circuit based on a Burr-Brown INA103 instrumentation amplifier. This amplifier has an input noise of 1 nV/√Hz and a gain-bandwidth product of 100 MHz. We operated with a bridge supply voltage of 5 V (i.e. 2.5 V across each resistor) and a gain of 100. Since the cantilever resistance was 5-30 kΩ (depending on length and doping concentration), the power dissipated in the cantilever was 1 mW or less.

The inferred values of $\Delta R/R$ per angstrom are plotted in Figure 5. Test results show an excellent agreement with theory, with $\Delta R/R$ up to 0.75 ppm/Å. In addition, these levels of sensitivity compare very favorably to those reported for other piezoresistive sensors, while our device has much lower stiffness. Table 1 shows a comparison between one of our cantilevers and one from ref. [2]. It is seen that while the two cantilevers have about the same $\Delta R/R$ per Å, our cantilever exhibits a more favorable force figure of merit due to its lower spring constant.

	L_1 (μm)	t (μm)	measured ΔR/R per Å	k [a] (N/m)	MDD [b] (Å)	FFOM [c] (nN)
This work	75	1	7.5×10^{-7}	1.5	0.5	0.075
Ref.[2]	170	3.8	8.5×10^{-7}	16.4	0.2	0.33

[a] Spring constant
[b] Minimum detectable deflection in the 10 Hz-1 kHz bandwidth
[c] Force figure of merit (= MDD × k)

Table 1: Comparison of piezoresistive cantilevers from this work and ref. [2]

Figure 6 shows the noise spectrum of a 75-μm long cantilever. The observed noise floor of 1.6×10^{-2} Å/√Hz is very close to the Johnson-limited noise floor of 1.3×10^{-2} Å/√Hz. The $1/f$ noise performance is quite good: with its knee at about 200 Hz, the integrated $1/f$ noise from 1 Hz to 200 Hz is only about 0.5 Å.

The inset of Figure 6 shows the response of the cantilever to a 100 Å modulation applied with the AFM. The bandwidth used was 1 Hz-100 kHz, within which the minimum detectable displacement was less than 10 Å. Since data pits on the polycarbonate disk are typically 100-200 Å in depth, piezoresistive readback of real data pits should be possible at this bandwidth.

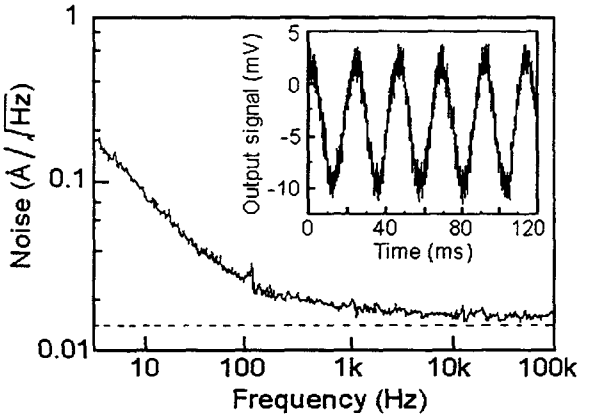


Figure 6: Noise spectrum of 75-μm piezoresistive cantilever. Theoretical Johnson noise floor is shown as dotted line. Figure 6 (Inset): Cantilever response to 100Å modulation.

Figure 7 shows a readback signal for one of our 75-μm piezoresistive cantilevers operated on a spinning test sample with 600-nm deep grooves. The bandwidth in this case was 300 kHz, and the linear velocity was approximately 120 mm/s. The waveform signal-to-noise ratio is excellent on these 600-nm deep grooves, although of course the feature depth will be much smaller for real data pits. Still, this demonstrates the basic feasibility of integrated piezoresistive readback on a rotating sample.

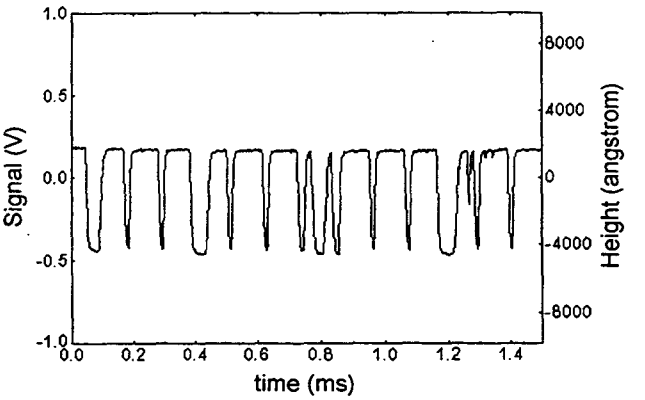


Figure 7: Readback signal obtained with 75-μm piezoresistive cantilever operated on spinning test sample with 600-nm deep grooves.

INTEGRATED HEATERS

The present cantilevers are made of single-crystal silicon, which has the advantage that a wide range of resistivity is accessible through appropriate doping. This is very convenient for electrical heating, as it allows us to make the heating element and the leads of the same material. Another advantage is that the cantilevers are intrinsically low stress and have high thermal conductivity. Other approaches to micromachined heating elements include polysilicon [6] and nickel silicide [7].

We create the heaters by implanting the cantilever with boron except for a gap at the end which is several microns wide (denoted by w_h in Figure 4). This lightly doped region forms the resistive heater. An electrical pulse applied to the cantilever will cause significant power dissipation in the heater region, giving rise to localized heating. After the pulse has ended, the heat is dissipated mainly by conduction down the legs to the support. The heavy implant concentration is about $10^{20}/cm^3$ while the "heater" implant is about $10^{17}/cm^3$. These boron implants are the same type as that used to form the piezoresistors, and can thus be carried out during the same process.

An analysis of the thermal properties of these cantilevers will be presented in the next section. We have demonstrated basic functionality of these cantilevers, as shown by the AFM micrograph in Figure 8. Sample data tracks were written with the resistive heater on a rotating polycarbonate sample. The writing was very reliable, and different-sized marks could be written by varying the pulse conditions. Typical pulse parameters were 16 V, 20 µs, with a 11 kHz repetition rate.

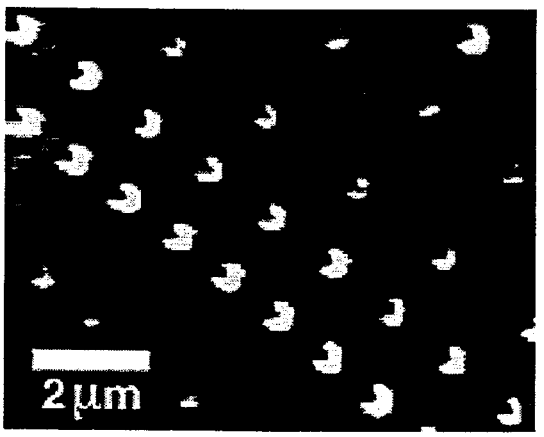


Figure 8: AFM image of sample data tracks

THERMAL AND ELECTRICAL ANALYSIS

We have measured the temperature variation of resistance of a 200-µm long cantilever with an 8-µm wide heater region. The cantilever chip was placed on a temperature-controlled chuck, and an HP4155A semiconductor parametric analyzer was used to apply a low test voltage across the cantilever and to measure the resulting current. The calculated resistance is shown in Figure 9, curve (a). Note that the observed resistance is the total resistance of the heater region and the cantilever legs. To find the resistance of the legs alone, we measured a "control" cantilever that has no heater region—what would have been the heater region is now doped as heavily as the legs and contributes very little to the overall resistance (see Fig. 4). The resistance of such a "control" cantilever therefore corresponds to that of the legs, and its measured value at different temperatures is shown in curve (b) of Figure 9. Subtracting (b) from (a) gives curve (c), which represents the resistance of the heater region alone. Curve (c) will be used in conjunction with another measurement described below (see Figure 10) to help estimate the transient temperature behavior of the heater.

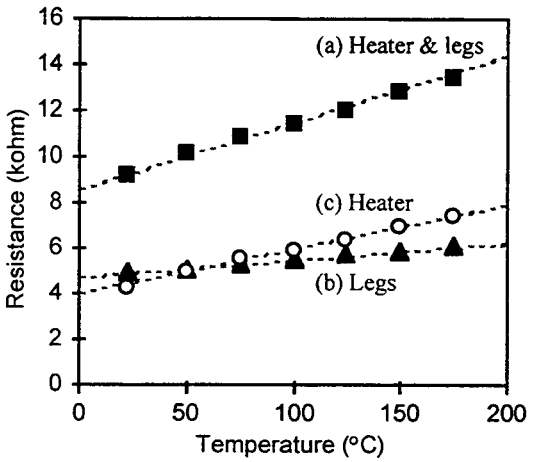


Figure 9: Resistance vs. temperature of 200-µm cantilever with 8-µm heater. Curve (a) shows total resistance of cantilever; curve (b) shows resistance of cantilever legs; curve (c) shows derived resistance of heater region alone (i.e. difference between (a) and (b)).

Figure 9 implies positive temperature coefficients of resistance for the legs and the heater on the order of 0.1%/°C and 0.3%/°C respectively. One would therefore expect that when a voltage pulse is applied to the cantilever, the resulting current would decrease as the cantilever heats up and increases in resistance. After the pulse, the resistance would gradually decrease to its room-temperature value as the cantilever cools. If a small DC test voltage is present across the cantilever after the pulse has ended, the resulting current would be expected to rise and approach a final room-temperature value as the cantilever cools. By measuring the instantaneous current during and after a voltage pulse, we can infer the transient temperature behavior of a heater-cantilever. The cooling rate of the heater corresponds to the thermal time constant and determines how fast successive marks can be written, i.e. the data rate.

In one experiment, a 11 V, 5 µs heating pulse was applied across a cantilever similar to the one used to obtain Figure 9. The pulse was superimposed on a 1 V DC offset. The value of 1 V was chosen so as not to cause noticeable heating. The applied pulse is shown in Figure 10a. The instantaneous current that flowed through the cantilever was determined from the voltage across a 100 Ω series resistor, and the oscilloscope trace is reproduced in Figure 10b. As predicted, the current decayed during the pulse, indicating a rise in heater temperature. The time variation of the heater resistance can be derived, and the heater temperature can be estimated using the data in Figure 10b as well as the resistivity-versus-temperature data in Figure 9, curve (c). The resulting estimate is shown in Figure 10c. This type of electrical measurement has previously been used for characterizing micromachined elements [6] as well as commercial AFM cantilevers [8].

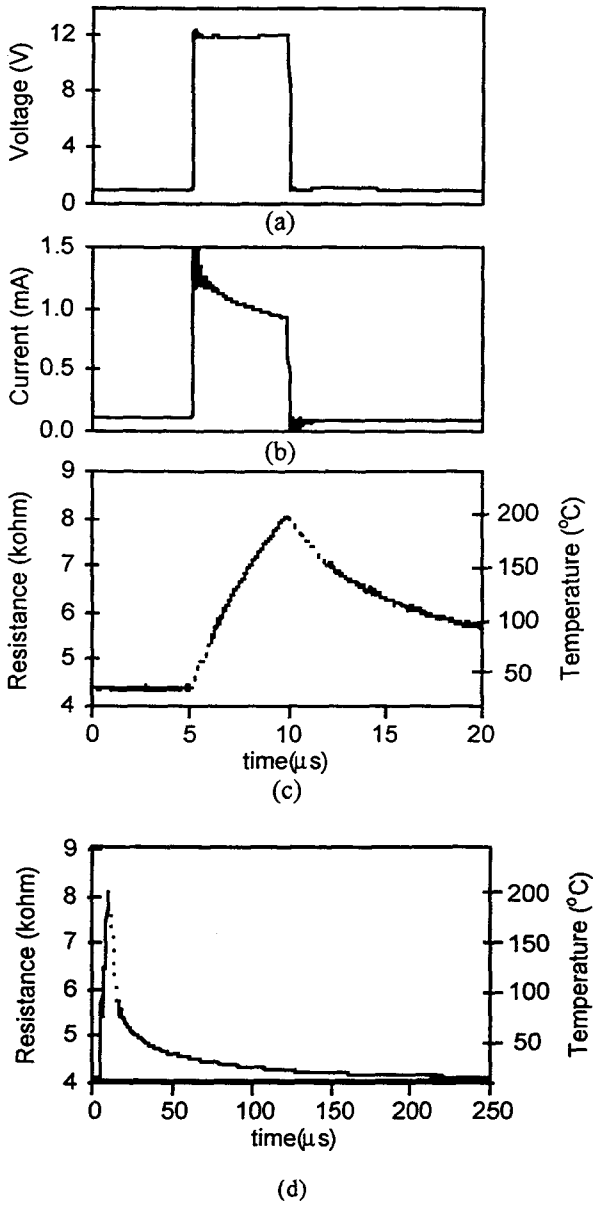


Figure 10: Electrical measurements for estimating heater temperature during heating pulse. Curve (a): Applied voltage pulse (amplitude is 12 V at time 5–10 μs and 1 V at all other times). Curve (b): Measured current determined from voltage across 100–Ω series resistor. Curve(c): Derived heater resistance and temperature, on 20-μs time scale. The sections of the curve at 5-6 μs and 10-12 μs represent interpolated data, since actual data was obscured by excessive ringing. Curve(d): Derived heater resistance and temperature on longer time scale of 250 μs, showing cooling time constant. Dotted lines show interpolated data for curves (c) and (d).

To investigate the cooling time constant, we measured the current that flowed after the pulse as a result of the 1 V offset voltage. The derived heater resistance and temperature is indicated in Figure 10, curve (d). As expected, the resistance started out from a relatively high value just after the pulse has ended and, as the heater cooled, gradually decreased towards its room-temperature value. It is seen that the thermal time constant is about 40 μs, compared with 350-450 μs for a commercial silicon cantilever without a localized heater **[8]**.

Unfortunately, the circuitry for measuring the current produces excessive ringing for several microseconds after the pulse, making it difficult to extract meaningful data during that period. This calls for a supplementary method of measurement that provides adequate time resolution under such conditons. One possible choice is laser thermometry.

Laser thermometry makes use of the fact that the optical reflectance of most materials varies with temperature. While this type of measurement is most commonly used with metals with up to picosecond-level resolution, in some cases it has been adapted for investigating semiconductor devices **[9]** such as SOI power transistors **[10]**. In our experimental setup, a low-power laser beam is focused on a heater-cantilever, and the reflected power is measured at different points in time. The results are used to estimate the heater temperature and the thermal time constant.

Because of the small laser spot size (on the order of 3 μm), laser thermometry will allow us to investigate the spatial distribution of temperature along the cantilever with sufficient time resolution to follow the propagation of the heat pulse. In future experiments, we will use this technique to confirm thermal models of heat propagation in these cantilevers.

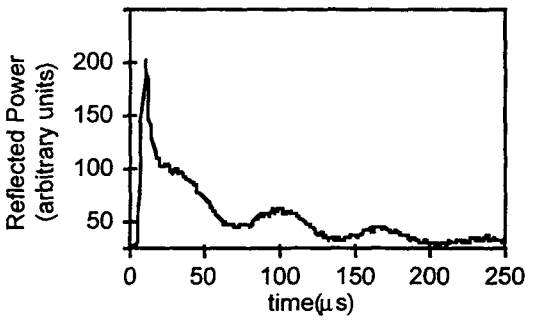


Figure 11: Reflected laser power from heater-cantilever, under same conditions as Figure 10.

Figure 11 shows the reflected laser power from a location near the end of the cantilever in response to a heating pulse. This curve shows two distinct phenomena: an oscillation at 14 kHz, and a decay with a time constant of about 40 μs. We believe that the overall decay is related to the temperature coefficient of the reflectivity of the cantilever; the time constant for this decay is consistent with the resistance measurements. In addition, we believe that the oscillation is due to a thermally induced vibration; the frequency of this oscillation is consistent with the resonant frequency of the cantilever. The cantilever has a 1000-Å oxide layer on top of a 1-μm silicon layer, so heating ought to produce a differential thermal expansion.

To confirm this hypothesis, we have used a vibrometer **[11]** to study the pure mechanical behavior of a cantilever in response to a heating pulse. In this technique, a laser interferometer is used to measure the motion of the cantilever tip with angstrom-level resolution. The output of the vibrometer operated in velocity mode is shown in Figure 12. A decaying sinusoidal oscillation is observed, confirming the mechanical origins of the oscillation in Figure 11. From Figure 12 the initial amplitude of the vibration can be estimated at roughly 350 Å, with a frequency of about 15 kHz. The difference in observed frequency between Figures 11 and 12 is most likely due to a slight mismatch in the length and thickness of the two cantilevers being measured.

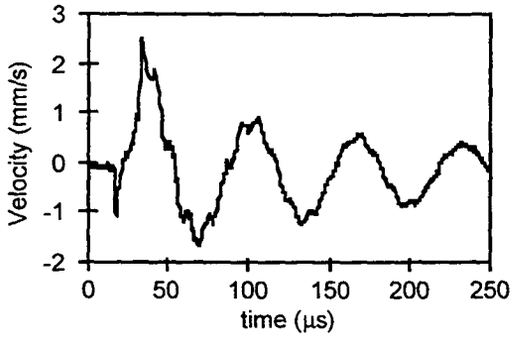


Figure 12: Vibrometer output signal showing mechanical vibration of cantilever tip.

As supplementary techniques to electrical measurements, laser thermometry and vibrometry make it possible to characterize the behavior of cantilevers in detail. Laser thermometry can provide spatially and temporally resolved measurements of cantilever temperature; laser vibrometry can do the same with regard to cantilever motion. Together, these two types of measurement will be useful for the development of improved cantilevers for thermomechanical data storage.

CONCLUSIONS

In this report, we have described two novel capabilities of AFM cantilevers. First, we have worked to add integrated sensing functionality onto low-stiffness cantilevers with integrated tips. A new process was developed which allows for considerably thinner cantilevers than previously possible, and should be scaleable to even thinner cantilevers. The measured sensitivity is in excellent agreement with predictions, so that we can calculate with confidence the type of cantilever needed to achieve a given level of performance. The current cantilevers are adequate to detect 100-200 Å deep features in a roughly 100 kHz bandwidth. They also show excellent $1/f$ noise behavior.

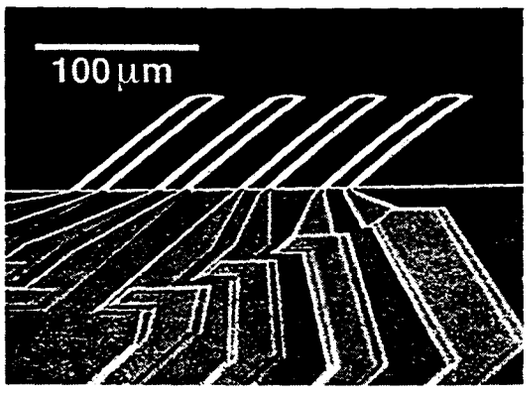


Figure 13: SEM image of fabricated cantilever array

A heater process was also developed on separate cantilevers which allows for thermal writing without the need for a laser. Electrical and laser thermometry measurements were made which indicate for our 200-μm heater-cantilevers a thermal time constant of about 40 μs, or 10 times shorter than for commercial silicon cantilevers. One of the advantages of micromachining is that the components are batch fabricated, so that arrays of tips can be easily made. Figure 13 shows an array of four cantilevers fabricated on the same wafer as the other cantilevers. The operation of cantilevers in parallel has already been demonstrated in AFM microscopy [12]. Ultimately, the same concept may be used in AFM data storage to improve the level of performance.

ACKNOWLEDGMENTS

We would like to thank J.D. Plummer of Stanford and L.S. Fan of IBM for providing valuable insight into the electrical characteristics of the heater-cantilever. We are also grate[illegible]o M. Tortonese of Park Scientific Instruments for helpful discussions, and to O.W. Käding for his work on the laser thermometry setup. We would like to acknowledge the support of an IBM Cooperative Fellowship, the National Science Foundation CAREER Program (ECS-9502046), the National Science Foundation Instrumentation for Materials Research Program (DMR-9504099), and the Charles Lee Powell Foundation. The work of H.J. Mamin and R.P. Ried is supported in part by ARPA contract number DABT63-95-C-0019. We would finally like to thank the National Science Foundation for its generous support of travel expenses for this conference.

REFERENCES

1. H.J. Mamin, L.S. Fan, S. Hoen, D. Rugar, "Tip-based data storage using micromechanical cantilevers," *Sensors and Actuators A, 48* (1994), pp. 215-219.
H.J.Mamin, D.Rugar, "Thermomechanical writing with an atomic force microscope tip," *Appl. Phys. Lett., 61*(1992), p.1003-5.
2. M. Tortonese, Ph.D thesis, 1994 (unpublished)
3. M. Tortonese, R.C. Barrett, and C.F. Quate, "Atomic resolution with an atomic force microscope using piezoresistive detection," *Appl. Phys. Lett., 62,* 8 (1993), pp. 834-6.
4. SOI wafers purchased from IBIS Corporation, phone (408) 662-9100.
5. Spreading resistance profile measurement by Solecon Laboratories, 2241 Paragon Drive, San Jose, California 95131.
6. C.H. Mastrangelo, J.H.-J. Yeh, R.S. Muller, "Electrical and optical characteristics of vacuum-sealed polysilicon microlamps." *IEEE Trans. on Elec. Dev., 39*, 6 (1992), pp. 1363-75.
7. J.H. Das, N.C. MacDonald, "Micromachined field emission cathode with an integrated heater," *J. Vac. Sci. Technol. B, 13,* 6 (1995), pp. 2432-5.
8. H.J. Mamin, "Thermal writing using a heated atomic force microscope tip," submitted to *Appl. Phys. Lett.* (1996).
9. P.W. Epperlein, "Micro-temperature measurements on semiconductor laser mirrors by reflectance modulation: A newly developed technique for laser characterization," *Jpn. J. Appl. Phys., 32,* 12A (1993), pp. 5514-22.
10. K.E. Goodson, Y.S. Ju, O.W. Käding, "Short-timescale thermal mapping of semiconductor devices," submitted to *IEEE Elec. Dev. Lett.* (1996).
11. Polytec Laser Vibrometer OVF-501-0, Polytec P.I. Inc., 3152 Redhill Suite 110, Costa Mesa, California 92626.
12. S.C. Minne, S.R. Manalis, C.F. Quate, "Parallel atomic force microscopy using cantilevers with integrated piezoresistive sensors and integrated piezoelectric actuators," *Appl. Phys. Lett. 67,* 26, (1993), pp. 3918-20.

ULTRASENSITIVE VERTICAL FORCE PROBE FOR MAGNETIC RESONANCE FORCE MICROSCOPY

Timothy Stowe, Kevin Yasumura, Thomas Kenny
Departments of Applied Physics and Mechanical Engineering,
Terman 551, Stanford University, California, 94305-4021

David Botkin, Koichi Wago, Daniel Rugar
IBM Research Division, Almaden Research Center,
San Jose, California, 95120-6099

ABSTRACT

This paper reports on ultrathin single crystal silicon cantilevers with integrated in-plane tips for use in magnetic resonance force microscopy. Single crystal silicon cantilevers 400-1700Å thick and 100-150 μm long have been fabricated from <100> oriented silicon-on-insulator wafers. These cantilevers are capable of detecting forces smaller than 10^{-16} N in a one hertz bandwidth and have Q values higher than 10,000 at pressures below 10^{-5} Torr. In addition to these extremely thin flexural mode cantilevers, 900 Å thick silicon nitride double-torsional mode oscillators have been fabricated as an alternative route to achieving high Q resonator performance.

INTRODUCTION

Atomic force microscope techniques have been adapted for the detection of a wide variety of surface phenomena [1-2]. Recently, a new technique known as magnetic resonance force microscopy (MRFM) [3] has been established as a viable method for performing three-dimensional magnetic resonance imaging with micrometer-scale spatial resolution. Both electron spin resonance (ESR) and nuclear magnetic resonance (NMR) have been demonstrated [4-8]. By improving the micromachined cantilevers used for MRFM, we hope to extend the sensitivity and the spatial resolution to detect and image single electron spins with nanometer-scale spatial resolution.

Figure 1 illustrates a future experiment for the detection of individual electron spins. The experimental setup shown is housed on a vibration isolation stage inside a vacuum can placed within a liquid helium dewar. Electron spins inside the sample are excited by a frequency modulated gigahertz magnetic field originating from a coil or microwave resonator. The frequency modulation periodically flips the direction of electron spins in resonance through a process known as adiabatic inversion [5]. Since the tip of the cantilever is coated with a thin ferromagnetic film there is a magnetic dipole-dipole interaction between the tip and electron spins. The oscillating spin direction excites the cantilever though this interaction force and the resulting motion is detected using a fiber interferometer aligned onto the cantilever paddle [9].

The interaction force which drives the cantilever is given by

$$F_x = \mu_z \frac{\partial B_z}{\partial x}, \qquad (1)$$

where μ_z is the magnetic moment of the electron spin and $\partial B_z / \partial x$ is the magnetic field gradient from the ferromagnetic

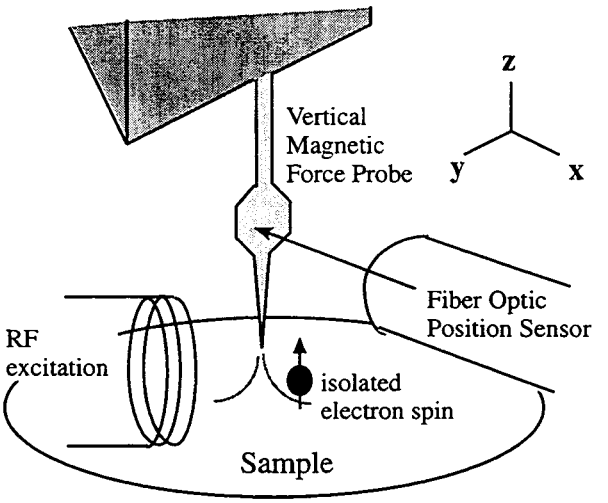


Figure 1: Experimental magnetic resonance force microscope setup.

tip. For a NiFe tip with a 500 Å radius of curvature placed 150 Å above the sample surface, the interaction force is on the order of 10^{-16} N for a single electron. As described below, detecting such a minute force requires an extremely soft and high Q cantilever. Such a cantilever operated in the standard AFM orientation, parallel to the sample substrate, would be highly susceptible to "snap down" since van der Waals forces can be as high as 10^{-9} N. However, if the cantilever is oriented perpendicular to the sample as shown in Figure 1 the snap down problem is reduced because the cantilever's stiff axis is then normal to the substrate [10]. In order to implement this geometry, it has been necessary to develop an in-plane tip with a radius of curvature under 500 Å. In addition, the magnetic moment at the tip needs to point along the length of the cantilever. This requirement can be met by using tips with a narrow apex angle because this provides favorable shape anisotropy for producing magnetization parallel to the length of the cantilever.

The force resolution is the most difficult requirement imposed by this experiment. The force resolution of a micromachined cantilever in non-contact mode is limited by thermo-mechanical noise [11]. In thermal equilibrium, the thermo-mechanical displacement for a single cantilever mode is related to temperature by the equipartition theorem

$$\frac{1}{2}k(\Delta x)^2 = \frac{1}{2}k_B T, \qquad (2)$$

0-9640024-1-8 $3.00 © 1996 TRF

Table I, Force resolution comparison of cantilevers at 10 K

Cantilever Description	Length (μm)	Width (μm)	Thickness	Frequency (kHz)	Q factor	Spring constant (10^{-3}N/m)	Force Resolution (10^{-17} N/√Hz)
thick silicon [13]	450	45	1.5 μm	9.8	200,000	70	4.3
silicon nitride	60	10	800 Å	30	16,000	0.8	1.2
target cantilever	110	5	500 Å	5.2	50,000	0.022	0.26

where k is the spring constant, k_B is Boltzmann's constant, and T is temperature. At 5 K, a cantilever with a spring constant of 10^{-3} N/m has thermo-mechanical displacement noise of 3 Å. By comparison, an interferometer can measure displacements smaller than 0.01 Å/√Hz [9]. Therefore, the thermo-mechanical noise, and not the measurement device, limits the minimum force that can be resolved from the cantilever's motion.

Taking into account the spectral distribution of the noise, the minimum detectable force in a bandwidth ΔB is

$$F_{min} = \sqrt{\frac{2kk_BT\Delta B}{\pi Q f_o}}, \qquad (3)$$

where Q is the quality factor and f_o is the resonance frequency. Not all of the parameters in equation 3 can be optimized independently in order to improve the force resolution. Our MRFM experiments require that the cantilever resonance frequency fall between 5-15 kHz. The low frequency limit was chosen to reduce the influence of environmental noise while the high frequency limit was chosen to optimize the magnetic resonance signal lifetimes during cyclic adiabatic inversion [13]. This finite signal lifetime, on the order of 0.1-1 seconds, limits the minimum bandwidth we can use for a single measurement. Given the constraints on the resonance frequency, the force resolution can be improved by minimizing the ratio of k/Q for a given cantilever geometry. For a cantilever thickness t, and width w, the length can be chosen such that the resonance frequency is fixed at 10 kHz. The dependence of the minimum detectable force on the cantilever geometry and Q can then be expressed as

$$F_{min} \propto \sqrt{\frac{wt^{3/2}}{Q}}. \qquad (4)$$

Clearly, the force resolution can be improved by reducing the cantilever thickness and width while maintaining high Q.

In this work we address the issue of making very thin cantilevers that maintain high Q values. Previous efforts [12] have shown that it is difficult to fabricate extremely thin cantilevers ($t \leq 2000$ Å) that have high Q values. One approach to obtaining a high Q is to use single crystal silicon rather than an amorphous material such as silicon nitride. Another benefit of using single crystal silicon is that internal stress gradients, that can arise in amorphous materials such as silicon nitride, are a concern when performing low temperature experiments. These stress gradients can cause curvature in a cantilever as it is cooled down to liquid helium temperatures, thereby causing an undesirable reduction in the fiber optic signal.

Previous measurements by Hoen *et al.* [12] on silicon nitride cantilevers below 600 Å thick have demonstrated Q values below 3,000 at room temperature and at 10 K. More recently we have measured Q values up to 16,000 for an 800 Å thick silicon nitride cantilever at 10 K. In contrast, we have repeatedly measured Q values above 50,000 on commercially available 1.5 μm thick single crystal silicon cantilevers at room temperature similar to those values reported for polysilicon [14-15]. At temperatures below 10 K, the Q of these single crystal silicon cantilevers climbs to 200,000 [13]. Despite this high Q, the actual force sensitivity of these cantilevers is limited by their relatively high stiffness.

The first two rows in Table I show a comparison of force resolution for these cantilevers and illustrate the tradeoffs between cantilever Q and cantilever stiffness. The third row represents cantilever performance that we wish to eventually achieve. Given the need for thin, high Q, low stress cantilevers we have investigated using bonded silicon on insulator (SOI) wafers for fabricating submicron thick single crystal silicon cantilevers.

An additional route to improving the Q performance of thin oscillators is to modify their geometry. Previous work has shown that double torsional oscillators have extremely high Q values. In this geometry clamping losses to the substrate are reduced [16-18]. We have modified the basic geometry of a double-torsional oscillator so that it can be used as a force probe. Because of the difficulty involved in fabricating extremely thin silicon structures, we investigated the Q performance of a 900 Å thick double-torsional silicon nitride oscillator and compared its performance with silicon nitride cantilever beams having similar frequencies and spring constants.

SILICON CANTILEVER PROBES

The cantilever processing begins with <100> oriented SOI wafers having a top silicon layer thickness of 2000-3000 Å. Fabrication proceeds according to the processing sequence outlined in Figure 2. The top silicon layer is thinned down using a wet oxidation (Figure 2a). This top oxide is then removed in a buffered oxide etch (BOE) and the silicon layer is patterned using an SF_6 plasma etch (Figure 2b). The details of how lateral tips are formed are described in the next section. Once the cantilevers have been patterned, polyimide is spun on the top side to protect them against scratching and as a final dry release layer. Openings in the backside oxide are patterned and the exposed silicon is etched anisotropically in 20% tetramethylammonium hydroxide (TMAH) (Figure 2c). TMAH was used rather than KOH because of its higher selectivity against etching the oxide mask and buried oxide layer which serves as a backside etch stop. The final release consists of removing the buried backside oxide in BOE and the remaining polyimide layer in an oxygen plasma (Figure 2d). Various cantilevers ranging in thickness from 400 Å to 1700 Å and up to 200 microns long were made without any noticeable curvature.

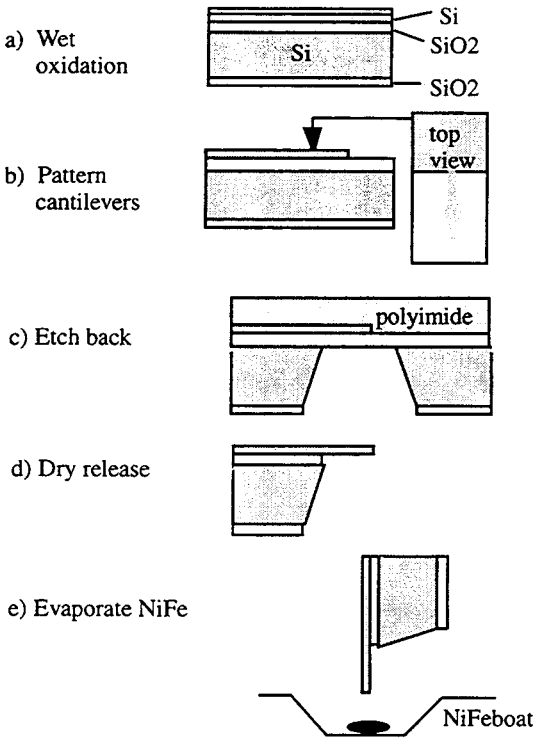


Figure 2: Processing Sequence

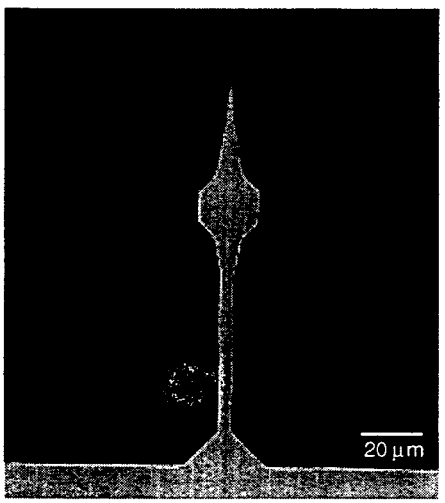


Figure 3: Finished "Spearhead" cantilever 120 µm long and 0.17 µm thick.

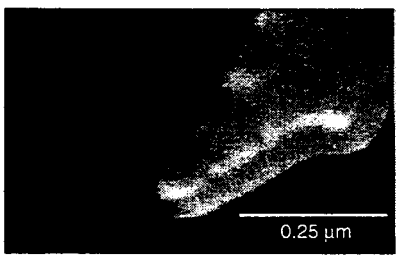


Figure 4: A cantilever tip formed by lateral oxidation sharpening.

In addition to the dry oxygen plasma release process, we have successfully experimented with BOE wet release processes without the top polyimide layer. In this case, the final release step consists of critical point drying [19] to avoid meniscus forces which lead to surface adhesion of the cantilevers.

FABRICATION OF IN-PLANE TIPS

Three possible methods were considered for lateral tip formation: lateral oxidation sharpening, convex corner undercutting, and edge intersection from a double patterning procedure (Figure 5). The lateral oxide sharpening method is similar to that used by Yao and MacDonald [20-21] for out of plane tips. First a layer of silicon nitride greater than 1000 Å thick is grown over the silicon and the cantilever geometry is

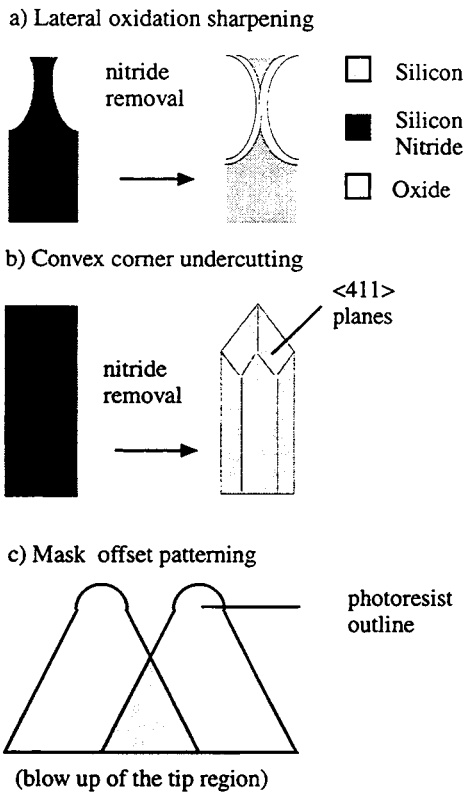


Figure 5: Three Lateral Tip Processes

patterned with a tip geometry shown in Figure 5a. The silicon nitride layer serves as an oxidation mask. Next a 900°C oxidation is used to laterally oxidize the underlying silicon, pinching off the silicon to form a tip. Finally, the nitride layer is removed in phosphoric acid. This tip sharpening process has yielded tip radii under 500 Å (Figure 3b) for tip apex angles of approximately 30 degrees. We found that smaller tip apex angles sometimes cause tearing in the top silicon nitride mask due to stress between the silicon nitride and oxide layers.

A second possible method for forming lateral tips uses an anisotropic etch to define the tip. Silicon nitride is grown over the silicon cantilevers and patterned so that it protects all areas of the cantilever except for the end, leaving two exposed convex corners. An anisotropic etch such as KOH will undercut

these exposed convex corners, forming a tip at two intersecting <411> planes in the silicon [22]. Although this method may be very suitable for other applications requiring in-plane tips we did not investigate this method further because the tip apex angle between the intersecting <411> planes is limited to 62 degrees which, as mentioned earlier, is not optimal with respect to magnetic shape anisotropy once a NiFe film is evaporated.

Finally, a very successful technique we have explored involves both a double exposure and double etch to define a single in-plane tip [23] (Figure 5c). First a 1.0 μm layer of photoresist defines the cantilever with a 0.5 μm rounded tip. The cantilever is patterned as usual in an SF_6 etch. The first photoresist mask is stripped and a second photoresist mask containing the same pattern is applied again but shifted laterally by 1 μm. The second dry etch forms a tip at the intersection of the sidewall formed during the first etch. This process can be used to form tips with radii below 500 Å and with high aspect ratios. In addition, these tips can be fabricated using many different materials including silicon nitride. The vertical sharpness of the tip is limited only by the thickness of the material. We have found this process to require significantly shorter processing times and has improved tip-to-tip uniformity.

Evaporation was used to place ferromagnetic material on the end of the tip. However because these cantilevers are extremely thin, simple evaporation onto one side of the tip can cause dramatic bending of the tip even for film thickness below 350 Å over tip lengths of only a few μm. This is caused by the tensile stress of the film that develops during deposition. Therefore, in order to place magnetic material such as Ni, Co, or NiFe at the end of a sharp tip, we have developed an edge-on evaporation technique shown in Figure 2e. Edge-on evaporation maximizes the amount of material deposited onto the very end of the tip and reduces material deposited along the length of the cantilever. Furthermore, edge-on evaporation balances the amount of material deposited on both sides of the tip. During evaporation, the side that receives the greater flux of material will tend to bend away from the source, increasing the deposition rate onto the opposite side. This self-correcting process causes no noticeable bending provided the cantilever is aligned perpendicular to the evaporation source before the start of the evaporation. Figure 3 shows a finished cantilever after edge-on evaporation.

DOUBLE-TORSIONAL SILICON NITRIDE PROBE

To fabricate the double-torsional probe, a layer of low stress silicon nitride 900 Å thick was grown on top of a <100> oriented silicon wafer and patterned as shown in Figure 6. The exposed silicon surrounding the silicon nitride oscillators was etched in a solution of 20% TMAH until the entire silicon nitride structure was undercut leaving it supported by the 4 μm wide beam running vertically through Figure 6. The oscillators were stored in water and finally released using critical point drying. The geometry was chosen such that moments of inertial about the central supporting beam are balanced. The inner torsional element has a tip which extends past the outer oscillator in order to allow sample approach. The high Q torsional mode results from counter oscillation of the outside frame and the inside probe. The fiber interferometer was aligned to the middle paddle located to the left of torsional beam support in Figure 6.

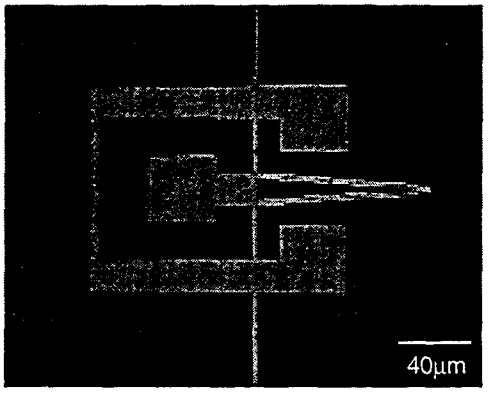


Figure 6: Optical micrograph of the double torsional force probe. The torsional beam supports the structure at both ends and is 4 μm wide and over 200 μm long.

Q MEASUREMENTS

A fiber optic interferometer was used in an evacuated bell jar to perform room temperature Q measurements. Q values were determined by measuring both decay times and amplitude vs. frequency curves. Figure 7 shows typical Q vs. pressure results for a bare silicon cantilever (no magnetic coating) that is 0.3 μm thick with a resonance frequency of 30 kHz. The lowest pressure shown in Figure 7, $2x10^{-3}$ Torr, was the lowest obtainable inside our bell jar and clearly shows that the highest measured Q value of 18,000 was still pressure limited.

Lower pressures (below 10^{-5} Torr) were obtained inside a low temperature vacuum apparatus. Q vs. temperature measurements are shown in Figure 8 for a bare silicon cantilever with the same dimensions as shown in Figure 3 and a resonance frequency of 14 kHz. The Q increased monotonically as temperature is decreased due to freeze-out of energy dissipation mechanisms. At 5 K and 10^{-6} Torr, Q values as high as 26,000 were measured. After coating this cantilever with 350 Å of NiFe, a Q of 22,000 was obtained at low temperature.

These Q values represent the largest we have measured for a cantilever below 0.2 μm in thickness, but they are still lower than the Q of 200,000 we have measured for 1.5 μm thick silicon cantilevers [13]. We attribute the lower Q in the thinner cantilevers to the greater importance of surface losses which may arise from surface dislocations and other microscopic imperfections. One possibility for increasing the Q is to do a forming gas anneal in order to reduce the surface dislocations and dangling bonds that may limit the cantilever Q.

In order to calculate the noise sensitivity of these soft cantilevers it is necessary to determine their spring constants. Attempting to measure load vs. displacement curves is impractical given that forces below 10^{-11} N would need to be applied directly at the tip. A more precise method for determining the spring constant is to directly measure the RMS thermal vibration noise. At 7K the RMS thermal noise amplitude was 2 Å, giving a spring constant of about $2x10^{-3}$ N/m. Neglecting variations in the beam shape from the paddle and tip, this value is within 10% of that predicted by simple beam theory.

Using the measured values for the spring constant, frequency, and Q, the calculated force resolution of the cantilever (shown in Figure 3) was $2x10^{-17}$ N$\sqrt{\text{Hz}}$ at 10 K. This

force resolution is the best we have measured and should be sufficient for single electron spin detection.

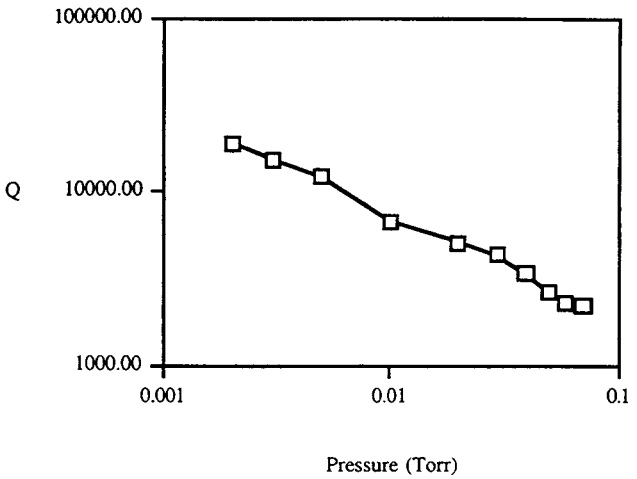


Figure 7: Q vs. Pressure for a 0.3 μm thick silicon cantilever

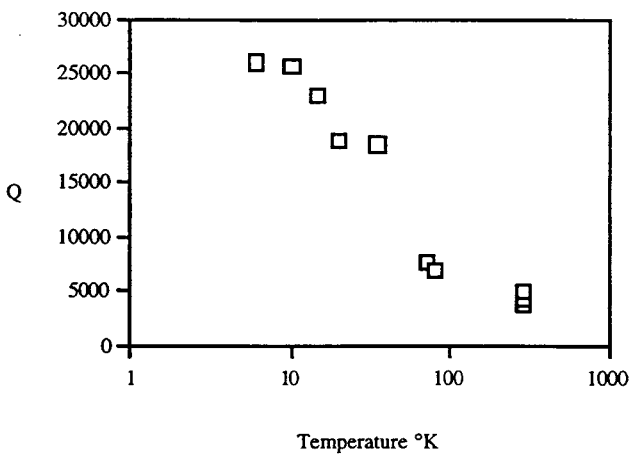


Figure 8: Q vs. Temperature for a 0.17 μm thick cantilever.

Lastly, the Q performance of the double torsional oscillator shown in Figure 6 was investigated. At 10 K there were four high Q modes ranging in frequency from 4.8 to 23 kHz. The highest Q modes at 4.8 kHz and 17.5 kHz had Q values of 20,000. This oscillator has the highest Q we have measured for silicon nitride oscillators of comparable thickness. this result is 25% greater than that obtained for a simple silicon nitride cantilever beam (Table I). The mode of oscillation (flexural vs. torsional) and cantilever geometry continue to play a role in determining the Q for these extremely thin oscillators However, we note that the Q improvement we measured is not as pronounced as improvements reported by Buser [17] for much thicker silicon oscillators.

SURFACE MEASUREMENTS

These cantilevers must maintain their high Q as the tip approaches a sample surface in order to maintain high force sensitivity. Sample approach measurements taken at room temperature inside a bell jar indicate that these cantilevers maintain relatively high Q until surface interactions begin to introduce damping 100-1000 Å from the sample. Approach measurements also show an increase in the natural resonance frequency of the cantilever due to van der Waals forces. This behavior is similar to that seen by Gregor, *et al.* [24] in which a tapered fiber optic tip was driven at resonance over a sample surface.

In order to verify that the end of the tip was magnetic after the edge-on evaporation, these cantilevers were used to image the bit patterns on a disk drive sample. The resonance frequency of the cantilevers was sensitive to magnetic field gradients in the lateral direction shown in Figure 9. Figure 10 shows a magnetic force microscopy (MFM) image of disk drive tracks taken approximately 5000 Å above the surface by measuring the phase of the cantilever oscillation while it was driven at a constant frequency of 12 kHz to a vibration amplitude of 100 Å.

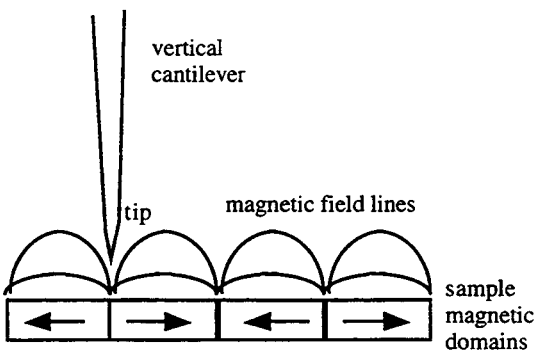


Figure 9: Illustration of the magnetic force microscopy technique.

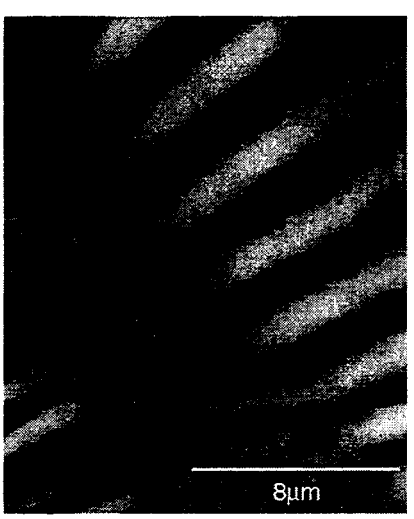


Figure 10: Imaged 8 μm wide disk drive track.

CONCLUSIONS

We have investigated both flexural and torsional mode oscillators for magnetic resonance force microscopy. In particular, we demonstrated ultrathin single crystal silicon cantilevers with sharp in-plane magnetic tips that are operated perpendicular to the substrate. Three techniques for making in-plane tips have been explored yielding tip radii below 500 Å and narrow tip apex angles. The predicted field gradients from these in-plane magnetic tips should produce interaction forces as high as 10^{-16} N for a single electron spin. In a one hertz detection bandwidth, the force resolution of these cantilevers approaches 10^{-17} N. Thus, single electron magnetic resonance force microscopy should be feasible.

ACKNOWLEDGMENTS

We acknowledge S. Hoen's work in the design and testing of our critical point dryer and are grateful for invaluable processing advice from Chris Storment, Ron Maynard, John Suh, and Steve Minne at the Center for Integrated Systems. T. Stowe received funding for this research from the John and Fannie Hertz Foundation. D. Botkin received partial support from the Office of Naval Research. This work has also been supported by the National Science Foundation CAREER Award (ECS-9502046), The Charles Lee Powell Foundation, and the Terman Fellowship. We would also like to acknowledge the National Science Foundation for travel support for this conference.

REFERENCES

1. G. Binning, C. F. Quate, and C. Gerber, *Physics Review Letters*, *56*, 930 (1986).

2. D. Rugar, P. Hansma, "Atomic Force Microscopy", *Physics Today, 43* (1990).

3. J. A. Sidles,*Phys. Rev. Lett., 58*, 2854 (1991).

4. O. Züger and D. Rugar, "First images from a Magnetic Resonance Force Microscope", *Applied Physics Letters*, *63*, 2496-2498 (1993).

5. J. A. Sidles, *et al.*, "Magnetic Resonance Force Microscopy", *Review of Modern Physics*, *67*, No 1, pp. 249-265 (1995).

6. D. Rugar, C. S Yannoni, J. A. Sidles, "Mechanical Detection of Mangetic Resonance", *Nature, 360*, pp.563-566 (1992).

7. D. Rugar, *et al.*, "Force Detection of Nuclear Magnetic Resonance", *Science*, *264*, pp.1560-1563 (1994).

8. P.C. Hammel, *et al.*, "Sub-Surface Imaging with the Magnetic Resonance Force Microscope", *Journal of Low Temperature Physics*, *101*, No. 1/2, pp.59-69 (1995).

9. D. Rugar, H. J. Mamin, P. Guethner, *Applied Physics Letters*, *55*, pg. 2588 (1989).

10. A. DiCarlo, *et al.*, "Magnetic force microscope utilizing an ultra-small-spring-constant vertically cantilevered tip", *Ultramicroscopy*, *47*, pp. 383-392 (1992).

11. T.B. Gabrielson, "Mechanical-thermal noise in micromachined acoustic and vibration sensors.", *IEEE Trans. on Electron Devices*, *40*, No.5, pp. 903-909 (1993).

12. S. Hoen, *et al.*, "Fabrication of Ultrasensitive Force Detectors", *Proceedings of the Solid State Sensor and Actuator Workshop*, pp. 209-212 (1994).

13. K. Wago, *et al.*, "Low Temperature Magnetic Resonance Force Detection", *J. Vac. Sci. Tech B*, March/April, (in press) (1996).

14. H. Guckel, *et al.*, "Polysilicon Resonant Microbeam Technology for High Performance Sensor Applications", *Proceedings of the Solid State Sensor and Actuator Workshop*, pp 153-156 (1992).

15. J. D. Zook, *et al.*, "Characteristics of polysilicon resonant microbeams", *Sensors and Actuators A*, *35*, pp 51-59 (1992).

16. R. N. Kleiman, *et al.*, "Single-crystal silicon high-Q torsional oscillators", *Review of Scientific Instruments*, *56*, pp. 2088-91 (1985).

17. R.A. Buser, *et al.*, "Resonant Silicon Structures", *Sensors and Actuators*, *17*, pp. 145-153 (1989).

18. R. A. Buser and N. F. de Rooij, "Very high Q-factor resonators in monocrystalline silicon torsional resonators", *Sensors and Actuators*, *A21-A23*, pp. 323-327 (1990).

19. G. T. Mulhern, D. S. Soane, and R.T. Howe, *1993 International Conference on Solid-State Sensors and Actuators: Transducers '93*, Yokohama, Japan (1993).

20. J. J. Yao, N. C. MacDonald, "Fabrication of high frequency two-dimensional nanoactuators for scanning probe devices", *J. Micromechanical Systems*, *1*, pp. 14-22 (1992).

21. J. J. Yao and N. C. MacDonald, "A micromachined, single-crystal silicon, tunable resonator", *J. Micromechanical Microengineering*, *6*, pp. 257-264 (1996).

22. H. Sandmaier, *et al.*, "Corner Compensation Techniques in Anisotropic Etching of (100)-Silicon using Aqueous KOH", *International Conference on Solid-State Sensors and Actuators* (1991)

23. Kobayashi, D., *et al.*, "Control of a micro tunneling unit under atomic force", *IEEE/RSJ International Conference on Intelligent Robots and Systems. Human Robot Interaction and Cooperative*, 2, (1995).

24. M. J. Gregor, *et al.*, "Probe-surface interaction in near-field optical microscopy: The nonlinear bending force mechanism", *Applied Physics Letters*, *68*., No. 3, pp. 307-309 (1996).

A Fully Integrated Magnetically Actuated Micromachined Relay

William P. Taylor and Mark G. Allen
Georgia Institute of Technology, School of Electrical and Computer Engineering
Atlanta, GA 30332-0250

Charles R. Dauwalter
Milli Sensor Systems and Actuators
93 Border Street, W. Newton, MA 02165-2013

ABSTRACT

A fully integrated magnetically actuated micromachined relay (microrelay) has been fabricated and tested. Fabricated devices range in size from 6 mm x 7 mm to 7 mm x 7.4 mm. A coil current of 600 mA actuates the microrelay, with a coil power dissipation of 320 mW. The ON resistance of the microrelays is typically less than 30 Ω, with values as low as 0.85 Ω observed under electromagnetic operation. A contact resistance of 0.017 Ω was observed under mechanical actuation.

INTRODUCTION

Due to parasitic effects and offset voltage injection in solid state relays, electromechanical relays remain widely used in a variety of applications such as data acquisition boards and aircraft and automotive control circuitry. Micromachined versions of these relays (microrelays) are desirable due to their potential for smaller size and batch fabrication.

Micromachined silicon switches were demonstrated in 1979 by Petersen [1]. These devices were bulk micromachined, electrostatically actuated, and contained metal contacts in order to reduce contact resistance and increase current carrying capacity. More recently, electrostatic polysilicon microrelays have been achieved [2,3]. These devices have a contact resistance of approximately 10 kΩ, and a carry current of less than 10 mA. More recent versions of electrostatically-driven microrelays with metallic contacts and lower contact resistance have also been reported [4-7]. These devices have the advantage of higher carry currents, and cycle lifetimes on the order of 10^8 [5]. Another recently-reported microrelay device employs the use of a mercury droplet to eliminate some of the contact resistance issues, such as contact force, wear, and oxidation [8]. This device was reported to have a contact resistance of less than 1 Ω, with a carry current of 20 mA.

Lower actuation voltages may be useful in many applications, such as automotive applications in which low-voltage, high current electrical systems are present. Magnetically-driven microrelays, which have lower actuation voltages, may satisfy these constraints. Previously reported magnetic microrelays [9-11] do not have fully integrated magnetic coils. Instead, these devices either require an external coil for electromagnetic actuation or integrated heating coils which act to demagnetize portions of an externally-driven magnetic circuit and results in a switching action. The use of a fully integrated magnetically actuated microrelay design combines the manufacturing, cost, and size benefits of full integration with the benefits of magnetic actuation. In addition, these devices offer the potential for plug-in replacement of conventional reed or other relays in current designs. Our approach is to use micromachining techniques to realize a magnetically-driven microrelay.

DEVICE CONCEPT

Figure 1 shows a schematic top view of the microrelay. The basic structure of the microrelay consists of a zigzag or meander coil which winds over an insulated lower magnetic core, made of a soft magnetic material, and through soft magnetic material side cores. The coil is electrically insulated from the magnetic materials. The entire coil and core structure, which is also electrically insulated on the top, can be thought of as an integrated driving electromagnet. There are two contacts on top of the electromagnet. In order to achieve a relay action, the two contacts are alternately electrically connected and disconnected by a movable conducting platform, which is integrated with the structure using surface micromachining.

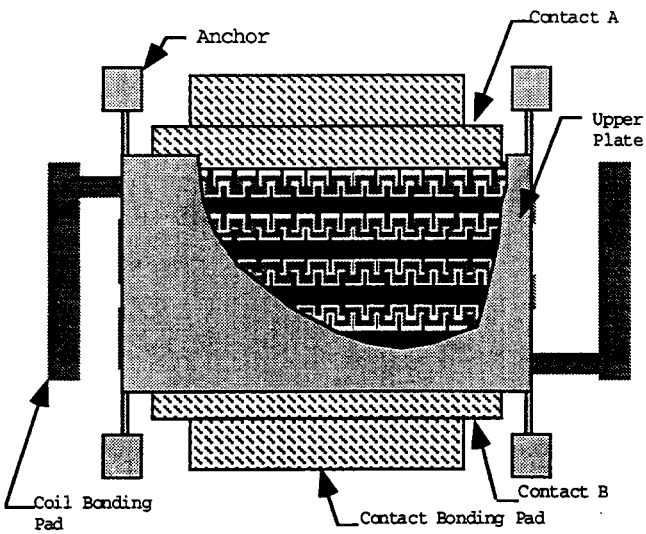


Figure 1: Schematic drawing of microrelay (top view)

The basic operation of the device is as follows. A current flows through the coil and generates a magnetic flux which creates a force on the upper magnetic movable plate. This force causes the upper plate to move downwards and connect the two contacts which results in a switching action. When the coil current is turned off, the mechanical stress in the arms generated by the downward movement of the plate forces the upper magnetic plate to rise from the contacts, thus opening the relay. Figure 1 shows a four-beam upper plate support configuration; however, other devices with varying beam numbers and geometries (including cantilever-type devices) have also been realized.

0-9640024-1-8 $3.00 © 1996 TRF

THEORY

A cross section of the generalized actuator is shown in Figure 2. Each of the slots of the structure contains N electrical conductors, each of which carries current I. Assuming linearity, an expression for the normal (y-direction) force between the stator and armature may be obtained as follows. For the actuator being described, the width of a tooth, a, and the actuator length, l, are both relatively large compared to the gap, g, so that the effects of fringing are small and are neglected in the analysis. The permeability of the stator and armature iron is assumed to be large enough that the reluctance of the iron can be neglected.

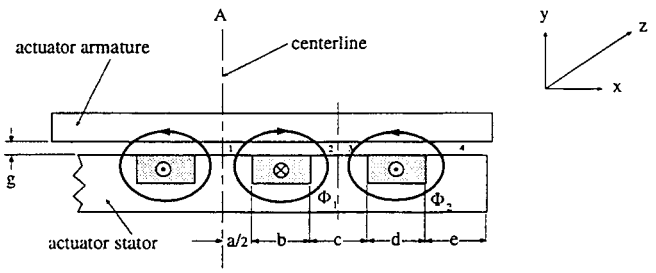


Figure 2: Side view of theoretical actuator with definitions of dimension variables included.

The magnetic reluctances of the air gap regions numbered 1, 2, 3, 4 in the figure are:

$$\Re_2 = \Re_3 = \frac{g}{\mu_0 (c/2) l} = \frac{2g}{\mu_0 c l} \tag{1}$$

$$\Re_1 = \frac{g}{\mu_0 (a/2) l} = \frac{2g}{\mu_0 a l} \text{ and } \Re_4 = \frac{g}{\mu_0 e l} \tag{2a,b}$$

The combined permeance of air gap regions 2 and 3 is:

$$P_{2,3} = \frac{1}{\Re_2} + \frac{1}{\Re_3} = \frac{\mu_0 c l}{2g} + \frac{\mu_0 c l}{2g} = \frac{\mu_0 c l}{g} \tag{3}$$

Similarly:

$$P_{1,4} = \frac{1}{\Re_1} + \frac{1}{\Re_4} = \frac{\mu_0 a l}{2g} + \frac{\mu_0 e l}{g} = \frac{\mu_0 l}{2g}(a+2e) \tag{4}$$

The reluctances of regions 2,3 and 1,4 are then:

$$\Re_{2,3} = \frac{1}{P_{2,3}} = \frac{g}{\mu_0 c l} \text{ and } \Re_{1,4} = \frac{1}{P_{1,4}} = \frac{2g}{\mu_0 l (a+2e)} \tag{5a,b}$$

The total air gap reluctance is:

$$\Re_{total} = \Re_{1,4} + \Re_{2,3} = \frac{2g}{\mu_0 l}\left\{\frac{1}{a+2e} + \frac{1}{2c}\right\} = \frac{2g}{\mu_0 l}\left\{\frac{2c+a+2e}{2c(a+2e)}\right\} \tag{6}$$

so that the total air gap permeance becomes:

$$P_{total} = \frac{1}{\Re_{total}} = \frac{\mu_0 l}{2g}\left\{\frac{2c(a+2e)}{2c+(a+2e)}\right\} = \frac{\mu_0 c l}{g}\frac{(a+2e)}{(a+2e)+2c} \tag{7}$$

The total magnetic flux is the magnetomotive force, $F = NI$, multiplied by the total air gap permeance, viz:

$$\Phi_{total} = NIP_{total} \tag{8}$$

The total flux is then:

$$\Phi_{total} = NIP_{total} = \frac{\mu_0 NIcl}{g}\frac{(a+2e)}{(a+2e)+2c} \tag{9}$$

The flux, λ, linking the conductors is:

$$\lambda = N\Phi_{total} = \frac{\mu_0 N^2 Icl}{g}\frac{(a+2e)}{(a+2e)+2c} \tag{10}$$

The inductance is defined as the flux linkages per unit current:

$$L = \frac{\lambda}{I} = \frac{\mu_0 N^2 cl}{g}\frac{(a+2e)}{(a+2e)+2c} \tag{11}$$

The magnetic energy, W, is:

$$W = \frac{1}{2}LI^2 = \frac{\mu_0 (NI)^2 cl}{2g}\frac{(a+2e)}{(a+2e)+2c} \tag{12}$$

The normal (in the y direction) force on this section of the armature is therefore:

$$F_y = \frac{dW}{dy} = \frac{-\mu_0 (NI)^2 cl}{2g^2}\frac{(a+2e)}{(a+2e)+2c} \tag{13}$$

The equations derived above were used to estimate the force generated on the upper plate. The relay air gap was sized to be small compared with other characteristic lengths, in order to minimize the effects of leakage flux. Typical relay geometries considered in this paper had initial forces on the upper plate in the 1-10 mN range as calculated using the above equations. This force must be used both to overcome support beam flexure resistance as well as provide contact closure. From [9], it was estimated that a contact closure force of at least 200 μN is required for reliable relay operation; the compliance of the upper plate support beams was calculated using this criterion. The coils were designed to carry a minimum of twice the estimated actuation current. Contacts were designed to obtain low contact resistance through the use of large contact areas. For this relay prototype, no effort was made to optimize the contacts for maximum lifetime or minimum resistance.

FABRICATION

Fabrication is based on standard polyimide mold electroplating techniques [12]. Figure 3 depicts the major fabrication steps. The fabrication begins by deposition of a seed layer onto the substrate followed by electroplating of a

permalloy lower magnetic core through a patterned polyimide mold. The lower core is then insulated with polyimide, and a copper coil is electrodeposited and insulated. The polyimide between the coil lines is removed with a standard plasma etching process, and the magnetic side cores are deposited in these cavities using the bottom magnetic core as the electroplating seed layer, forming the electromagnet. The entire electromagnet is then further insulated and planarized. Figure 4 shows a close-up of the completed electromagnet prior to contact and movable plate deposition. Contacts are deposited above the electromagnet, and a photoresist sacrificial layer is patterned over the contacts and electromagnet. An electroplating seed layer is deposited over the photoresist, and the upper magnetic plate is electrodeposited through a second photoresist mold. All photoresist and seed layers are then removed to release the upper magnetic plate. A photomicrograph of the completed and released microrelay is shown in Figure 5.

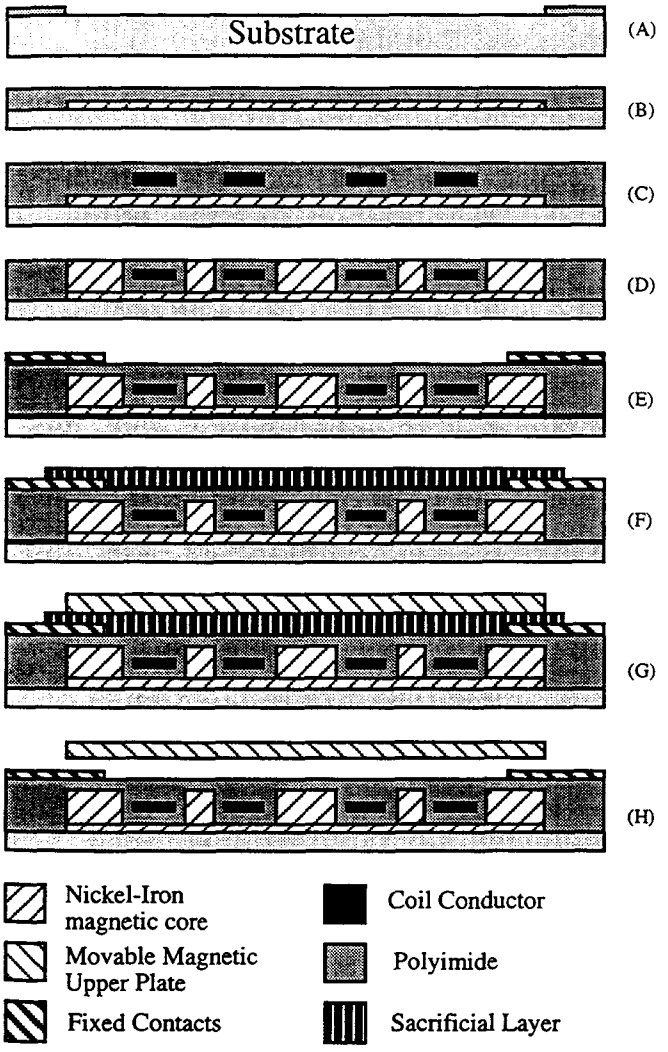


Figure 3: Fabrication sequence for the microrelay. (a) after deposition of plating mold for lower magnetic core; (b) after deposition of lower magnetic core; (c) after deposition of coil; (d) after formation of side cores; (e) after contact deposition; (f) after sacrificial layer deposition; (g) after deposition of upper magnetic plate; (h) completed device.

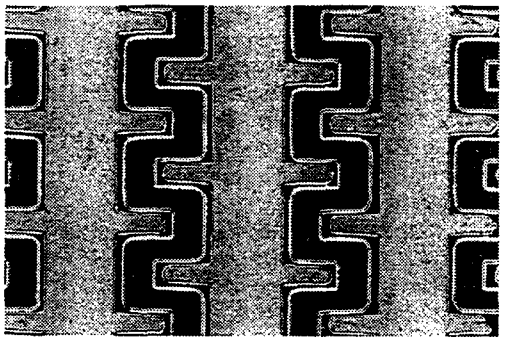


Figure 4: Photomicrograph of a fabricated electromagnet, top view. The darker serpentine regions are the coils; the light regions are the permalloy side cores.

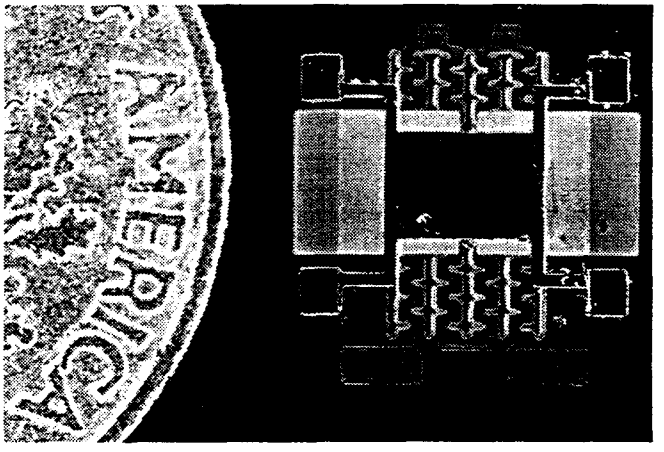


Figure 5: Photomicrograph of a completed magnetic microrelay, top view. The relay is shown next to a dime for comparison. The upper magnetic plate is suspended by four arms over the contacts. The upper magnetic plate is 2.2 mm by 3.4 mm. The beams are 200 μm in width, and 2 mm in length (1 mm from plate to the turn in the beam, and 1 mm from the turn in the beam to the anchor point).

RESULTS

The average coil resistance was less than 1 Ω, depending on coil size and thickness. The microrelay coils have been tested up to 1.8A DC for several minutes and showed no signs of failure. Two types of plate suspension were tested: a four-arm bridge suspension as shown in Figure 5, and a two-arm cantilever type suspension (i.e., the relay as shown in Figure 5 but with the upper plate supported by two arms, both of which were on the same side of the suspended plate). The beam geometries for each type of suspension were varied. In addition, two operating modes of the relay were investigated: one in which the movable plate bridged both lower contacts, and one in which the movable plate formed one of the contacts and the lower contacts formed the second contact. The first operation mode has higher contact resistance but no current must flow through the movable plate support arms; the second operation mode has lower contact resistance since only one interface is formed upon relay closure. All types of suspensions and operating modes produced successfully actuating and functional relays. As expected, the more flexible cantilever-type suspension required lower actuation currents, and the second operation mode had lower contact resistance. For a typical cantilever-type device, with beam dimensions of 200 μm width and 2 mm length, actuation was achieved with a 500 mA coil

current, although the lowest contact resistances were observed at 600 mA. For a bridge-type device, i.e., with four suspension arms, actuation required 1200 mA. For a typical cantilever type device with a 0.87 Ω coil, the actuation power (i.e., power consumed by the coil) was 320 mW at the 600 mA actuation current.

Measurements of the microrelay contact resistance under magnetic actuation were performed in air using a DC multimeter. "ON" resistances as low as 0.85 Ω, with typical values of 5-10 Ω, were seen repeatedly. These results were observed on several different relay designs, including cantilever and bridge suspensions. The "OFF" resistance was infinite; no current flow was observed with voltages in excess of 10 V applied across the contacts.

A long-term reliability test of the relay was performed using a computer-controlled actuation and measurement setup. Figure 6 shows reliability test results for a typical cantilever-type device. In Figure 6, the "ON" resistance of the relay as a function of number of trials is shown. For the first 9,000 trials, the device is operating in the first mode (plate bridging two contacts); for the second 8,000 trials, the device is operating in the second mode (plate acting as one contact); the shift in contact resistance between mode operation shows the importance of the contact surfaces on the contact resistance of the microrelay. A vertical line extending to the top of the graph indicates an open circuit when current is applied to the driving coils, i.e., a failure of the relay to close. After an initial burn-in period, the device showed more than 14,000 operations with only one missed switching operation (defined as a contact resistance in the "ON" state of more than 300 Ω). Above 15,000 operations, several more failures were observed. The test was arbitrarily terminated at 17,000 operations, at which time the relay was still functional.

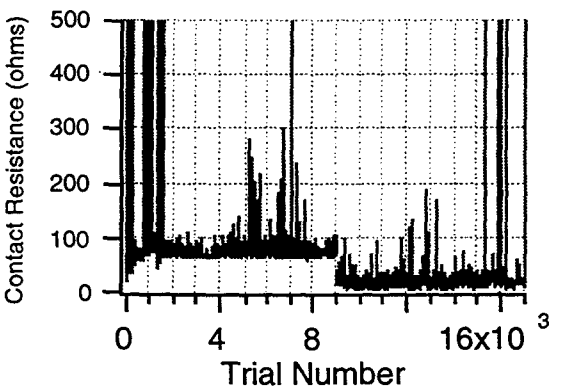


Figure 6: Lifetime results for a cantilever beam upper magnetic plate relay. The actuation current was 600 mA.

Testing the relay with an active load which could be varied between 10 and 60 mA has also been performed. A cantilever type relay operating at 600 mA coil current was able to repeatedly switch a 10 mA DC load. The maximum load switched by any relay design was 60 mA at a coil current of 800 mA. The actuation time was measured to be less than 10 msec. Release times were measured to be 25 msec. Localized contact damage was observed upon application of a 30V DC signal. The contacts when closed mechanically (non-switching) had a contact resistance of 0.017 Ω and could carry 4.5 A of current for 2 minutes before failure occurred, and 3 A repeatedly without failure. Although these large carry currents have not yet been switched, it should be noted that these carry currents are well above the maximum 1 A of conventional reed relays, and comparable to small size armature relays.

CONCLUSIONS

A fully integrated magnetically actuated micromachined relay has been fabricated and tested. Actuation of the microrelay occurs at 600 mA coil current, 320 mW coil power for a cantilever type upper plate. These coil power levels are comparable with armature relays, and within a factor of ten when compared to typical reed relays. The contact resistance of the microrelay is less than 30 Ω. Preliminary lifetime tests show that the microrelay can operate over 14,000 cycles with only one failed switching action. The microrelay has shown the ability to repeatedly switch up to 60 mA electromagnetically, and in excess of 3A with mechanical actuation.

In these prototype microrelays no attempt has been made to address contact issues. In addition, the coil has not been optimized for minimum resistance. The resultant improvement in efficiency when the coil and contacts are optimized may lead to fully integrated magnetically actuated micromachined relays with characteristics that compare favorably to conventional reed and armature devices.

REFERENCES

1. K. E. Petersen, "Micromechanical Membrane Switches on Silicon," *IBM J. Res. Develop., vol. 23, no. 4,* 376-385 (1979).

2. M. A. Gretillat, P. Thieubaud, N. F. de Rooij, and C. Linder, "Electrostatic Polysilicon Microrelays Integrated with MOSFETs," *Proc. IEEE Microelectromechanical Sys. Conference*, Oiso, Japan, pp. 97-101 (1994).

3. M. A. Gretillat, P. Thieubaud, C. Linder, and N. F. de Rooij, "Integrated circuit compatible electrostatic polysilicon microrelays," *J. Micromech. Microeng., 5,* 156-160 (1995).

4. S. Roy and M. Mehregany, "Fabrication of electrostatic nickel microrelay by surface micromachining," *Proc. IEEE Microelectromechanical Sys. Conference*, Amsterdam, Netherlands, pp. 353-357 (1995).

5. J. Drake, H. Jerman, B. Lutze, and M. Stuber, "An Electrostatically Actuated Microrelay," *Proc. Transducers '95*, Stockholm, Sweden, vol. 2, pp. 380-383 (1995).

6. M. Sakata, "An Electrostatic Microactuator for Electro-Mechanical Relay," *Proc. IEEE MEMS Workshop*, Salt Lake City, 149-151 (1989).

7. J. J. Yao and M. F. Chang, "A Surface Micromachined Miniature Switch for Telecommunications Applications with Signal Frequencies up to 4GHz," *Proc. Transducers '95*, Stockholm, Sweden, vol. 2, pp. 384-387 (1995).

8. J. Simon, S. Saffer, and C.J. Kim, "A Micromechanical Relay with a Thermally-Driven Mercury Micro-Drop," *Proc. IEEE Microelectromechanical Sys. Conference*, San Diego, pp. 515-520 (1996).

9. H. Hosaka, H. Kuwano, and K. Yanagisawa, "Electromagnetic Microrelays: Concepts and Fundamental Characteristics," *Proc. IEEE Microelectromechanical Sys. Conference*, Ft. Lauderdale, pp. 12-17 (1993).

10. E. Hashimoto, H. Tanaka, Y. Suzuki, Y. Uenishi, and A. Watabe, "Thermally Controlled Magnetization Actuator (TCMA) using Thermosensitive Magnetic Materials," *Proc. IEEE Microelectromechanical Sys. Conference*, Oiso, Japan, pp. 108-113 (1994).

11. H. Hosaka, H. Kuwano, and K. Yanagisawa, "Electromagnetic microrelays: consepts and fundamental characteristics," *Sensors and Actuators A, vol. A40,* 41-47, (1994).

12. A. B. Frazier, C. H. Ahn, and M. G. Allen, "Development of Micromachined Devices Using Polyimide-Based Processes, *Sensors and Actuators A (Physical), vol.A45, no.1,* 47-55, (1994).

A Thin Film Teflon Electret Technology for Microphone Applications

Tseng-Yang Hsu, Wen H. Hsieh and Yu-Chong Tai
Department of Electrical Engineering
California Institute of Technology
Pasadena, CA 91125, USA

Katsushi Furutani
Department of Mechanical System Engineering
Toyota Technological Institute
12-1, Hisakata 2-chome, Tempaku-ku
Nagoya 468, Japan

ABSTRACT

A micromachining compatible Teflon electret technology has been developed. The electret material used is thin film spin-on Teflon. A pulsed electron gun, called the Back-Lighted Thyratron (BLT), is used for charge implantation. A charge compensation method is used to measure the charge density of the electret. The electret has a saturated charge density on the order of 10^{-4} C/m^2. A prototype silicon micromachined electret microphone has been fabricated and tested. An open circuit sensitivity of 0.3 mV/Pa has been achieved for a hybrid microphone package.

INTRODUCTION

An electret is a dielectric that produces a permanent external electric field which results from permanent ordering of molecular dipoles or from stable uncompensated surface or space charge. Electrets have been the subject of study for their charge storage characteristics as well as for their application in a wide variety of devices such as acoustic transducers, electrographic devices, and photocopy machines. To incorporate electret technology into micromachining technology is an attractive idea for the fabrication of small, self-powered devices. We have developed and demonstrated a technology to make thin film electrets using micromachining technology. The use of such a technology has been demonstrated in a micromachined microphone which can produce a signal without the need for external biasing. This reduces the system volume and complexity.

ELECTRET FABRICATION

The electrets fabricated in our laboratory are made from Du Pont spin-on Teflon® AF 1601S. The Teflon dielectric film is prepared by spinning at 2 krpm and baking at 250 °C. The resulting dielectric film is approximately 1 μm thick with surface roughness of less than 2000 Å across the wafer. For time spans longer than usual processing times, the adhesion of the Teflon film to different material surfaces such as silicon, silicon dioxide, silicon nitride, copper, gold, and chrome is satisfactory in the presence of chemicals frequently used in MEMS fabrication, such as water, photoresist developers, acetone, alcohol, HF and BHF. If necessary, the Teflon film can be patterned using oxygen plasma.

To form electrets, we have investigated the new use of the Back-Lighted Thyratron for electron implantation into the thin film Teflon. The BLT [1-2] structure we have designed (Figure 1) consists of two electrode plates with hollow back cathode and hollow back anode. The two electrodes facing each other have a diameter of 75 mm and a center aperture of 5 mm. The electrodes are separated by a 5 mm thick insulating plexi-glass plate. The structure is filled with a low pressure gas such as hydrogen or one of the noble gases to a pressure of 50 to 300 mTorr. The device is triggered optically by an ultraviolet light pulse applied to the back of the cathode. The ultraviolet light generated by a flashlamp passes through a quartz window into the back of the cathode. This initiates the formation of an electron beam which is directed towards the thin film Teflon sample.

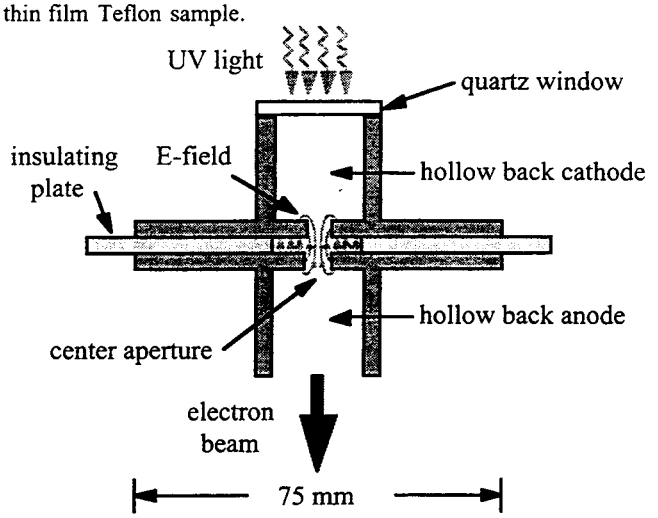


Figure 1. The Back-Lighted Thyratron.

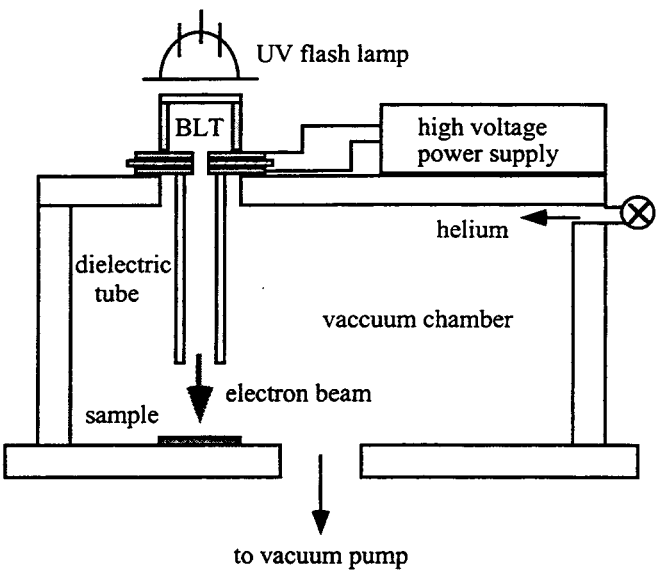
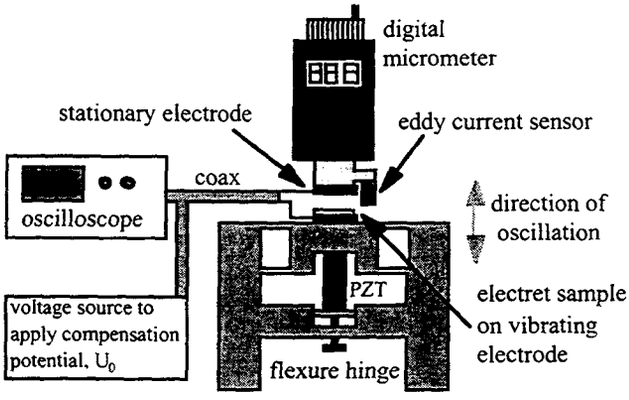
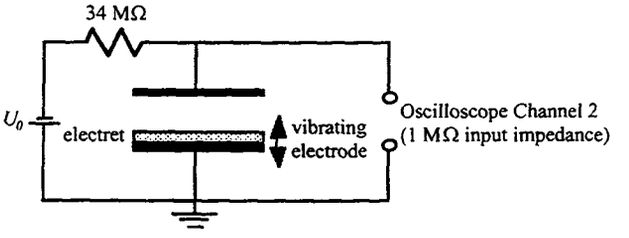


Figure 2. The BLT charge implantation system.

Our in-house BLT is constructed on top of a vacuum chamber with a triggering UV flashlamp at a distance of 2 cm away from the quartz window as shown in Figure 2. The BLT cathode is biased at a high negative potential for beam acceleration. The electron beam pulse is directed to the sample which is 12 cm away from the beam exit. With a beam divergent angle of 6°, the beam diameter is 1.75 cm at the sample surface [2]. Integrating a dielectric tube at the beam exit has the effect of collimating and focusing the electron beam [3]. The bias potential is adjusted according to the desirable range of electrons in the Teflon film. In the case of 1 μm thick Teflon film, the electron beam energy is set at 10 keV which gives an implantation depth of approximately 1 μm [4].



(a)

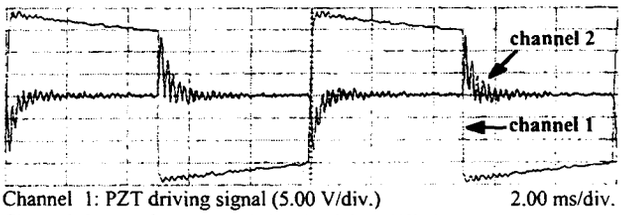


(b)

Figure 3. *(a) The setup used to measure the charge density of electret sample. (b) Charge compensation circuit.*

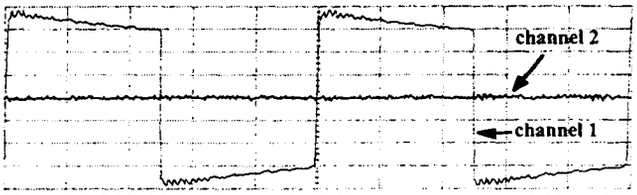
ELECTRET CHARACTERISTICS

To measure the charge density, a setup consisting of a PZT stack and a micrometer controlled stationary electrode was constructed (Figure 3a). To confine displacement in the z-direction only, the PZT is integrated into a flexure hinge made of 304 stainless steel and machined by electrical discharge machining (EDM). The movable part of the flexure hinge weighs 30 g and has a spring constant of 1.53 x 10^6 N/m. The PZT driver deforms 15 μm at 100 V and can be driven by a maximum voltage of 150 V. The linearity of the displacement of the PZT caused by hysteresis is 10%. The PZT is driven by a unit consisting of a periodic source and an amplifier. The amplifier is a class-B push-pull type amplifier specially designed for capacitive loads. An eddy-current sensor is integrated into the micrometer for monitoring and double checking the dynamic and static displacement. The test sample is prepared using a 1.2 cm × 1.2 cm silicon die evaporated with 2000 Å of Cr/Au. A 1 μm thick layer of Teflon® AF 1601S is coated on the Au surface and is then implanted with 10 keV electrons using the BLT at 420 mTorr of helium.



Channel 1: PZT driving signal (5.00 V/div.) 2.00 ms/div.
Channel 2: sample generated signal (50.0 mV/div.)

(a)



Channel 1: PZT driving signal (5.00 V/div.) 2.00 ms/div.
Channel 2: sample signal compensated by 3 V (5.00 mV/div.)

(b)

Figure 4. *(a) A PZT driving signal and the induced sample signal. (b) The same PZT signal and the compensated sample signal.*

The electret sample is fixed on top of the vibrating flexure hinge. The signal generated by induced charges on the stationary electrode due to the vibrating electret is then displayed on the oscilloscope (Figure 3b). By applying a compensation potential, U_0, between the two electrodes the net electric field in the air gap between the vibrating and stationary electrode can be reduced to zero. The signal generated by the induced charges thus becomes zero. The effective surface charge density, ρ_{eff}, of the electret sample is then given by [5]:

$$\rho_{eff} = \varepsilon_0 \, \varepsilon \, U_0 / t$$

where ε_0 is the permitivity of air, ε = 1.9 is the relative permitivity of the Teflon film and t is the electret thickness. Figure 4a shows a typical PZT driving signal at 100 Hz and the signal generated by the electret sample. Figure 4b shows the generated ac signal being compensated to zero by the application of a compensation potential, U_0. Depending on the number of electron pulses, the charge density of an electret sample ranges from 2×10^{-5} C/m^2 to 2×10^{-4} C/m^2. The maximum charge density we have obtained is comparable to what has been reported for Teflon films [4-6].

It is found from experiment that at room temperature the electret initially undergoes a 10-20% drop in total charge density a few hours after the implantation, but then stabilizes afterward as shown in Figure 5. Some samples were monitored at room temperature over a period of six months (the longest observation time so far up to the point of this paper) and no detectable charge decay was observed. Samples have also been tested for charge decay at elevated temperatures. Curve (a) of Figure 6 shows the charge density of a sample at 100 °C for 16 hours. The initial significantly large 40% drop in the charge density is due to the elevated temperature. However, even at 100 °C the charge stabilizes after the initial drop to a rate which is not measurable within the time span of the experiment. The same electret sample is then monitored for charge decay at 120 °C as shown by curve (b) of Figure 6. Again there is an initial drop in charge density, but the charge stabilizes after a few hours. The same trend is observed for the same sample at 140 °C, as illustrated by curve (c) in Figure 6. Although only short term data has been available so far, the

charge decay data obtained at room and elevated temperatures suggest that a stable electret can be formed using Teflon® AF and the BLT.

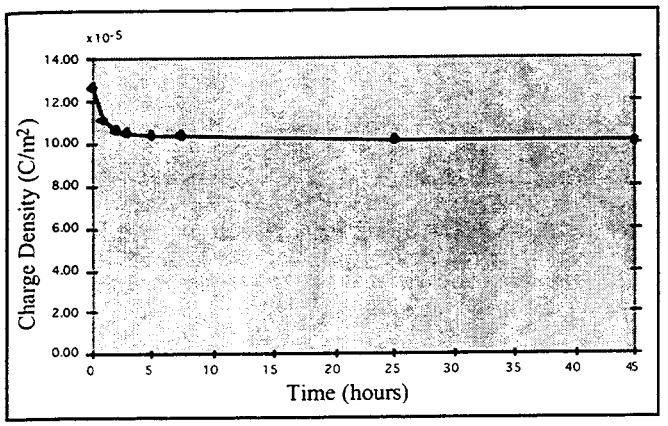


Figure 5. Charge density in Teflon® AF electret at room temperature.

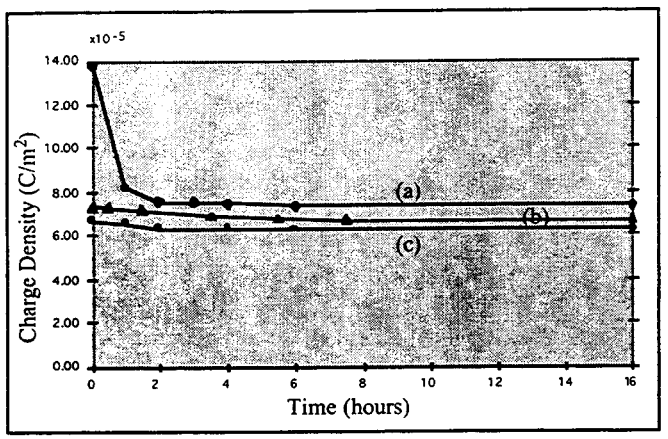


Figure 6. Charge density in Teflon® AF electret at elevated temperatures.

ELECTRET MICROPHONE

To demonstrate the self-powering capability of a MEMS compatible electret device, a micromachined electret microphone has been fabricated and tested. The fabrication of the membrane part of the microphone begins with a wafer coated with 1 μm thick low stress LPCVD nitride. It was first anisotropically back-etched to form a 3.5 mm × 3.5 mm free standing nitride diaphragm. The front side of the diaphragm was then evaporated with 2000 Å thick Cr/Au through a physical mask to form the membrane conductor. The Teflon film was then spun on. Electrons of 10 keV energy were implanted into the Teflon using the BLT. The back plate electrode was fabricated starting with a wafer coated with 3 μm thermal oxide. Portions of the oxide layer were etched to create etching windows which extended to the silicon substrate. A timed KOH etch followed, creating a 3 μm deep recess in the silicon substrate. This forms the air gap of the capacitive microphone. A 3 μm thick thermal oxide was then grown. A 40 × 40 array of cavities for reducing the air streaming resistance [7] was then formed by anisotropic followed by isotropic etching through the patterned oxide. Each cavity has a 30 μm diameter opening and a half-dome shaped hole 80 μm in diameter and 50 μm deep. Lastly, a 2000 Å thick Cr/Au electrode in the back plate was evaporated using a physical mask. Figure 7 shows the process flow of the micromachined microphone. The fundamental resonant frequency of the microphone membrane with Cr/Au electrode and Teflon film was measured using a laser Doppler vibrometer. The fundamental resonant frequency was found to be around 8 kHz. The microphone membrane and back plate are shown in Figure 8a. A portion of the back plate cavity array with conducting gold is shown in Figure 8b. The hybrid microphone package is illustrated in Figure 9. The two halves of the microphone are mechanically clamped together and are enclosed in a metal box that provides EM shielding.

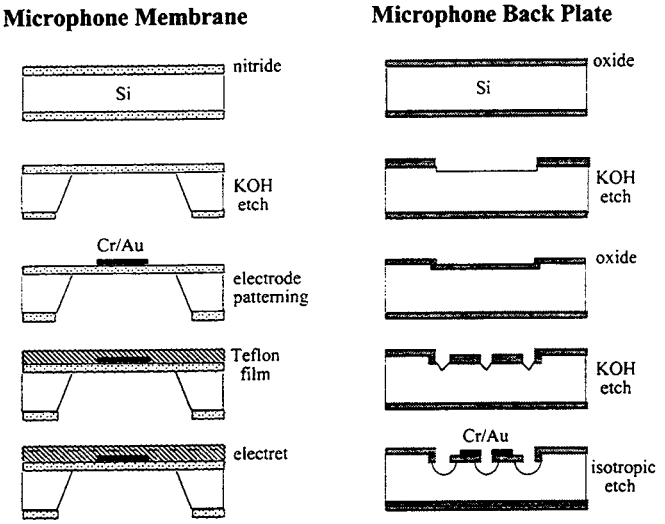
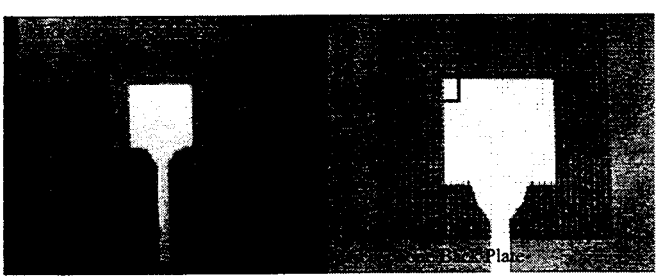
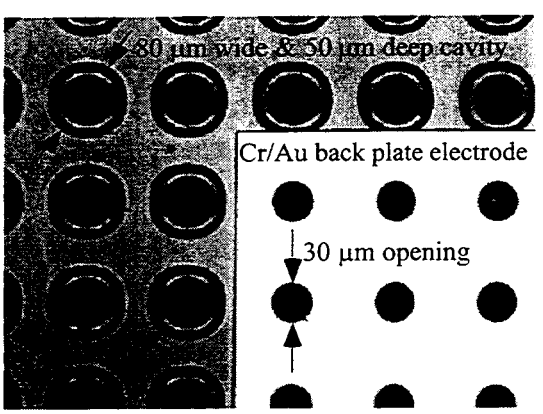


Figure 7. The process flow for the electret microphone.



(a)



(b)

Figure 8. (a) The microphone diaphragm and back plate. (b) Part of the microphone back plate cavity array with the Cr/Au electrode.

To reduce the stray capacitance, the electrode area was designed so that it only covered a fraction of the membrane and back plate area [8]. Only a 2 mm × 2 mm square Cr/Au electrode was used to cover the center part of the 3.5 mm × 3.5 mm diaphragm and 4 mm × 4 mm perforated back plate. The fraction of the back plate area occupied by the cavity openings is 0.07. The streaming resistance, R_a, is calculated to be 0.03 Ns/m [7, 9]. The cut-off frequency (f_c = 13.57 σ h /{2π R_a}, where σ = 100 MPa is the nitride diaphragm stress and h = 1 μm is the diaphragm thickness [9]) is calculated to be approximately 7.6 kHz.

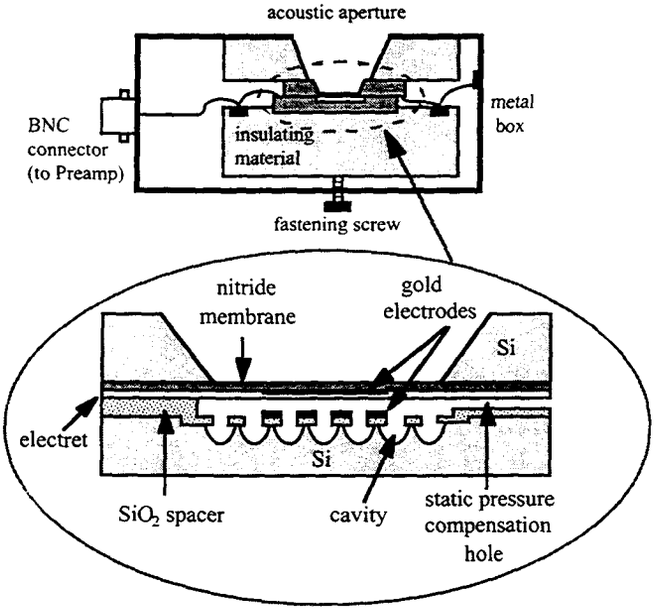


Figure 9. *The hybrid microphone package.*

The theoretical capacitance of the microphone is 7 pF with the 4.5 μm air gap, the 1 μm thick Teflon electret and an electrode area of 4 mm^2. Using a Hewlett Packard 4192 LF Impedance Analyzer the measured capacitance of the microphone package was 30 pF. The discrepancy in capacitance values can be attributed to stray capacitance between the Cr/Au electrodes and silicon substrates and between the two clamped halves of the microphone.

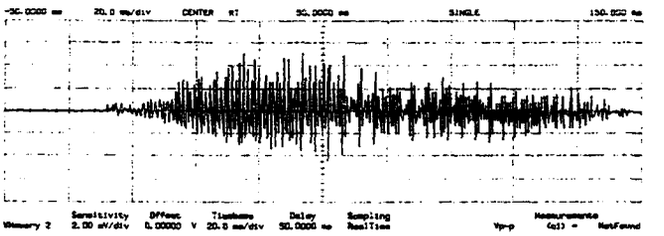


Figure 10. *An oscilloscope display of a human voice detected by the hybrid microphone without an amplifier.*

The microphone was able to detect the sound from a loud human voice without the use of an amplifier (Figure 10). When the microphone was connected to an EG&G PARC Model 113 Pre-amp (gain set at 1000) and was excited by a Brüel & Kjaer Type 4220 Pistonphone operating at 250 Hz and 123.9 dB (re. 20 μPa) amplitude, the oscilloscope displayed a 250 Hz, 190 mV peak-to-peak amplitude signal. The estimated open-circuit sensitivity of the microphone package is 0.3 mV/Pa. Taking into account the 15 pF input capacitance of the pre-amplifier and the stray capacitance of the microphone package, the open-circuit sensitivity of the microphone itself should be higher than the estimated value. The open-circuit sensitivity of the microphone can also be estimated by calculating the deflection of the electret diaphragm and the output voltage due to a sound pressure [5, 10]. Assuming piston-like movement of the conducting area of the nitride diaphragm, calculations indicate that an open-circuit sensitivity of 7 mV/Pa is achievable.

CONCLUSIONS

We have demonstrated the technology to make high charge density, thin film electrets using Du Pont Teflon® AF and the Back Lighted Thyratron. Experiments have been conducted to demonstrate the stability of the Teflon electret at room temperature and at elevated temperatures. A prototype hybrid micromachined electret microphone has been made and preliminary tests and calculations suggest potentially high sensitivities.

ACKNOWLEDGMENT

This research was funded by the David and Lucile Packard Foundation and by the NSF Presidential Young Investigator Award under NSF Grant No. ECS-9157844.

REFERENCES

1. G. F. Kirkman and M. A. Gundersen, "A Low Pressure Light Initiated Glow Discharge Switch for High Power Applications", *Appl. Phys. Lett. 49*, 494 (1986).

2. T-Y Hsu, "A Novel Electron Beam Source Based on the Back-Lighted Thyratron", Ph.D. dissertation, University of Southern California, 1992.

3. T-Y Hsu, K. Frank, and M. A. Gundersen, "Collimating Structure for Intense Pseudospark Electron Beams", *IEEE International Conference on Plasma Science*, New Mexico, June 6-8, 1994.

4. B. Hilczer, and J. Malecki, "Electrets", Elsevier Science Publishers, 1986.

5. G. M. Sessler, editor, "Electrets", Springer-Verlag Berlin Heidelberg, 1987.

6. Y. Tada, "Experimental Characteristics of Electret Generator, using Polymer Film Electrets", *Japan J. Appl. Phys. 31*, 846 (1992).

7. Z. Skvor, "On the Acoustical Resistance due to Viscous Losses in the Air Gap of Electrostatic Transducers", *Acoustica 19*, 295 (1967/68).

8. J. A. Voorthuyzen, A. J. Sprenkels, A. G. H. Van Der Donk, P. R. Scheeper and P. Bergveld, "Optimization of Capacitive Microphone and Pressure Sensor Performance by Capacitor-electrode Shaping", *Sensors and Actuator A, 25-27*, 331 (1991)

9. P. R. Scheeper, A. G. H. van der Donk, W. Olthuis, and P. Bergveld, "Fabrication of Silicon Condenser Microphones using Single Wafer Technology", *J. of Microelectromechanical Systems, vol. 1 (3)*, 147 (1992).

10. M. G. Allen, M. Mehregany, R. T. Howe, and S. D. Senturia, "Microfabricated Structures for the In Situ Measurement of Residual Stress, Young's Modulus, and Strain of Thin Films", *Appl. Phys. Lett. 51 (4)*, 241 (1987).

A Micromachined Silicon Condenser Microphone With On-Chip Amplifier

Jonathan J. Bernstein and Jeffrey T. Borenstein

The Charles Stark Draper Laboratory, MS 37
555 Technology Square, Cambridge MA 02139-0803

ABSTRACT

Monolithic silicon microphones have been fabricated which include on-chip JFET buffer amplifiers, providing a complete microphone plus electronics on a single silicon chip. Test data is presented showing sensitivities as high as 40 mV/Pa (-48 dB ref. 1V/μBar), with routine sensitivity of 15 mV/Pa (-57 dB ref. 1V/μBar) and bandwidth of 70 Hz to 20 kHz. Noise measurements show that electronics noise from the gate bias resistor is dominant, leading to $1/f^2$ noise spectral density.

The fabrication of these microphones uses both surface and bulk micromachining (Fig. 1). Surface micromachining (precision electroforming) is used to make the perforated bridge electrodes, while bulk micromachining is used to form the diaphragm and through-hole. Diaphragms of P^+ doped silicon from 1 to 2 mm in diameter have been tested. A monolithic ion implanted JFET and polysilicon resistors are fabricated on-chip.

An analysis of the theoretical attainable sensitivity of a micromachined condenser microphone will be presented. A theoretical Figure of Merit proportional to the microphone signal to noise ratio is derived, which allows the microphone design to be optimized.

INTRODUCTION

Today's microphones are of four types: dynamic, condenser, condenser - electret, and piezoelectric. Most previous micromachined microphone work **[1-14]** has focused on condenser microphones because of their high sensitivity, simplicity, potential for miniaturization, and compatibility with thin film fabrication techniques.

Condenser microphones use a thin membrane as one plate of a variable capacitor, and can achieve sensitivities as high as 100 mV/Pa with a 25 mm diameter diaphragm. A bias voltage is applied to the sense capacitor, which produces an output voltage when attached to a high input impedance buffer amplifier.

Condenser-electret microphones use an electret material in the sense gap to produce the bias electric field, hence no polarizing voltage is required. Electrets render the bias voltage unnecessary, but it is not clear if the electret process is compatible with on-chip IC processing. Electrets are a major advantage in low cost battery powered systems where providing the bias voltage (5 to 20 volts) is an undesirable cost.

The work reported here concerns the fabrication of silicon condenser microphones and their optimization using a monolithic JFET source follower amplifier.

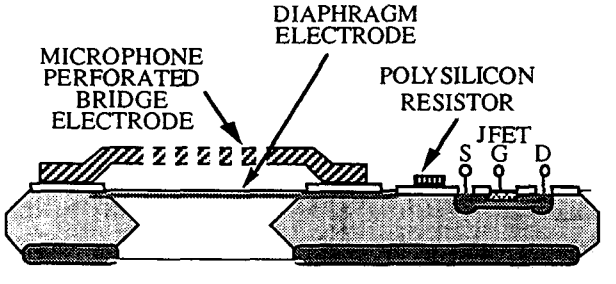


Fig. 1. Cross section of monolithic microphone with on-chip JFET and poly-resistor.

THEORY

The theory of condenser microphones is well developed in the literature [**15**]. What follows is a procedure for optimizing the design given system constraints such as bias voltage, buffer input capacitance, maximum SPL and package volume.

The diaphragms used in this work are sufficiently complicated that they can only be analyzed by finite element analysis. However, the theoretical analysis of the microphone is best carried out using a simple spring-mass-damper model in which the diaphragm is assumed to move as a rigid piston. The equivalent compliance C_{SP} (m^3/Pa) of a real diaphragm can be calculated (using FEA) as the volume displacement divided by the static pressure.

The acoustic compliance C_{SP} is given in terms of an average spring constant k_{SP} (N/m) as:

$$C_{SP} = \frac{A^2}{k_{SP}} \tag{1}$$

The compliance of the package interior is given in terms of the internal volume V_{CAV}, atmospheric pressure P_0 and the ratio of specific heats γ:

$$C_{CAV} = \frac{V_{CAV}}{\gamma \cdot P_0} \tag{2}$$

The series compliance of the diaphragm and cavity is

$$\frac{1}{C_{SER}} = \frac{1}{C_{CAV}} + \frac{1}{C_{SP}} . \tag{3}$$

The gap capacitance is $C_{gap} = \frac{\varepsilon_0 A}{x_0}$. The total inertance L_{TOT} (kg/m^4) is given by the sum of the diaphragm plus radiation inertance:

$$L_{TOT} = \frac{\rho_{Si} \cdot t}{A} + \frac{1.7 \cdot \rho_{air}}{\sqrt{\pi \cdot A}} \tag{4}$$

where $\rho_{Si} * t$ is the mass/area of the diaphragm. The fundamental resonant frequency is given by

$$\nu_0 = \frac{1}{2\pi\sqrt{L_{TOT} \cdot C_{SER}}} . \tag{5}$$

The collapse voltage of the capacitor is given by

$$V_C^2 = \frac{8 \cdot x_0^3 \cdot A}{27 \cdot \varepsilon_0 \cdot C_{SP}} . \tag{6}$$

The sensitivity M (V/Pa) of a condenser microphone is given by [15, 16]

$$M = \left(\frac{8 \cdot V_{Bias} \cdot x_0^2}{27 \cdot \varepsilon_0 \cdot V_C^2}\right) \cdot \left(1 + \frac{C_{SP}}{C_{CAV}}\right)^{-1} \cdot \left(1 + \frac{C_L}{C_{Gap}}\right)^{-1} \tag{7}$$

0-9640024-1-8 $3.00 © 1996 TRF

Solid-State Sensor and Actuator Workshop
Hilton Head, South Carolina, June 2-6, 1996

where C_L is the load capacitance including both stray and buffer amplifier input capacitance. The advantages of writing the sensitivity in this form are that the dependence on cavity volume and load capacitance is very clear. If the package is made too small, or the load capacitance is made too large, the signal is reduced.

Eq. 6 can be used to eliminate C_{SP} in favor of the collapse voltage, which is a typical product specification. This allows the sensitivity to be plotted as a function of area A and zero bias gap x_0 , with bias voltage, collapse voltage, cavity volume, and load capacitance as parameters. The number of parameters can be further reduced by taking V_{bias} as a fixed fraction of V_C , e.g., $V_{BIAS} = 0.6 * V_C$.

Fig. 2 shows a plot of sensitivity vs. area and gap, using a collapse voltage of 2 volts and a bias voltage of 1.2 volts, a cavity volume of 15 mm^3, and a load capacitance of 0.6 pF. Such small bias voltages are interesting because of the possibility of using these microphones in small battery powered applications such as hearing aids. The maximum sensitivity is 32 mV/Pa, which is quite respectable for a hearing aid microphone.

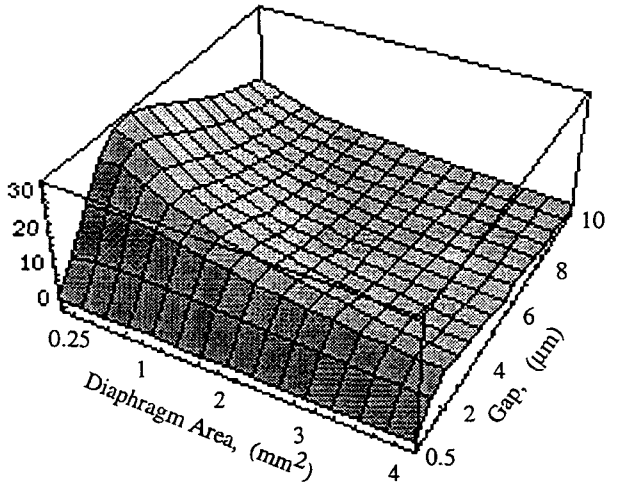


Fig. 2. Sensitivity (mV/Pa) vs. diaphragm area and sense gap for a condenser microphone with V_{Bias} = 1.2 V and package volume = 15 mm^3.

Fig. 3 shows sensitivity vs. area and gap using a package volume of 40 mm^3, a collapse voltage of 8.3 volts and a bias voltage = 5 volts, all other parameters the same as Fig. 2. The peak sensitivity is 56 mV/Pa. This is much higher than available in current commercial electret microphones with wide bandwidth (20 kHz).

The noise of a condenser microphone with source follower buffer amplifier is usually dominated by resistor and JFET noise. Gate leakage current at the JFET and other current leakage terms lead to white current noise which can be lumped in with the gate resistor noise term. With careful design, the Brownian motion limit (due to air-damping of the diaphragm) can be achieved for higher frequencies.

The voltage noise V_n^2 (V^2/Hz) from the JFET and gate resistor is [10]

$$V_n^2 = \left(\frac{C_{Gap}+C_L}{C_{Gap}}\right)^2 \cdot \left(\frac{K}{f}+\frac{8kT}{3\cdot g_m}\right)+\frac{4kT}{R_g\cdot\left(\omega\cdot C_{Gap}\right)^2} \quad (8)$$

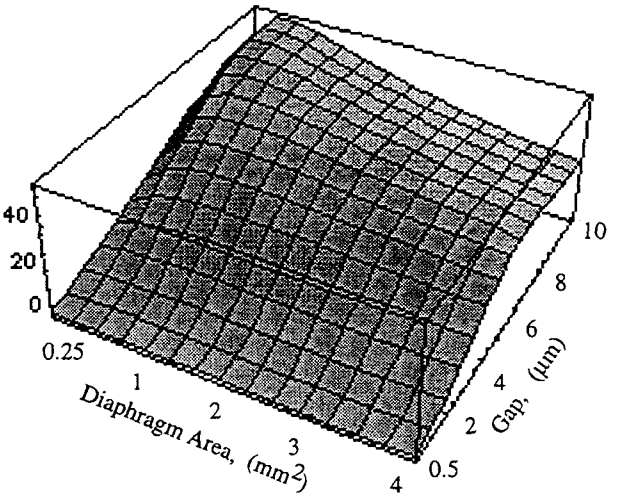


Fig. 3. Sensitivity (mV/Pa) for microphone with bias voltage of 5 V, package volume = 40 mm^3.

where $\omega = 2\pi f$, g_m is the JFET transconductance, k is Boltzman's constant, C_L is the buffer plus stray capacitance and K is the flicker noise coefficient of the JFET.

From Eqn. 8, we see that the resistor noise term (which is usually dominant) is inversely proportional to the microphone capacitance, hence the S/N ratio of the microphone is proportional to the product of sensitivity and output capacitance. The charge figure of merit $FOM_Q = M^*C_{GAP}$ is therefore the quantity to maximize if maximum S/N is the goal. Given an input sound pressure P_{in} and considering only the gate resistor noise, we have

$$\frac{S}{N} = \frac{P_{in}\cdot M\cdot C_{Gap}\cdot\sqrt{R_g}}{\omega\sqrt{4kT}} \quad . \quad (9)$$

Fig. 4 shows FOM_Q for the 5 volt bias case. The optimum design for S/N is shifted towards larger capacitance, resulting in an optimum at x_0 = 4 µm and A = 4 mm^2, whereas the sensitivity is maximized at x_0 = 9 µm and A = 0.8 mm^2. The sensitivity at optimum S/N drops roughly in half to 25 mV/Pa, but the S/N is increased by a factor of 5.5 compared to the sensitivity maximum. If a post amplifier with noise level lower than the source follower is available, then the overall S/N of the system can be improved significantly by choosing the optimum FOM_Q design.

The microphone must operate at a maximum pressure P_{max} without collapsing or excess distortion. If we assume that the gap changes 10% at the maximum pressure, we can derive a very simple and useful equation bounding the performance of the microphone:

$$P_{MAX} = \frac{A\cdot x_0}{10\cdot C_{SER}} \quad (10)$$

$$M \le \frac{V_{Bias}}{10\cdot P_{MAX}} \quad (11)$$

Therefore once the bias voltage and maximum SPL are given, the maximum sensitivity is determined.

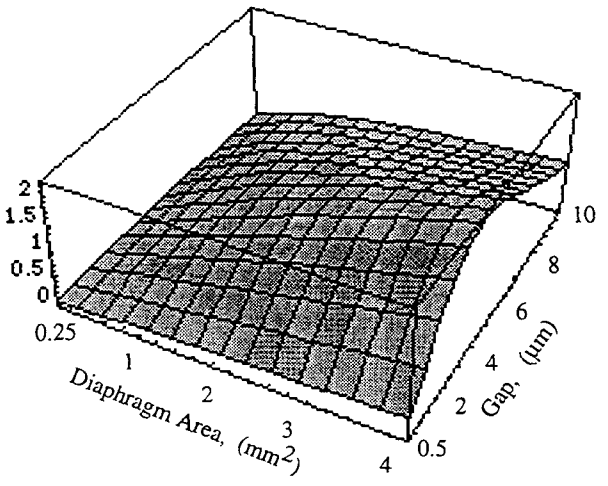


Fig. 4. Charge Figure of Merit (FOM_Q, 10^{-13} C/Pa) for microphone. Parameters same as Fig. 3.

EXPERIMENTAL DETAILS

Fabrication

Fabrication started with 3" n-type silicon wafers. A thermal field oxide 700 nm thick was used to mask a P^+ solid source boron diffusion used for the acoustic membranes. The same diffusion was used for the source and drain of the JFETs.

Polysilicon 0.5 µm thick was deposited for both doped and undoped poly-resistors. These were implanted with phosphorus through a photoresist mask and then dry patterned by RIE using SF_6.

Channel and gate of the P-channel JFETs were formed by ion implantation of boron and arsenic, respectively. After an implant activation anneal, metal 1 (TiW/Au) was deposited and patterned using liftoff photolithography.

Electroplated gold bridge electrodes were formed using standard thick photoresist processing [**13, 17**]. A seed layer of Ti/Au was sputtered on a temporary photoresist spacer layer. The plating mold resist (25 µm thick) was spun on top of the plating seed layer, and patterned. Gold was electroplated into the mold to a thickness of 20 microns. The plating mold, seed layer, and spacer layer were then stripped in acetone, gold etch, Ti etch, and positive resist stripper respectively.

The final fabrication step is a six-hour anisotropic etch in EDP (Ethylene-Diamine, Pyrocatechol, and water), which creates the through-wafer holes and frees up the diaphragm. V-grooves are formed which allow the chips to be snapped apart without dicing.

Test Procedures

The microphone sense capacitors were characterized by an I-V (snap-down) test and by capacitance-voltage curves, using an EG&G PAR 410 C-V meter controlled by an HP 4145A semiconductor parameter analyzer. The snap-down voltage was measured using 40 MΩ series resistors to avoid harming the electrodes.

JFETs were characterized using the HP 4145A to determine SPICE parameters such as LAMBDA, R_{OUT} and I_{DSS}. A program was written to plot both $\sqrt{I_D}$ and I_G vs. V_G, with V_D held at -9 volts. The intercept of the $\sqrt{I_D}$ vs. V_G curve with zero current was used to determine the pinchoff voltage. Gate leakage current is measured at the same time. DC-coupled gain of the source followers was measured from gate to source as 0.91 to 0.94. Using a series 5 pF capacitor, the AC coupled gain was measured as 0.84 to 0.86. This allows a computation of the effective input capacitance of the active source follower as 0.5 pF.

Microphones were packaged in custom small kovar packages with holes drilled for acoustic inlets. Wirebonding was used to attach the JFET to the four external leads. The acoustic inlets were covered with small circles of clean room paper to prevent dust from entering.

An acoustic test station was used to calibrate sensitivity. Two 9-cm thick layers of Sonex (Illbruck, Inc.) acoustic foam line a $(60\ cm)^3$ wood box containing a small headphone speaker, reference microphone and DUT. Miniature optical rail was used to hold the speaker and a black thread cross-hair used as a position reference to repeatably locate the microphones. Microphones were suspended by their cables from the top of the test chamber.

The speaker is driven by the output of an HP 3562A Dynamic Signal Analyzer. A B&K 4135 condenser microphone and type 2804 pre-amp were used as a reference microphone with sensitivity of 3.2 mV/Pa. The DUT and reference microphone were positioned adjacent to each other and 13 cm from the speaker. Simultaneous measurement did not introduce any errors below 20 kHz, compared to measuring each microphone alone and forming the ratio of the responses.

TEST RESULTS

JFET's

The ion implanted JFET test results are summarized in Table 1. Leakage current on some wafers averaged 0.25 pA, while on other wafers was as high as 30 - 40 pA. The variation in leakage current is probably due to heavy metal decoration of dislocations resulting from the boron diffusion.

Low noise JFET's are typically formed in epitaxial silicon with no implants, however, good low leakage JFET's can be formed by ion implantation [**18-20**]. An implanted channel and gate were chosen for this process to avoid an epi step and enable tight control over threshold voltage. Pinchoff voltage is chosen at 1 V to allow FET operation with a single dry cell, although a larger bias voltage (5-10 V) is typically required to obtain good sensitivity.

Table 1: JFET Parameters

Parameter	Typical Value
W/L	576µm/3µm
Gate resistance	12 GΩ
Source resistance	50 kΩ
V_{PO}	1.0 V
I_{DSS}	680 µA
Early Voltage	20 V
Effective Input Capacitance	0.5 pF
DC coupled gain	0.91 - 0.94

Figure 5 shows a microphone chip with a 1.8 mm diaphragm and radial support springs. Several alternative JFET buffer amplifiers are located on the device edges. Fig. 6 is a close-up of a microphone with a 1 mm diaphragm, showing the bridge electrode and an oval-shaped JFET.

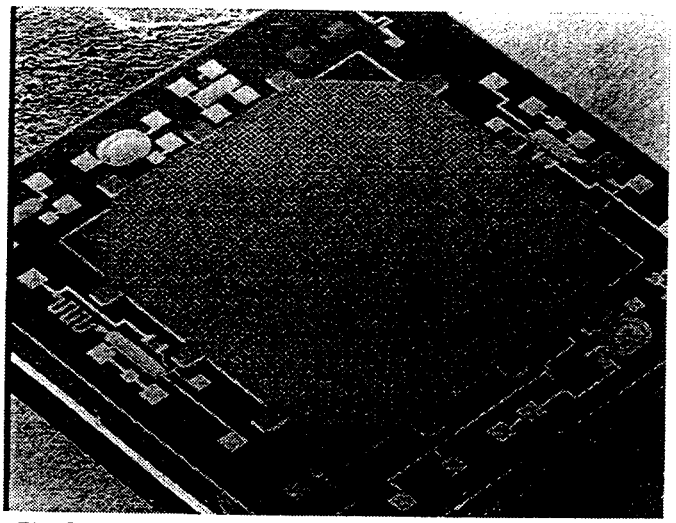


Fig. 5. Microphone with 1.8 mm Diaphragm on a 3 mm chip, showing perforated, electroplated bridge electrode and various designs of buffer amplifiers at periphery.

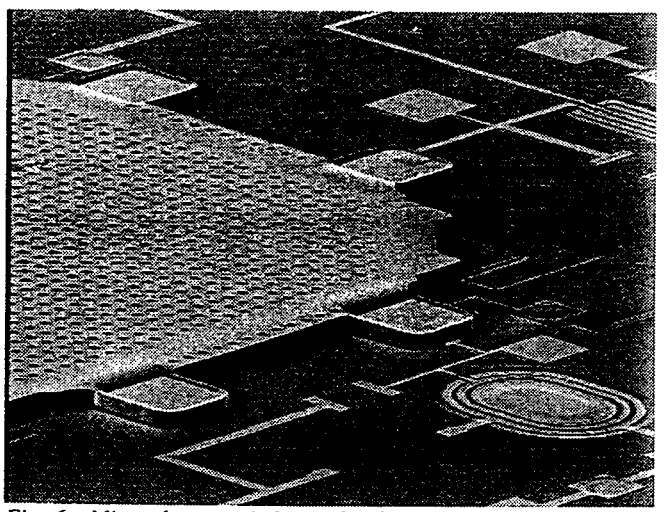


Fig. 6. Microphone with 1 mm diaphragm, showing close-up of oval JFET and perforated bridge electrode.

Sensitivity vs. bias voltage is shown in Fig. 7 for the 1.8 mm diameter radial spring design. This data was taken using hybrid JFET chips sawn from JFET process development wafers.

The frequency responses of 1.0 mm and 1.8 mm diameter diaphragm microphones is shown in Fig. 8a and b. The fundamental resonance of each is at 15 kHz, and sensitivity is 16 mV/Pa mid-band.

Environmental testing was performed on prototype microphones by several outside laboratories. The microphones survived both "cap-gun" and "balloon-burst" stress testing. Microphones were exposed to 95% R.H. at 60°C for 2 weeks, with no change in sensitivity. Total Harmonic Distortion was less than 0.5% at 1 kHz and 100 dB SPL. These microphones and pre-amps had no protective passivation layers, hence units with passivation layers are expected to better withstand environmental stress.

NIST [21] also performed calibration tests on early prototypes, confirming the sensitivity data of Draper's in-house calibration tests. THD was shown to be < 1% at 1 kHz and 114 dB SPL. A-weighted noise of these devices was as low as 25 dB SPL.

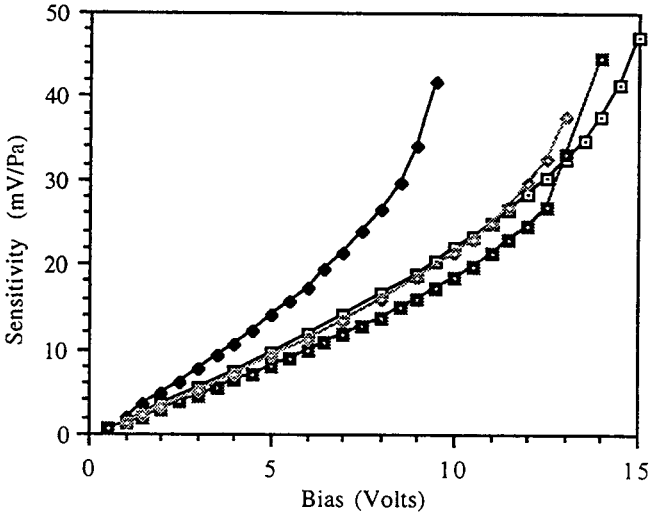


Fig. 7. Sensitivity vs. bias voltage for several microphones with 1.8 mm diameter diaphragms and 4 radial springs.

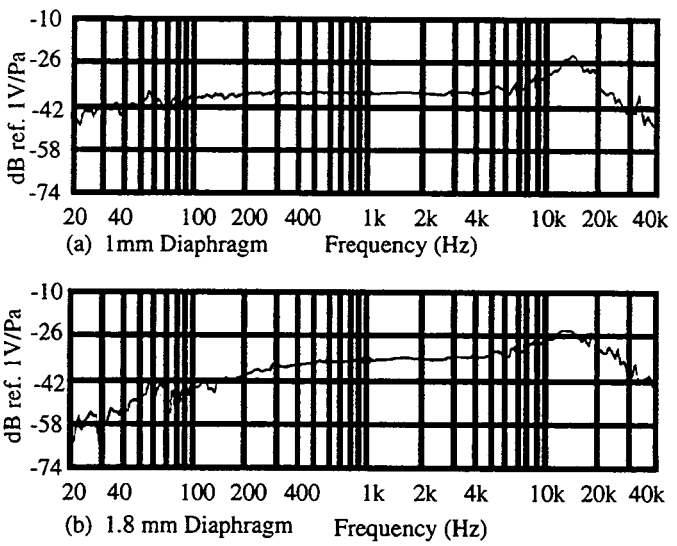


Fig. 8. Frequency response of microphones with (a) 1 mm diaphragms, and (b) 1.8 mm diaphragms.

CONCLUSIONS

Monolithic silicon microphones with on-chip JFET source followers have been fabricated and tested. The monolithic buffer amplifier results in the lowest stray capacitance and good control over leakage currents. Future developments may include CMOS circuitry if the trend towards system integration on a chip continues, or a hybrid system may be more economical.

According to theoretical analysis, a bias voltage of 5 to 10 volts is required to achieve high sensitivity, wide dynamic range and high signal to noise ratio. A voltage multiplier circuit will likely be necessary for single cell applications such as in-the-ear hearing aids.

The sensitivity and bandwidth of these microphones can be optimized for any application by changing the diaphragm compliance, area, gap and damping. The bandwidth demonstrated on prototype devices (20 kHz) is wider than standard hearing aid microphones (typically 6 kHz), yet their sensitivity is higher than the low cost electrets. These microphones could find a niche in voice recognition or other

applications where both high fidelity and moderate cost are required.

ACKNOWLEDGMENTS

The authors gratefully acknowledge support from Noise Cancellation Technologies, Inc. The authors would also like to thank Alex Murphy for microphone testing, and the entire Draper fabrication team for processing the microphones.

REFERENCES

1. D. Hohm and G. Hess, "A subminiature condenser microphone with silicon nitride membrane and silicon back plate", J. Acoust. Soc. Am. 85 (1), pp 476-480, (Jan. 1989).
2. P. Murphy, K. Hubschi, N. de Rooij, and C. Racine, "Subminiature Silicon Integrated Electret Capacitor Microphone", IEEE Trans. on Electr. Insul., V. 24 No. 3, pp. 495-498, (June 1989).
3. J.A. Voorthuyzen, P. Bergveld and A.J. Sprenkels, "Semiconductor based Electret Sensors for Sound and Pressure", IEEE Trans. on Electr. Insul., Vol 24 No 2, pp 267-276, (Apr. 1989).
4. J. Bergqvist and F. Rudolf, "A New Condenser Microphone in Silicon", Sensors and Actuators, A21-A23, pp. 123-125, (1990).
5. P.R. Scheeper, J.A. Voorthuyzen, and P. Bergveld, "PECVD Silicon Nitride Diaphragms for Condenser Microphones", Sensors and Actuators B, 4, pp 79-84, (1991).
6. W. Kuhnel and G. Hess, "Micromachined subminiature condenser microphones in silicon", Sensors and Actuators A, 32, pp 560-564, (1992).
7. A.G.H. van der Donk, P. Bergveld, and J.A. Voorthuyzen, "Optimal design of an electret microphone metal-oxide semiconductor field-effect transistor amplifier", J. Acoust. Soc. Am. 91 (4), Pt. 1, pp. 2261-2269, (Apr. 1992).
8. J.J. Bernstein, "A Micromachined Condenser Hydrophone," Solid State Sensor and Actuator Workshop, Hilton Head Island, June 1992.
9. E. Graf, W. Kronast, S Duhring, B. Muller, and A. Stoffel, "Silicon membrane condenser microphone with integrated field effect transistor", Sensors and Actuators A, 37-38, pp 708-711 (1993).
10. P. Bergveld, "The merit of using silicon for the development of hearing aid microphones and intraocular pressure sensors", Sensors and Actuators A, 41-42, pp. 223-229, (1994).
11. S.S. Lee, R.P. Ried, and R. M. White, "Piezoelectric Cantilever Microphone/Speaker", 1994 Solid State Sensor and Actuator Workshop (Hilton Head, SC) pp. 33-37, (June 1994).
12. R.P. Ried, E.S. Kim, D.M. Hong, and R.S. Muller, "Piezoelectric Microphone with On-Chip CMOS Circuits", J. Microelectromechanical Syst. Vol. 2, No. 3, pp. 111-120, (Sept. 1993).
13. J.J. Bernstein, M. Weinberg, E. McLaughlin, J. Powers and F. Tito, "Advanced Micromachined Condenser Hydrophone", Technical Digest of the Solid State Sensor and Actuator Workshop, Hilton Head Island, pp 73-77 (June 1994).
14. J.J. Bernstein, "A Micromachined Silicon Condenser Microphone and its Potential for Hearing Aid Applications", 130th Meeting of the Acoustical Society of America, St. Louis, MO, (Nov. 1995).
15. L.E. Kinsler, A.R. Frey, A.B. Coppens, and J.V. Sanders, "Fundamentals of Acoustics", 3d edn., (J. Wiley and Sons, New York, 1982), pp. 193, 349-351, 376-380.
16. R. Brander, Beltone Laboratories, private communication.
17. B. Lochel, A. Maciossek, H.J. Quenzer, and B. Wagner, "Ultraviolet Depth Lithography and Galvanoforming for Micromachining", J. Electrochem. Soc., Vol 143 No. 1, pp. 237-244, (Jan. 1996).
18. C.J. Kircher, "Comparison of leakage currents in ion-implanted and diffused p-n junctions", J. Appl. Phys., vol. 46, pp. 2167-2173, (1975).
19. J.P. Close and L.W. Counts, "A 50-fA Junction Isolated Operational Amplifier", IEEE Journ. of Solid State Circuits, v. 23, No. 3 (June 1988).
20. T. Jung, "A fully integrated cryogenic, charge sensitive low noise preamplifier using N-channel JFETs and polysilicon resistors," PhD thesis, EECS, University of Wisconsin-Madison, 1994.
21. V. Nedzelnitsky, "Micromachined silicon microphones and hearing aids: Performance and Potential", 130th Meeting of the Acoustical Society of America, St. Louis, MO, (Nov. 1995).

TOUCH MODE CAPACITIVE PRESSURE SENSORS FOR INDUSTRIAL APPLICATIONS

W.H.Ko, Qiang Wang and Yang Wang

Department of Electrical Engineering & Applied Physics
Case Western Reserve University
Cleveland, OH 44106

ABSTRACT

The principle, simulation, design, characteristics and application of touch mode capacitive pressure sensors are presented. In touch mode operation, the diaphragm of the capacitive pressure sensor is touching the substrate structure. The advantages of this mode of operation are: near linear output, large over-range pressure and robust structure that make it capable to withstand harsh industrial field environment.

INTRODUCTION

Capacitive pressure sensors are known to have no turn-on temperature drift, high sensitivity, robust structure and less sensitive to environment effects. However its output is nonlinear with respect to input changes and the sensitivity in the near linear region is not high enough to ignore many stray capacitance effects. In the normal mode of operation, the diaphragm is kept at a distance away from the substrate as shown in Fig. 1-a. If the sensor is deigned to operate in the pressure range where the diaphragm is allowed to contact the substrate with a thin layer of insulator (t_m), as shown in Fig. 1-b, then the device is a touch mode capacitive sensor. The touch mode device was developed to withstand harsh industrial environment, and with one or two orders of magnitude higher sensitivity than the normal mode near linear operation, so that some of the stray capacity effects can be neglected. This paper presents the principle, computer simulation, design, characteristics and application of touch mode capacitive pressure sensors.

PRINCIPLE

The basic element of a capacitive pressure sensor is an equivalent parallel plate capacitor with clamped edges, where the diaphragm would deform responding to a differential pressure applied to two sides of the diaphragm, as shown in Fig. 1. The capacitance, neglecting the fringe effect, is

$$C = \varepsilon \frac{A}{d} \tag{1}$$

where ε is the permittivity of the media between the two plates; A is the area of the electrode plate; d is the gap space between the two plates. The upper plate of the capacitor, known as the diaphragm, deforms when a differential pressure between the external environment and the inside chamber is applied. The general equation relating the deflection of a rectangular diaphragm, without residual stress, in normal operation region can be expressed as:

$$D\left(\frac{\partial^4 w}{\partial x^4} + 2\frac{\partial^4 w}{\partial x^2 \partial y^2} + \frac{\partial^4 w}{\partial y^4}\right) = P + N_x \frac{\partial^2 w}{\partial x^2} N_y \frac{\partial^2 w}{\partial y^2} + 2N_{xy} \frac{\partial^2 w}{\partial x \partial y} - x\frac{\partial w}{\partial x} - y\frac{\partial w}{\partial y} \tag{2}$$

where D is flexural rigidity, $D=E/[12(1-\nu^2)]$; P is the differential pressure; w is deflection at point (x,y); Nx, Nx and Nxy are direct and shear stresses in a plane parallel to the plate surface; E is Young's modulus; and ν is Poisson's ratio.

In the normal operation mode of a capacitive sensor, the diaphragm does not contact the substrate electrode. The output capacitance is nonlinear due to its inverse relationship with the gap ($d_0 - w$, d_0 is initial gap), which is a function of pressure P, as given in equation (2). This nonlinearity becomes significant for large deflection, [$(w_o/h) > 0.3$, w_0 is center deflection], and large sensing capacitance regions. Many efforts have been made to reduce the nonlinear characteristics of capacitive sensors either by modifying the structure of sensors or by using special non-linear converter circuits. In the touch mode capacitive sensor, as shown in Fig. 1-b, the major component of the sensor capacitance is that of the touched area where the effective gap is the small thickness of the thin insulator layer, t_m, on the substrate electrode. In this touch mode operation region, the capacitance varies with pressure nearly linearly and the sensitivity (dC/dP) is much larger than that in the near linear region of a normal mode device. A typical C-P characteristic of a capacitive pressure sensor covering normal and touch mode regions is shown in Fig. 2. It has four regions, i.e. normal, transition, linear and saturation regions. The touch mode capacitive pressure sensors (TMCPS) operate in the region III-linear region. They were developed to meet the manufacturing and operational conditions of industrial applications. Its cross-section is shown in Fig. 1-b. After the diaphragm touches the substrate and as the pressure increases, the sensor capacitance is mainly determined by the capacitance of the touched area instead of the capacitance in the untouched "normal operation area". In region III, Fig. 2, the change of touched area is almost proportional to pressure, thus the C-P characteristics is nearly linear. The support of the substrate to the diaphragm after touch, enables the devices to have very large over-load protection. In summary, the advantages of TMCPS are: nearly linear C-P characteristics, large overload protection, high sensitivity and simple robust structure that can withstand industrial handling and environment.

COMPUTER SIMULATION — FINITE ELEMENT MODELING

It is important to have a good understanding of the deflection, stress and strain of the diaphragm in order to design capacitive pressure sensors properly. The output capacitance of a

0-9640024-1-8 $3.00 © 1996 TRF

capacitive pressure sensor can be calculated by integration over the deformed diaphragm with a series of equivalent parallel plate elementary capacitors where the effective gap is the difference between the zero pressure gap and the local deflection. In the normal mode, the deflection can be calculated from equation (2) by numerical approximation analysis [1, 2]. In the touch mode device, however, equation (2) is no longer valid. Hence the finite element modeling (FEM) of diaphragms is used to determine the touch mode deflection of the diaphragm. A uniform mesh of 24 x18 elements, with 8 nodes for each element, in a quarter of the diaphragm is used as shown in Fig. 1-c. By geometrically nonlinear computing using ABAQUS, the deflection, stress and capacitance versus pressure at each node can be calculated for a set of sensor parameters *a, b, h,* and *d,* as shown in Fig. 1, and material properties *E* and *v*. GAP element in ABAQUS is used to model the touch mode operation. Post data analyses based on information obtained on deflection at each node is developed to find the capacitance-pressure behaviors to predict the performance of simulated capacitive sensors [3].

From the FEM simulation, the performance of the capacitive sensor can be estimated as the device parameters are varied. The dimensionless deflection at the center point of diaphragms is shown in Fig. 3, where the normalized deflection ratio (w_o/h) is plotted against the normalized pressure (Pa^4/Eh^4) for two (b/a) ratios. The curves in Fig. 3 can be used to determine the touch point pressure (P_t) of known diaphragms by letting the center deflection equal to the initial gap (d_0). For a set of (a,b) and diaphragm material, the touch point can be selected by (h) and (d) using curves in Fig. 3. After touch, the stress in the touched area is tensile and the stress in the untouched area has three sub-regions as shown in Fig. 4. Fig. 5 shows the deflection along y-axis as the pressure is increased in equal steps. The deflection along x-axis has similar behavior. Beyond the touch point, the shape of the untouched portion of the diaphragm remains nearly the same. However, the touched area *A* equals to ($a + bP - cP^3$) where a,b,c are constants [3], it increases nearly linearly at first, then approaches saturation for larger pressures. This explains the C-P characteristics in region III and IV in Fig. 2.

The average sensitivity (dC/dP) is defined as the slope of the straight line obtained by linear curve fitting over the working pressure range. The nonlinearity is defined as

$$\frac{|\Delta C_+| + |\Delta C_-|}{2(C_{max} - C_{min})} 100\% \qquad (3)$$

where ΔC_+ and ΔC_- are the maximum and minimum deviation from a straight line characteristic, respectively; C_{max} and C_{min} are the maximum and minimum capacitance in the operation range, respectively.

It is found that the sensitivity is proportional to ($1/t_m$). The sensitivity is also inversely proportional to the thickness of the diaphragm for a fixed touch point as shown in Fig. 6. Fig. 7 shows (dC/dP) sensitivity and nonlinearity versus touch point pressure.

These simulation results can predict the trend of the sensor's performance when the device parameters are varied. The understanding and quantitative curves can be used as a design tool to arrive at a specific desired performance. The results of the diaphragm deflection and stress in both touched and untouched regions are useful for micro-actuators incorporating diaphragms.

FABRICATION AND PACKAGE

For molded in-package pressure monitoring used in industries, the sensors have to survive high temperature, high pressure manufacturing processes, and to operate in harsh environment of repeated shock, vibration, stress, and temperature cycling for tens of years. It desires high sensitivity and good linearity to simplify the reading circuitry. The TMCPS was developed to meet these requirements. A rectangular diaphragm was selected to reduce the effect of the maximum stress near the corners on the sensor performance. With the help of finite element modeling, the parameters of TMCPS can be designed to meet device performance specifications and fabrication constrains. A flow chart indicating major fabrication process steps for the sensor chips is given in table 1. An isolation layer, t_m, is deposited over the substrate electrode to provide isolation in touched range and to aid the anodic bonding of silicon to the glass substrate to form a hermetically sealed cavity for absolute pressure measurement. This layer plays an important role in the TMCPS. Sputtered #7740 Pyrex glass is used as the insulating layer. The substrate material could be glass or silicon. A special process was developed to fabricate hermetic electrical feed-throughs to connect the electrodes inside the hermetically sealed cavity to external bonding pads. With carefully selected thickness of the isolation layer, #7740 glass substrate with thin sputtered glass can be anodically bonded to the silicon to achieve hermetical sealing [4].

The individual sensor chip is to be packaged on printed circuit board or ceramic substrate together with other components of the system. Multiple layer package process was developed to protect the sensor from shock and interfering stress, and yet to transmit the pressure within the desired accuracy over the life of many decades. The resulting package is shown in Fig. 8. It was molded at 300°C, 400% overpressure for several hours and then evaluated in the field with no degradation in performance [4,5].

EXPERIMENTAL RESULTS

Hundreds of TMCPS, with device parameters and material properties given in Table 2, have been fabricated in Electronics Design Center, CWRU, for laboratory and field tests in the last two years. A typical P-V characteristic is shown in Fig. 9. The CP-10 CMOS C-V converter circuit [6] is used in the sensor evaluation to convert capacitance into voltage. The measured non-linearity, as defined in equation (3), is 0.6% in the rang of 80-110 PSI.

The hysteresis was found in some sensors. The hysteresis of sensors is defined as the difference between the outputs when the pressure is cycled upward and downward, which can be expressed as follows:

$$H = \frac{\max|V_F - V_B|}{V_{max} - V_{min}} 100\% \qquad (4)$$

where V_F and V_B are sensor's outputs when the forward and backward pressure are applied, respectively.

The hysteresis of TMCPS is mainly due to the quality of the isolation layer. Since the diaphragm presses against this layer during the operation, non-elastic deformation of the sputtered glass layer results in the sensor hysteresis. However, after a few high pressure cycles (400% overpressure), the glass layer is stabilized hence the hysteresis is reduced significantly. Two

hysteresis curves of a typical sensor is shown in Fig. 10. After a few 400PSI pressure cycles, the hysteresis can hardly be observed.

The sensors have been successfully molded in the industrial packages and are undergoing simulated life tests in proving grounds. Fig. 11 shows a TMCPS' P-V characteristic before and after molding process, which consists of 300°C and 400% of full scale overpressure for one to two hours [5].

CONCLUSION

The FEM and computer simulation, design, fabrication, evaluation and field testing of touch mode capacitive pressure sensors in industrial applications are reported. The absolute pressure measuring device is demonstrated to have the advantages of good stability, low power consumption, robust structure, large overload ability and high (dC/dP) sensitivity. The basic device design can be used to measure pressure from 10^{-4} to 10^{3} PSI full scale with only changes of several process parameters. It can withstand manufacturing package temperature up to 300°C for several hours after packaging, and may have 200% to 200,000% full scale over pressure protection. This new mode of operation may supplement the piezoresistive and normal mode capacitive pressure sensor in industrial applications where mechanical and electrical stability are important.

The initial step in developing a CAD simulation program for capacitive pressure sensor has been made. The modeling of touch mode capacitive sensors (neglecting the build-in stresses of the diaphragm) was developed to obtain a set of data base for a family of sensors over a range of device parameters. From the data accumulated, the effects of device parameters on the device performance can be estimated. These results would be useful to capacitive sensor designers in estimating the performance of device before fabrication and also may serve as a guide for adjustment of parameters to arrive at a desired performance in mass production.

The design and simulation of the diaphragm would be applicable to diaphragms for other sensors, actuators, and microsystems. The understanding on the deflection and stress of a diaphragm under various loads, with small or large deflection, will be useful for diaphragm design to accurately predict the behavior of the element under normal or touch modes operation.

REFERENCES

1. S.P. Timosheko and S. Woinowsky-Krieger, Theory of Plates and Shells, 2nd Ed. (McGraw Hill, New York, 1970).

2. M.D. Giovanni, Flat and Corrugated Diaphragm Design Handbook, (Mercel Dekker, New York, 1982) p. 178-206.

3. Xiang X. Huang, "Modeling of Large Deflection and Touch Mode Capacitive Sensors", MS Thesis, EEAP Dept., Case Western Reserve Univ., Cleveland, Ohio, USA, (1994).

4. Yang Wang, "Industrial Capacitive Pressure Sensor", MS Thesis, EEAP Dept., Case Western Reserve Univ., Cleveland, Ohio, USA, (1994).

5. Wen H. Ko, "Capacitive absolute pressure sensor for industrial applications" patent pending, filed 1995

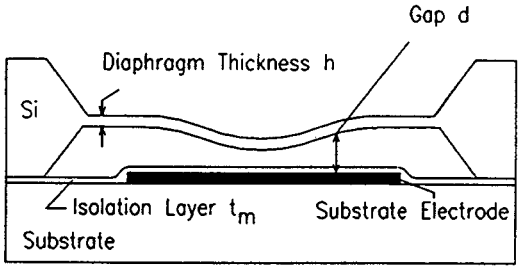
6. W.H.Ko and G.J.Yeh, "An integrated interface circuit for capacitive sensors", *Microsystem Technology*, (Springer International) Vol.1, No.1, Oct. 1994, pp.42-47.

Table 1. TMCPS Process Flow Chart

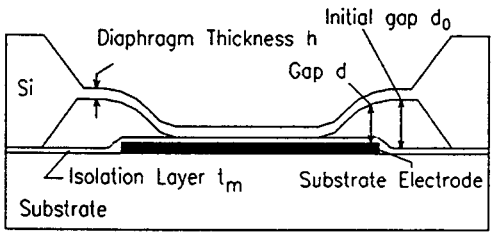
Step	Name	Description
Silicon Process		
1	Oxidation	Generate masking layers
2	Photolithography I	Define Gap area
3	RIE Silicon etch	Form gap
4	Boron diffusion	Define the thickness of diaphragm
5	Photolithography II	Define the etch back area
Substrate (glass) Process		
6	Electrode I	Form substrate electrode by sputtering Cr/Pt/Cr
7	Isolation	Sputter 7740 glass over the substrate electrode
8	Electrode II	Form contact pad for the diaphragm
9	Pad	open windows for wire bonding pads
Final Process		
10	Anodic bonding	Bond silicon wafer to the substrate wafer
11	Silicon etch	Form diaphragm
12	Dicing	Dicing sensor chip

Table 2. Touch Mode Capacitive Pressure Sensor Parameters

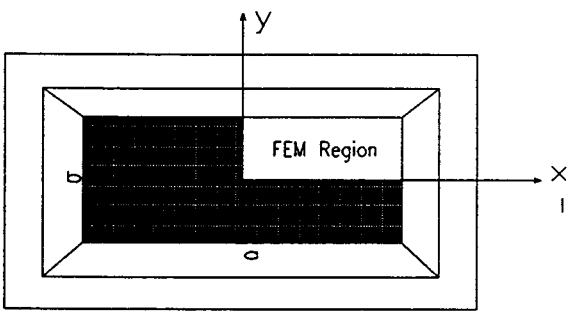
Parameter	Symbol	Value	Unit
Diaphragm Length	a	1500	μm
Diaphragm Width	b	457	μm
Diaphragm Thickness	h	5.0-8.0	μm
Gap	g	5.0-8.0	μm
Insulator Layer Thickness	t_m	0.01-3.0	μm
Young's Modulus of Si (P^+)	E	1.3E+11	Pa
Poison's Ratio	ν	0.3	



(a) Normal mode of operation



(b) Touch mode of operation



(c)

Figure 1 .Typical capacitive pressure sensor, structure a,b - dimension of the diaphragm; d - gap, h - thickness;

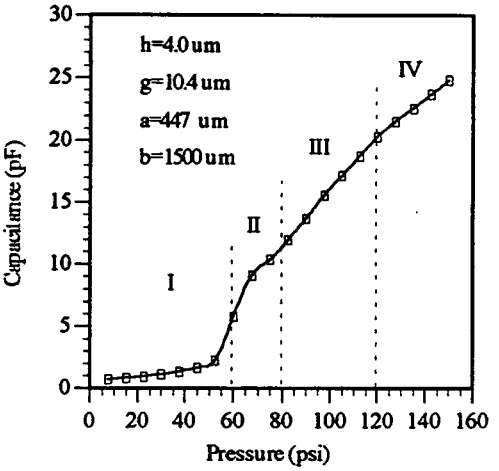


Figure 2. Typical C-P Characteristics of a TMCPS
I normal region ; II transition region;
III linear region; IV Saturation region.

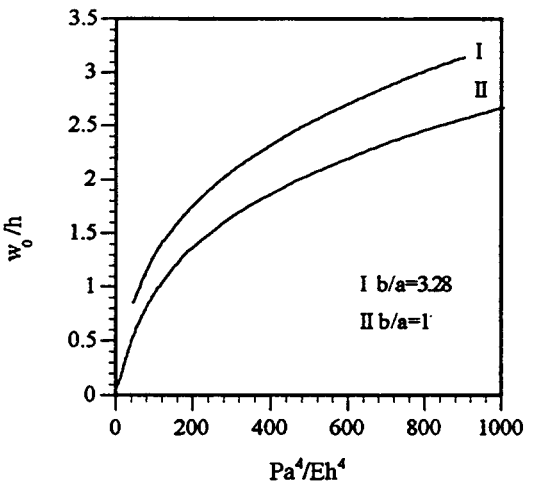


Figure 3. Dimensionless Center Point Deflection

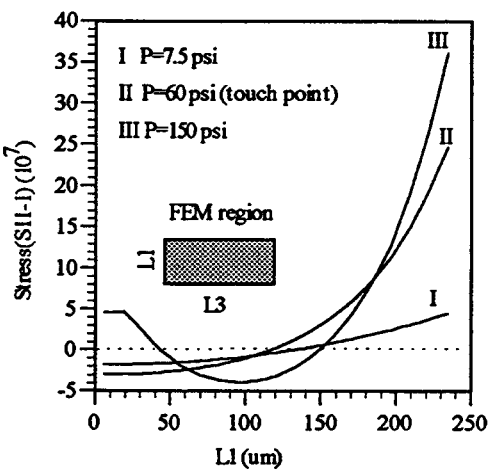


Figure 4. Stress versus coordinate L1 Touch point P_t=60 psi
a=457 μm b=1500 μm h=6.0 μm g=7.95 μm

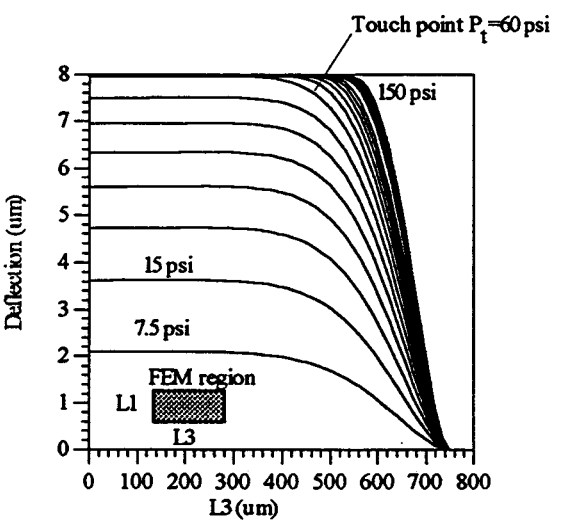


Figure 5. Deflection with different pressures along coordinate L3
a=447 μm b=1500 μm gap=7.95 μm h=6.0 μm

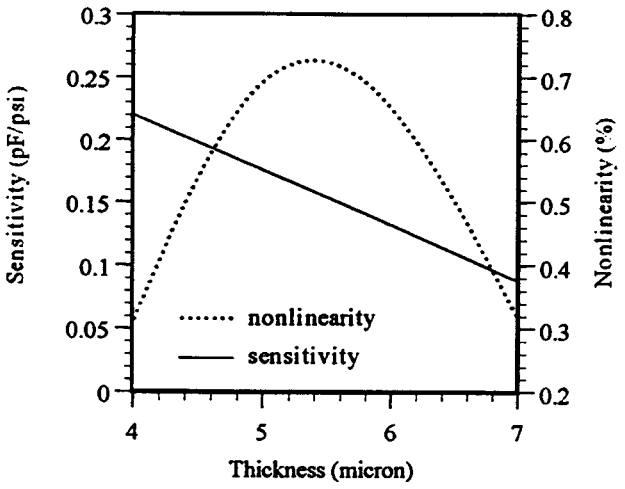


Figure 6. Sensitivity and nonlinearity versus diaphragm thickness

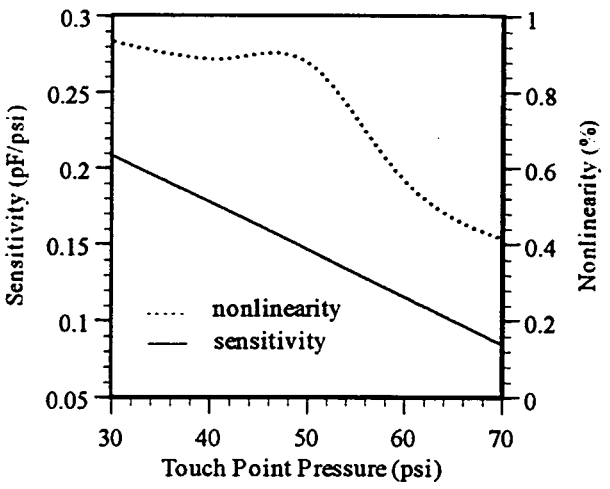


Figure 7. Sensitivity and nonlinearity versus touch point P_t

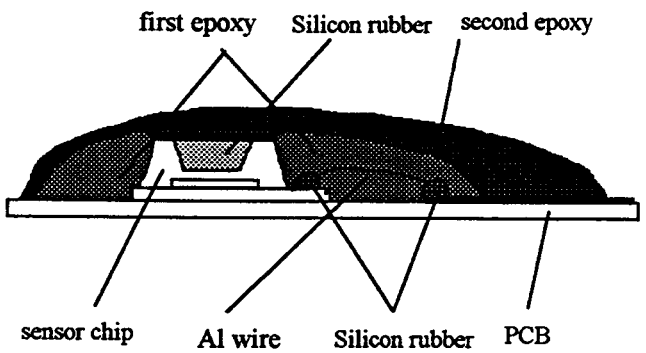


Figure 8. Package of TMCPS for the industry application

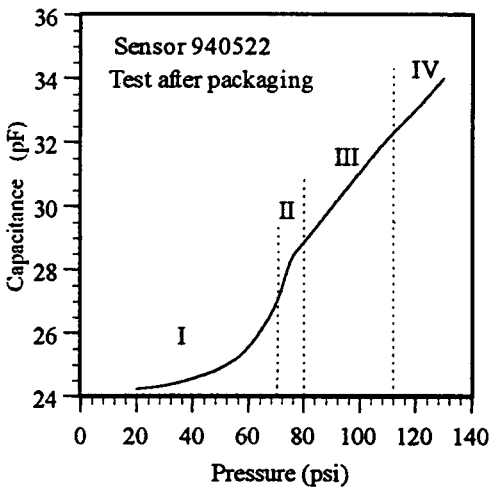


Figure 9. Typical C-P characteristics of a TMCPS
I normal region II transition region
III Linear region IV saturation region

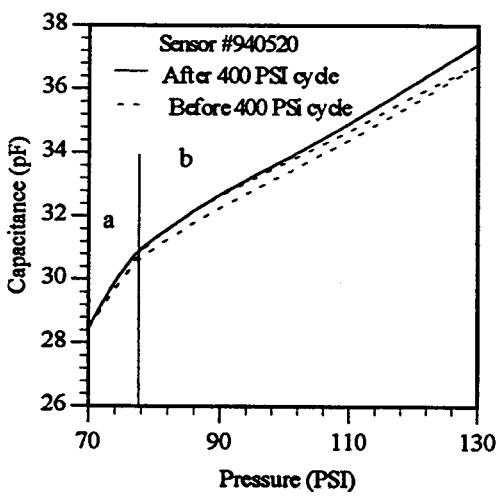


Figure 10 Hysteresis in the touch mode operation, region (a) - normal mode; region (b) - touch mode

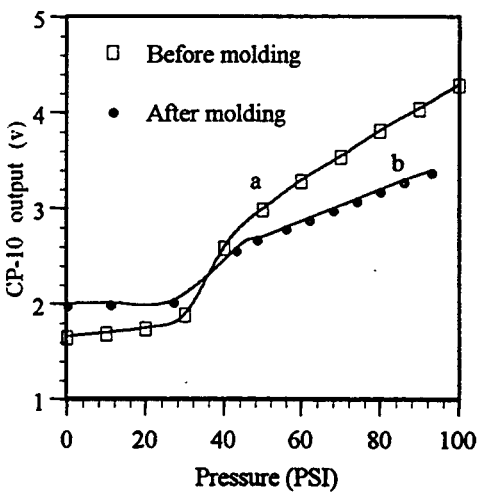


Figure 11. Field test results of sensor #9210084 molded in an industrial package. Curve (b) includes some pressure attenuation due to overall package.

INTEGRATED THERMAL CONDUCTIVITY VACUUM SENSOR

Erno H. Klaassen and Gregory T.A. Kovacs
Center for Integrated Systems, CIS-X 202
Stanford University, Stanford, CA 94305-4075

ABSTRACT

A thermal conductivity vacuum sensor with a novel on-chip control system is presented. The sensing structure is a small circuit suspended over a cavity etched in the substrate. An integrated control system maintains a constant temperature difference between this thermally isolated sensor and the substrate. The control loop allows for a digitally adjustable operating temperature of the sensing structure. The range of the vacuum sensor extends from 0.8 Pa (6 mTorr) to approximately 9.2×10^4 Pa (690 Torr). The system was fabricated in a completely unmodified foundry CMOS process, dissipates only 5 mW of power, and occupies an area of approximately 0.3 mm^2.

INTRODUCTION

Thermal conductivity gauges include the Pirani, thermocouple, and thermistor gauges. These devices measure pressure indirectly by sensing changes in the thermal conductivity of the ambient gases [1]. The thermal conductivity of a gas can be expressed as

$$k = \frac{1}{3} m \lambda n v_{ave} c_v \qquad [1]$$

where m is the molecular mass, λ is the mean-free-path, n is the molecular density, v_{ave} is the average molecular velocity of the molecules, and c_v is the specific heat at constant volume of the gas. The thermal conductivity of a gas, as a material property, does not change with pressure, since the pressure dependence of λ and n cancels. Thermal conductivity vacuum gauges therefore must limit the mean-free-path by some geometrical constraint.

Thermal vacuum sensors, like other thermally based instrumentation, can benefit from miniaturization. The reduction of parasitic thermal losses and thermal time constants can yield lower power consumption and faster response. These benefits, as well as the prospect of providing single-chip sensing systems have led to the development of a number of micromachined vacuum sensors over recent years [2,3,4,5,6,7]. Only some of these sensors have on-chip circuitry, and most are made with a custom fabrication process.

The fully integrated thermal vacuum sensor presented here is fabricated using post-processing methods on an unmodified foundry CMOS process [8]. The sensor measures the pressure-dependent thermal losses from a hot element, which consists of an area of single crystal silicon suspended by silicon dioxide support beams. The device is fabricated with an electrochemical etch in silicon-doped tetramethyl ammonium hydroxide (TMAH). Figure 1 shows a cross section of the undercutting etch that thermally isolates the circuitry. The details of the fabrication process can be found elsewhere [9,10]. Elements with high thermal isolation, yet with relatively small size, can be fabricated in this fashion. In addition, the suspended silicon can contain efficient heating and temperature sensing elements.

The suspended element is maintained at constant temperature relative to the substrate. The power necessary to do so is a function of thermal conduction through the support beams, but also has a convective component. When the mean free path of the molecules is dictated by the spacing between the silicon and the etched cavity, the convective component becomes a measure of air pressure, as shown in Eqn. 1.

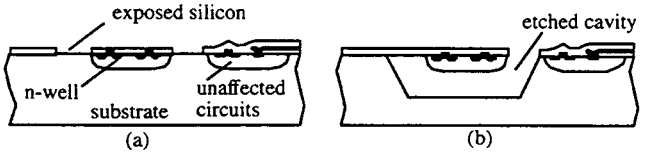


Figure 1. (a) Cross section of a vacuum sensor upon completion of the CMOS process. (b) Cross section of a completed vacuum sensor after the post-processing etch. The suspended silicon contains a PMOS heater and a diode temperature sensor.

SENSING ELEMENT DESIGN

An electrical analog of the thermal behavior of a thermally isolated sensor is shown in Figure 2. The thermal conductances are modeled as electrical conductors, heat dissipation is depicted by a current source, and voltages inside this circuit reflect temperature. The reference, indicated as an electrical ground, is the temperature of the silicon substrate around the sensor. The conductive losses are split into two components. The first is the unchanging conductive loss through the support beams of the element, which is dominated by the 0.5 µm thick aluminum traces in the beams. To minimize this parasitic element, the beams are made long, and the aluminum traces are made minimum width (3 µm). The other component is a pressure dependent thermal conductance through the air between the heated circuitry and the silicon substrate. In order to maximize the ratio of the two conductors, the heated element was placed close to the corner of the etched pit, so that the dominant conduction through the gas occurs between the silicon element and the (111) side-wall.

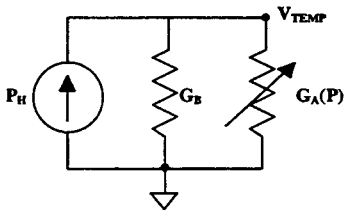


Figure 2. Electrical analog of the thermal behavior of the sensing element. The heat dissipated in the suspended circuit is modeled as a current source, while the temperature is denoted as the nodal temperature V_{TEMP}. The reference point in the circuit is the substrate temperature.

Figure 3 shows a top view of a vacuum sensing element. Devices were fabricated with beam lengths of 240 µm and 340 µm. The heated element requires a temperature sensor and a heater. In previous micromachined references, these functions were often performed with a single resistor [3,5,6,7]. In this technology, however, separate devices can be used for each of

0-9640024-1-8 $3.00 © 1996 TRF

these transduction operations between the electrical and thermal signal domains. A diode-connected bipolar transistor was used as the temperature sensor. When forward-biased with a constant current, the voltage across these devices has a temperature sensitivity of -2 mV/K. A polysilicon or diffused resistor could be used as a heater, but since the power dissipated in a resistor has a square-law dependence on voltage across or current through the device, special circuitry would be necessary to provide an output proportional to the conductive losses in the sensor. The on-chip feedback amplifier would also need to drive a resistive load. A better solution is to use a PMOS heating transistor. The power dissipated in such a device is proportional to source-drain current for a constant source-drain voltage and the amplifier only needs to drive the gate capacitance of the heater.

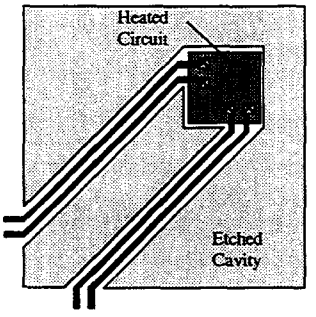


Figure 3. *Illustration showing a top view of a sensing element. The circuitry at the tip of the device is heated.*

SYSTEM DESIGN

Figure 4 shows a simplified schematic of the vacuum sensor system. The shaded devices are located on the sensing element. The output voltage between the two forward-biased diodes is proportional to the temperature difference between the sensor and the substrate. This temperature difference must be maintained at a constant value. The source-drain current through the PMOS transistor is then proportional to the power dissipated to maintain this temperature set-point.

The set-point could be implemented by adjusting the ratio of the current densities of the two diodes. This creates a voltage offset between the two diodes that is driven to zero by the feedback amplifier, thus generating a temperature difference between the sensor and the substrate. The voltage between two diodes (at equal temperatures) is given by

$$V_O = v_t \ln\left(\frac{I_{D1}}{I_{S1}}\right) - v_t \ln\left(\frac{I_{D2}}{I_{S2}}\right) = v_t \ln\left(\frac{n}{m}\right) \qquad [2]$$

where v_t is the thermal voltage (kT/q), I_D is the current through the diodes, I_S is the reverse saturation current, n is the ratio of I_{D1} to I_{D2}, and m is the junction area ratio of the two devices. The offset cannot be varied over a very wide range, due to the natural logarithm in the expression. Another problem is that the voltage is proportional to absolute temperature, which makes the operating point of the element a strong function of ambient temperature.

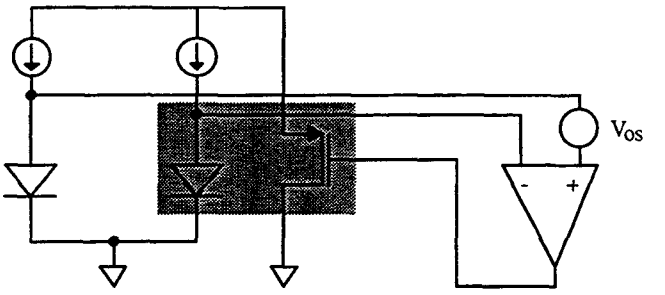


Figure 4. *Simplified schematic of the vacuum sensor system. The shaded area is the thermally isolated circuit. The output voltage between the two forward biased diodes is proportional to the temperature difference between the sensor and the substrate.*

A better way to set the operating temperature of the element is to introduce an offset inside the feedback amplifier. One way to do this is to effect a threshold voltage difference in the input devices of an operational amplifier. A topology that realizes this is shown in Figure 5.

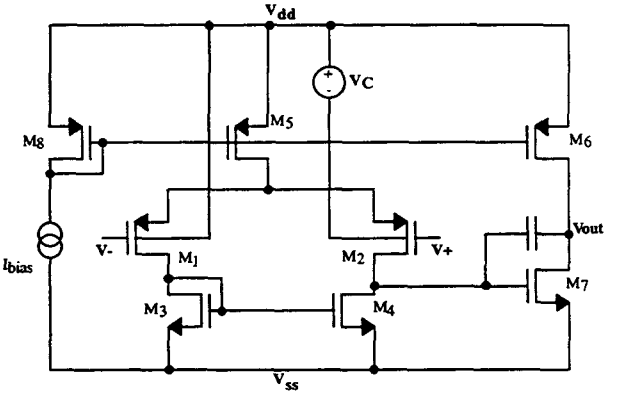


Figure 5. *Two-stage CMOS operational amplifier with adjustable offset.*

The n-well that contains the non-inverting input transistor M_2 is held at a voltage V_C below the well containing the other input transistor. The threshold voltage of an MOS transistor is modulated by the body effect, and for a PMOS device can be written as

$$V_T = V_{T0} - \gamma\left[\sqrt{|2\varphi_f| - V_{SB}} - \sqrt{|2\varphi_f|}\right] \qquad [3]$$

where V_{T0} is the threshold voltage without the body effect, $2\phi_f$ is the surface potential at inversion, V_{sb} is the source-to-bulk voltage, and γ is the body effect coefficient. The difference in offsets is

$$|\Delta V_T| = \gamma\left[\sqrt{|2\varphi_f| - V_{SB1}} - \sqrt{|2\varphi_f| - V_{SB2}}\right] \qquad [4]$$

With an offset introduced in this fashion, the common-mode rejection of the amplifier is degraded from one with a balanced input pair. The common-to-differential mode rejection ratio (CMRR), however, remains high if the differential gain is sufficiently large. This CMRR rejects the effects of common-mode changes in the temperature sensor output that result from ambient temperature changes.

The offset of the amplifier as a function of the offset control voltage V_C was measured and is plotted in Figure 6.

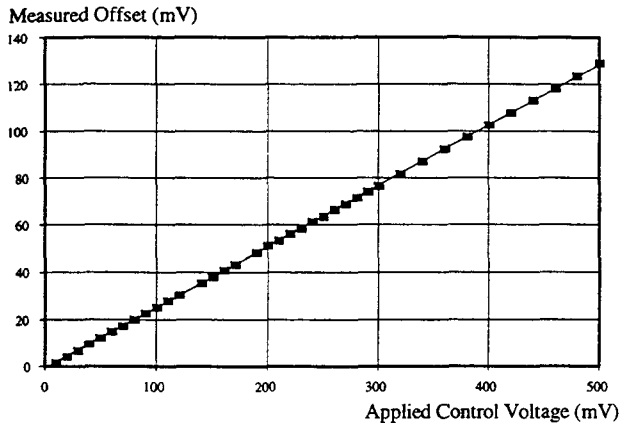


Figure 6. Measured input-referred offset of the amplifier in Figure 5 as a function of the control voltage V_C.

In order to generate the offset-control voltage, a digital-to-analog converter (DAC) was implemented on-chip. This allows for effective external control of the operating temperature of the sensing element by digital means, such as a microprocessor. The schematic of the DAC is shown in Figure 7. The DAC is based on a current-mirror consisting of bipolar transistors Q_1 and Q_2. These transistors are made by using a p-type base implant available in the CMOS process. The MOS devices at the emitters of the NPN transistors operate in the linear region, and act as an equivalent resistance. Kirchoff's voltage law around the loop containing the linear MOS transistors and the two bipolars yields (ignoring base currents)

$$v_t \ln\left(\frac{I_1}{I_2}\right) + I_1 R_A = I_2 R_B \qquad [5]$$

where I_1 and I_2 are the collector currents of Q_1 and Q_2, respectively. R_A is the effective resistance of linear MOS device M_6, and R_B is the source-drain resistance of the parallel combination of M_7 through M_{12}. With proper device sizing and biasing, the contribution of the natural logarithm term can be neglected, and the expression above simplifies to

$$I_2 \approx \frac{I_1 R_A}{R_B} \qquad [6]$$

The current ratio in the mirror is therefore determined by the ratio of the resistances at the emitter nodes. The resistance R_B is set by a digital code applied to the binary-weighted set of NMOS transistors M_7-M_{12}.

The entire vacuum sensor system is shown in Figure 8. The power dissipated in the PMOS heater is proportional to the current through this device, which is the current output of the vacuum sensor.

RESULTS

Figure 9 shows an SEM photograph of an etched vacuum sensor element. The beams on this sensor are 240 μm long. The temperature-sensing diode is located at the far corner of the element, and the PMOS heater (two separate devices) is located along the inner edge of the suspended silicon island.

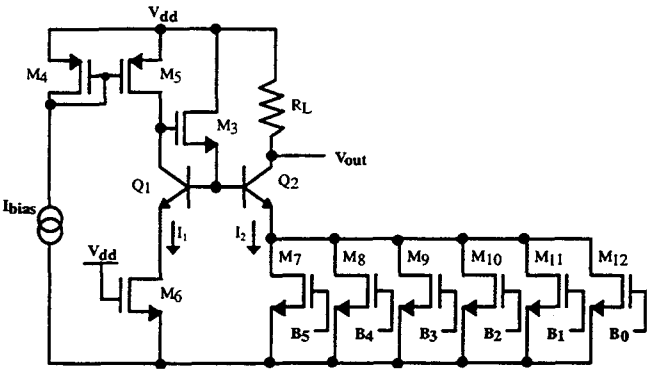


Figure 7. DAC that generates the offset control voltage for the feedback amplifier.

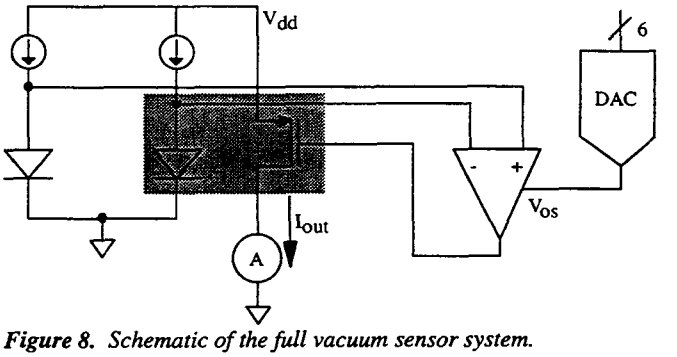


Figure 8. Schematic of the full vacuum sensor system.

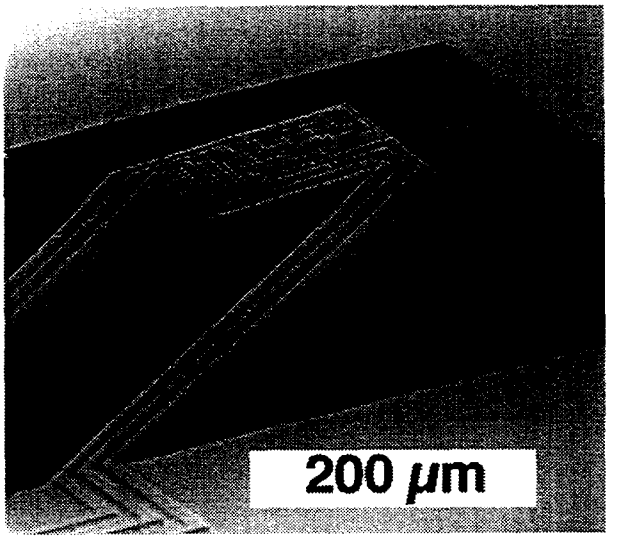


Figure 9. SEM of an etched vacuum sensor element.

Figure 10 shows the output of the sensor system for a device with 340 μm long beams, for different operating temperatures. The pressure was measured with an MKS Baratron, a bellows-type capacitive vacuum sensor. As the graph shows, the effective sensitivity is larger for higher operating temperatures, since all the points on the curve scale with the set-point. The output is the current (μA) through the PMOS heater for a fixed source-drain voltage of 3V. For this sensor, the ratio of the thermal resistance in air to that at low vacuum was approximately 5.9.

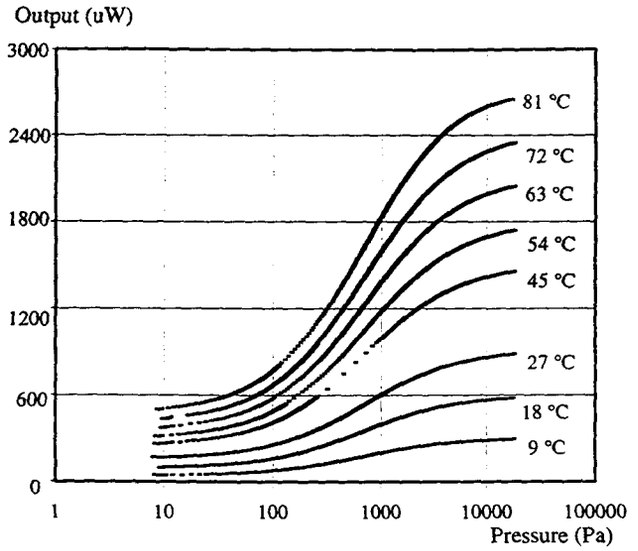


Figure 10. Sensor output over pressure for various operating temperatures relative to the substrate.

Figure 11 shows the performance of the sensor over its entire pressure range. The pressure is measured with an MKS Baratron below $1x10^4$ Pa and with a Hastings 760 Thermocouple gauge for pressures above this point. The output is plotted as

$$I_N = \frac{I(P) - I_{LP}}{I_{FS}} \qquad [7]$$

where $I(P)$ is the current at a given pressure, I_{LP} is the current at low pressure (0.1 mTorr), below the range of the sensor, and I_{FS} is the current at ambient [2]. This plot is for a set-point of 90 °C above the substrate temperature. The vacuum sensor has a range of 0.8 Pa (6 mTorr) to approximately $9.2x10^4$ Pa (690 Torr).

The entire system occupies less than 0.3 mm^2 and consumes approximately 5 mW of power (for a setpoint of 90 °C above the substrate temperature).

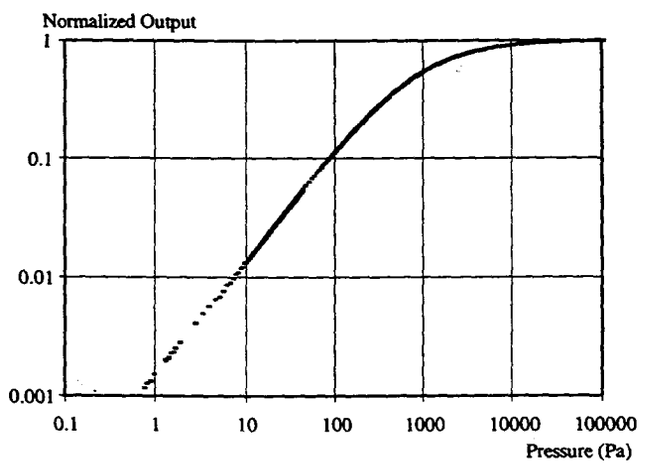


Figure 11. Plot of normalized sensor output over pressure.

CONCLUSIONS AND OUTLOOK

The fully integrated vacuum system presented here has a useful pressure range of more than five decades. It was built using an unmodified foundry CMOS process, which makes relatively inexpensive fabrication possible. The device could be further improved by providing some type of cover close to the suspended element [2]. This would increase the conductive losses through the air relative to that through the beams. Another possible improvement is to fabricate the device on wafers containing a buried oxide layer. This buried oxide layer would serve as an etch-stop during the undercutting etch, and would limit the depth of the anisotropically etched cavity.

REFERENCES

1. L. Carpenter, Vacuum Technology, An Introduction, American Elsevier, New York, 1970, pp. 49-55.

2. A. van Herwaarden, "Thermal vacuum sensors based on integrated silicon thermopiles," Ph.D. Thesis, Delft University of Technology, 1987.

3. C.H. Mastrangelo and R.S. Muller, "Fabrication and Performance of a Fully Integrated μ-Pirani Pressure Gauge with Digital Readout," In Proceedings of Transducers '91, San Francisco, pp. 245-248, 1991.

4. S.D. James, R.G. Johnson, and R.E. Higashi, "A Broad Range Absolute Pressure Microsensor" In Technical Digest, IEEE Solid State Sensor and Actuator Workshop, Hilton Head, pp. 107-108, 1988.

5. O. Paul, A. Haeberli, P. Malcovati, and H. Baltes, "Novel Integrated Thermal Pressure Gauge and Read-Out Circuit by CMOS IC Technology," In Technical Digest, IEEE International Electron Devices Meeting, San Francisco, pp. 131-134, 1994.

6. N.R. Swart and A. Nathan, "An Integrated CMOS Polysilicon Coil-Based Micro-Pirani Gauge with High Heat Transfer Efficiency," In Technical Digest, IEEE International Electron Devices Meeting, San Francisco, pp. 135-138, 1994.

7. U. Bonne and D. Kubisiak, "Burstproof, thermal pressure sensor for gases," in Technical Digest, Solid State Sensor and Actuator Workshop, held June 13-16, 1994, Hilton Head Island, SC, pp. 78-81.

8. Mosis 2.0 μm Analog CMOS process. 4676 Admiralty Way, Marina del Rey, CA 90292-6695.

9. R. J. Reay, E. H. Klaassen, G. T. A. Kovacs, "Thermally and Electrically Isolated Single Crystal Silicon Structures in CMOS Technology," IEEE Electron Device Letters, no. 15, pp. 399-401, 1994.

10. E. H. Klaassen, R. J. Reay, C. Storment, J. Audy, P. Henry, A. P. Brokaw, G. T. A. Kovacs, "Micromachined Thermally Isolated Circuits," in Technical Digest, Solid State Sensor and Actuator Workshop, 1996, Hilton Head Island, SC.

MICROMACHINING AND MICROGENETICS: WHAT ARE THEY AND WHERE DO THEY WORK TOGETHER?

Michael Albin, Reid Kowallis, Enrico Picozza, Yefim Raysberg, Chuck Sloan, Emily Winn-Deen, Tim Woudenberg and Jerry Zupfer
Applied Biosystems Division of Perkin Elmer Corp.,
850 Lincoln Centre Drive, Foster City, CA 94404.

ABSTRACT

PCR provides specific amplification of very small quantities of DNA by employing DNA polymerase extension under tightly-controlled temperature cycling. The reactions are typically performed in 10's to 100's of microliters in plastic tubes. The role of silicon-based processing in improving upon the PCR process (cost and performance) by reducing sample requirements, increasing throughput, and improving temperature cycling uniformity and ramping rates are contrasted with issues regarding surface chemistry and user interfaces at small scales. A platform for preparing test-arrays and cycling of multiple sites with direct monitoring of temperature AND reaction progress is demonstrated in both silicon and molded, plastic parts. The relevance to genetic screening is addressed.

INTRODUCTION

The polymerase chain reaction (PCR) [1] has demonstrated widespread utility in the area of genetic analysis. This simple process, which mimics the cell's natural ability to reproduce DNA, has been used as a core technology by many researchers to specifically amplify small target regions of DNA and to produce sufficient material for further analysis. Two primers, one for each DNA strand, are used to direct selective chain replication by DNA polymerase. As each chain is replicated it creates a copy of the original template. Thermal denaturation of these copies from the original template results in four strands which can be replicated in the next cycle. By employing successive cycles of high temperature thermal denaturation and lower temperature primer annealing and chain extension, exponential amplification of the target sequence bracketed by the two primers is achieved. The use of thermal stable DNA polymerases has facilitated PCR automation in a programmable heating block or thermal cycler.

The early years of PCR (1985-1993) have focused on low throughput usage in a vast array of research projects. The current trend is toward an increasing number of applied markets using PCR for routine analysis of specific DNA targets. As research work evolves into screening programs, sample throughput must increase dramatically. Rapid statistical data analysis and a reduction in cost per test are also needed. This need must be met by simple, low-cost instrumentation and systems integration from specimen preparation through amplification, product detection and data interpretation. The primary segments of the market where such simple, high throughput PCR systems are expected to dominate are drug efficacy monitoring, infectious disease testing in humans, animals, plants, food and the environment, human genetic disease testing, biodiversity monitoring, and DNA identification. With conventional systems the cost per test, calculated on a relatively high volume reaction (25-50 µL), gives a fairly high, average value of $1 per reaction.

Recent improvements in PCR amplification include development of special thin walled amplification tubes and glass capillaries which facilitate thermal exchange between the cycling block and the reagents, greatly reducing total PCR cycle time, introduction of a heated lid which eliminates the need to cover the reaction with oil in order to prevent evaporation during thermal cycling, development of homogeneous hot start techniques, and the ability to run PCR *in situ* on tissue sections mounted on microscope slides.

After completion of the PCR reaction the products must still be detected. In most early work, products from amplification reactions were evaluated by simple gel electrophoresis using gels stained with an intercalating dye such as ethidium bromide. Later they were quantitated by gel electrophoresis on a fluorescent DNA sequencer. Other methodologies have been developed to enable colorometric, multiplex detection capability. Higher density multiplex detection will be possible with some of the hybridization array chips under development now by companies such as Affymetrix [2] and Nanogen [3]. These arrays contain many capture probes and are designed to provide a yes/no answer as to whether a particular sequence is present in the target DNA. Both depend on performing PCR as an upstream step prior to detection of the PCR product by hybridization.

Many future PCR applications will require both real time detection and self-contained devices. Tubes may disappear and be replaced with miniature well arrays on a silicon chip (a sort of mini-microtiter plate). Several groups have demonstrated the feasibility of performing PCR in a chamber made of silicon [4,5]. Small volume PCR (5-10 uL) has been carried out in these chips using a computer controlled Peltier heater/cooler. The next step in miniaturization for PCR will include integration of PCR and amplification product detection. Here we report the first integrated PCR thermal cycling and detection system in chip format.

EXPERIMENTAL DETAILS

The system concept for high-throughput, genetic screening based-upon real-time, quantitative detection for PCR is depicted in Figure 1. A translation stage holds a slide

0-9640024-1-8 $3.00 © 1996 TRF

Solid-State Sensor and Actuator Workshop
Hilton Head, South Carolina, June 2-6, 1996

that contains individually addressable, cycling regions (for each chip) that are imaged at the end of each cycle. The detector consists of a tungsten bulb illumination source, a CCD camera and appropriate focusing/collection optics. The following sections detail aspects of the system and its performance.

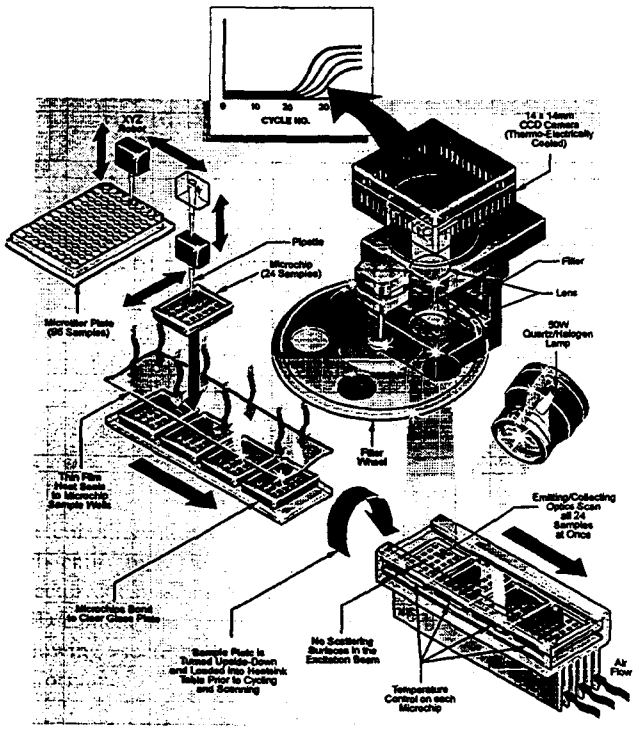


Figure 1. Schematic of PCR system: chips anodically bonded to a glass surface are loaded robotically; sealed with a polymer membrane; flipped for optical detection through the glass; thermal-cycled with on-chip resistive heaters or external peltiers; and imaged as cycling proceeds.

Chip Design:

Figure 2 shows designs of silicon test chips for PCR. The top row shows top/bottom views of a 14mm x 14mm structure consisting of eight 5 μL wells with a 4000Å thermal oxide coating. The larger entry holes (0.5 mm x 0.5 mm) allow for manual loading of individual wells. They are offset to minimize contamination as reagents are loaded manually. The bottom row illustrates the well layout and two designs in which resistive traces (aluminum) are deposited to provide a direct means to 'power' the devices.

Proportional, Integral, Differential (PID) control over sample temperature is established by one of two alternative means (on-chip and off-chip methods). The off-chip method uses a Marlow Industries model SE5020 temperature controller to drive a ITI Ferro Tec model 6300 peltier device. A maximum drive current of ±5amps enables the device to be ramped at ±4°C/s (average rate over a 55°C-95°C ramp). This design relies on a platen of uniform temperature which is very close to the sample temperature when at steady state. A 0.5°C uniformity for the chip can be realized in this format. The drawbacks are that the ramping is slower than can be realized with on-chip heaters, and that there is an issue of longevity of the thermal electric scheme which is currently employed.

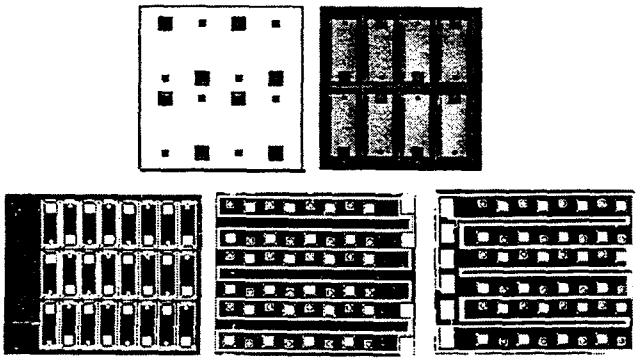


Figure 2: Layouts of chips for PCR testing - Top: 8 - 5μL wells, Bottom: 24 - 2μL wells.

The on-chip method uses one or two zones of resistance heating which can establish a maximum power dissipation of 25 Watts over the 200 mm^2 area chip. This results in a measured average temperature ramp rate of 10°C/sec over the transistion from 55°C to 95°C. Most of the time during this transition is spent closing out the ramp in a manner which does not allow more than 0.5°C of overshoot. These designs work by continuously applying a thermal loss to the silicon chip, such that the 10°C/sec down rate is accomplished merely by turning off the heater power. Hence at 95°C the heat flux through the chip is large. Unless the sourcing and sinking of heat is quite uniform, there will be significant non-uniformity in the steady state temperature profile. The temperature uniformity at different stages were assessed with InfraRed thermography (Fig. 3). The images depict system performance in actual use configuration at 95°C. The results indicate a steady-state difference in temperature of 5.2°C from the hottest to coldest point in the left design (single element) and 2.0°C in the subsequent modification. This is still not sufficiently uniform for precision PCR work.

Figure 3. Thermal images of two different on-chip heater designs as shown in Figure 2 (InfraMetrics Model 740 Infra Red camera configured with a 77K HgCdTe detector).

Chip loading:

The intent for an initial system would be to put the burden on the manufacturer in terms of pre-loading the microstructures so that the user would merely need to introduce the sample(s). A dispensing robot (Asymtek, Automove 402) coupled to drop-on-demand, ink-jet printing heads provide a tool for precisely delivering fluids to the inlets of the small structures (Figure 4). The system specifications are as follows:

Robot: X, Y-axis - work table, 12"x12"; resolution: 001"; velocity: 0-20 inches/sec.; repeatability, +/- 0.001". Z-axis - travel, 3.5"; resolution, 0.001"; velocity, 0-8 inches/second.

Options: tip locator, offset camera, strobe drop camera, on-axis camera, gravimetric drop calibration.

Printing Head: Type - drop-on-demand, Piezo-Electric; Wetted Surfaces - Glass, Teflon, Polypropylene; Minimum Drop Volume - 125 pL; Maximum Flow: 1 µl/min; Max Duty Cycle: 66%.

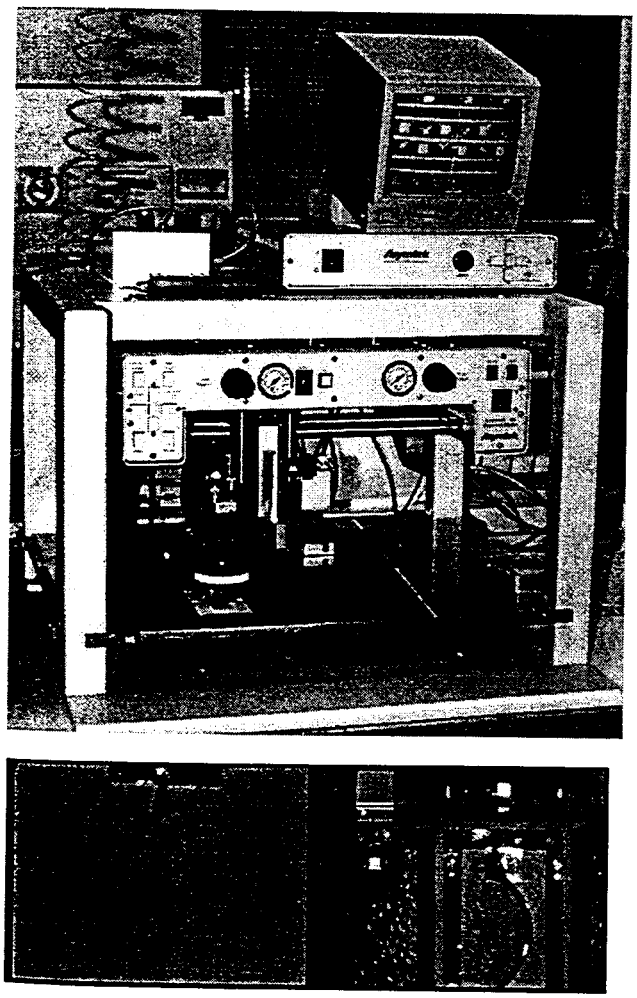


Figure 4. Chip loading configuration: Top - Robotic work station - screen image displays portion of slide with 0.5mm x 0.5mm entry ports for fluid loading; wells = 2 µL, slides are located beneath CCD camera at the system base. Bottom - droplet formation at tip; fluid at entry port entering well.

Sample loading:

One concept for a user interface to reagent-loaded arrays is to interconnect a number of wells that would serve to address one sample for multiple test screening. Such a structure is shown in Figure 5. In this configuration, 49-1µL wells are connected by a serpentine channel. The device is evacuable, thereby allowing the user to input a sample which can then be transferred to all connecting wells.

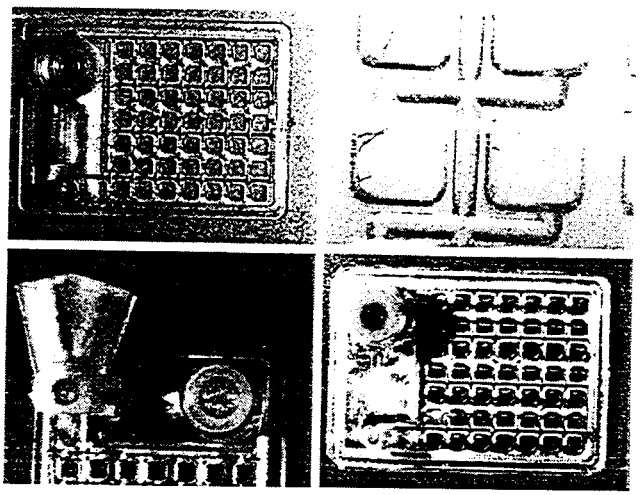


Figure 5. Multi-well plate and injection manifold: Top - Device and enlargement of interconnecting channel; Bottom - vacuum, injection ports. Wells are 1.2 x 1.2 x 0.75 mm deep (1.1 µL); channels are 250µm wide x 250 µm deep.

In this format, the process steps consist of:

- Filling the devices with reagents by the manufacturer
- applying a vacuum to the device
- Drawing sample into the evacuated device
- Sealing the access ports prior to thermocycling

Amplification and Detection:

Some of the key issues in performing DNA amplification in small, silicon devices include materials compatibility (does the chemistry work the same in silicon as in polypropylene?), geometry (how does the larger surface/volume area effect performance?), and thermal characteristics (does the higher thermal conductivity permit faster, more robust chemistry?). Materials properties of substrates evaluated for PCR are given in Table I.

In general, PCR in silicon devices has worked best using 4000-5000Å of thermal oxide grown on the silicon surface; however, this does not provide a general equivalence to polypropylene tube performance. Studies of the interaction of various species present in the PCR reaction mix indicate that enzyme binding to the surface is a critical aspect of the differing performance. Fig. 6 illustrates the results of analysis (ESCA, FISONS 22i XL) of Taq polymerase

binding to the surface of silicon with a thermal oxide coating versus untreated polypropylene.

Table I. Materials Properties Relevant to PCR

MATERIAL	TH. COND. (W/m/°K)	MAX USE (°C)	ELEC. RES. (ohm/cm)	PCR OK (bare/treated)
AIR	0.01			
GLASS	0.76	260	5 E 07	NO/YES
POLY CARB	0.20	126	8 E 16	YES/YES
POLY PROP	0.33	110	>1 E 17	YES/YES
SILICON	150.00	1410	4 E 06	POOR/YES
ALUMINUM	204.00	650	4 E -6	NO/YES
COPPER	391.00	1080	2 E -6	NO/???
PCR BUFFER	0.65			

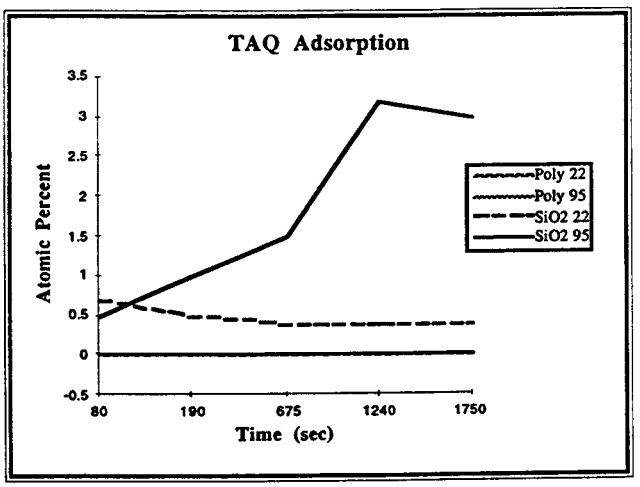


Figure 6. Analysis of enzyme interaction with polypropylene and silicon dioxide surfaces. Inset indicates material and temperature. Atomic % represents the number of atoms of an element in a volume of 100 atoms near the surface.

Figure 7 illustrates a method that gives a signal to allow real-time homogenous detection of specific PCR products [6,7]. This technique makes use of the 5' -3' endonuclease activity of AmpliTaq® DNA polymerase to release fluorescence from a reporter probe during PCR. The reporter probe is doubly labeled with a fluorescent reporter dye and a fluorescence quencher dye. If the target of interest is present, the probe specifically anneals to the target DNA between the forward and reverse primer sites. During primer extension the DNA polymerase allows the enzyme to both extend the primer and cleave the probe thereby releasing the reporter dye from the quencher. This fluorogenic process can be followed in real time, allowing specific detection and quantitation of up to 3 sequences simultaneously over a dynamic range of 5 orders of magnitude in starting target copy number.

The results of an automated run on a system as depicted in Figure 1 are shown in Figure 8. Positive samples and negative controls were run on distinct devices as cross-talk between silicon wells is a critical issue during manual sample loading in this configuration. The top figure shows real-time images of 8-5µL wells at different time cycles. The 582 nm

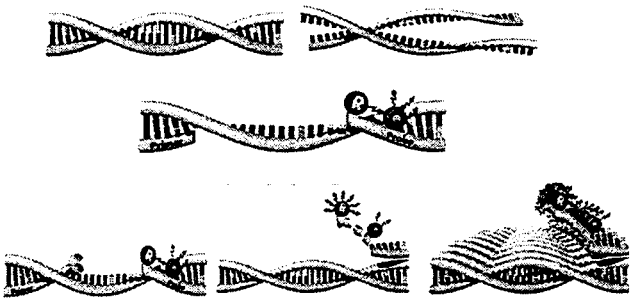


Figure 7. Illustration of PCR reaction and TaqMan™ format to generate real-time fluorescence signal: Top - denature DNA (95°C), Center: Anneal Primer and Probe (72°C); Bottom - polymerize --> fluorescence increases, repeat x-times.

images represent reference wavelength data; 518 nm represents the fluorescent product. The middle panel is the analyzed data showing the exponential growth of amplified product. It is significant to note the offsets between the sets of traces: for a ten-fold change in starting copy number of DNA, an offset of 3-4 cycle numbers for a given intensity is observed as predicted. The last panel shows an off-line analysis of final products by capillary electrophoresis. The product (294 base pairs) appears at ~6 min. and shows minimal side products. Tube controls gave ~2-fold higher signals.

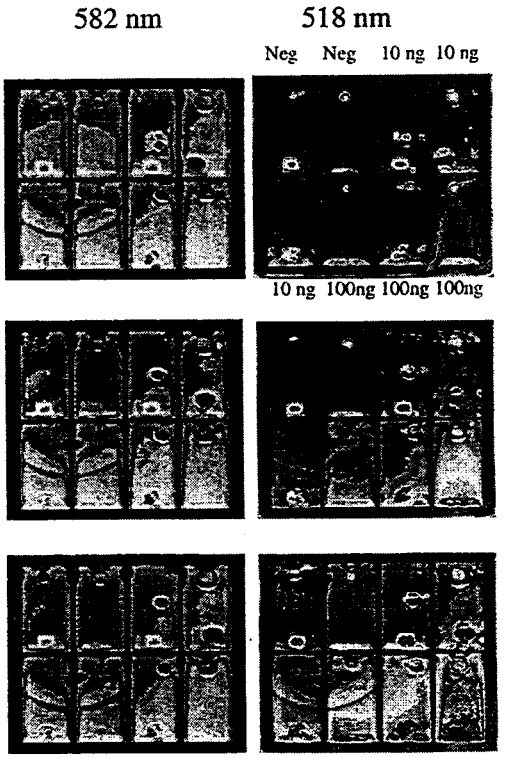


Figure 8a. Amplification of β-actin: 40 cycles (96°C, 30 sec; 54°C, 30 sec; 72°C, 2 min.) with peltier control. Real-time fluorescence images of PCR well arrays; Top = cycle 1, Middle = cycle 22, bottom = cycle 40.

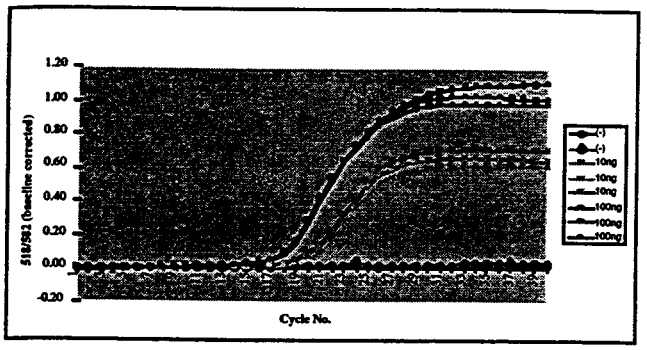


Figure 8b. Data analysis of Figure 8a.

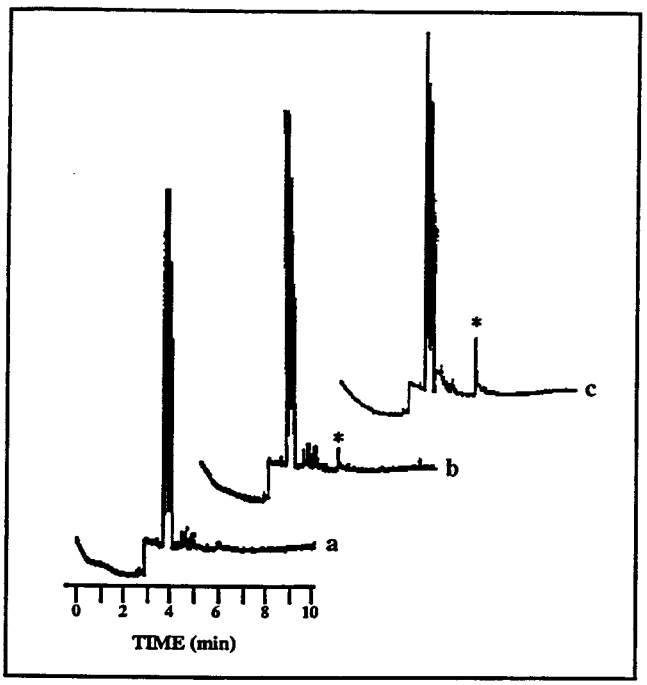


Figure 8. Off-line capillary electrophoresis analysis of PCR reaction products of Figure 8a.

CONCLUSIONS

A number of factors influence the formats in which genetic analysis is performed and the directions of current research. Speed and cost of analyses, simplicity (*i.e.* integration of analysis components), robustness and throughput are all key issues. In this work we have demonstrated that PCR can work efficiently in silicon-based devices; however, to be cost effective additional efforts in materials/process development must occur. Additional efforts will also focus on incorporation of sample preparation aspects and alternative loading and detection formats (such as capillary electrophoresis shown above) **[8-10]**. Much of this progress will rely on the development of fluid transfer and device interconnect methodologies.

REFERENCES

1. R.K. Saiki, S. Scharf, F. Faloona, K.B. Mullis, G.T. Horn,, H.A. Erlich, and N. Arnheim, "Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia" *Science* 230, 1350-1354 (1985).

2. R.J.Lipshutz, D.Morris, M.Chee, E.Hubbell, M.J.Kozal, N.Shah, N.Shen, R.Yang, and S.P.A.Fodor, "Using Oligonucleotide Probe Arrays to Assess Genetic Diversity", *BioTechniques*, 19, 442-447 (1995).

3. Cambridge Healthtech Institute, "The Human Genome Project: Commercial Implications", San Francisco, CA, 3/4-6/96 (1996).

4. M. A. Northrup, "DNA Amplification with a Microfabricated Reaction Chamber" Transducers '93, 924-926, (1993).

5. P. Wilding, M.A. Shoffner, and L.J. Kricka, "PCR in a Silicon Microstructure" *Clin. Chem., 40,* 1815-1818, (1994).

6. P.M. Holland, R.D. Abramson, R. Watson, and D.H. Gelfand, "Detection of Specific Polymerase Chain Reaction Product by Utilizing the 5' to 3' Exonuclease Activity *of Thermus aquaticus* DNA Polymerase", *Proc. Natl. Acad. Sci.*, 88, 7276-7280 (1991).

7. L.G. Lee, C.R. Connell, and W. Bloch, "Allelic Discrimination by Nick-translational PCR with Fluorogenic Probes", *Nuc. Acids Research*, 93, 3761-3766 (1993).

8. S.C. Jacobson and J.M. Ramsey, "Integrated Microdevice for DNA Restriction Fragment Analysis", *Anal. Chem.*, 68, 720-723 (1996).

9. "Micro Total Analysis Systems", A.van den Berg and P. Bergveld (eds.), Kluwer Academic Publishers, The Netherlands (1995).

10. D.J. Harrison, K. Fluri, K. Seiler, Z. Fan, C. effenhauser and A. Manz, "Micromachining a Miniaturized Capillary Electrophoresis-Based Chemical Analysis System on a Chip" Science, 261, 895-897, 1993.

ACKNOWLEDGMENTS

This research was supported by The Department of Commerce, NIST-ATP. The following sites contributed technical resources: EG&G IC Sensors (silicon parts), MicroFab Technologies (ink-jet heads) and SAIC (surface analysis).

Miniaturized Genetic-Analysis System

Rolfe C. Anderson, Gregory J. Bogdan, and Robert J. Lipshutz
Affymetrix, Inc., 3380 Central Expressway, Santa Clara, CA 95051

ABSTRACT

We describe a miniature cartridge for carrying out all required DNA target-preparation reactions for hybridization-based sequence analysis in 10 µl volumes. Using this cartridge we demonstrate a hybridization-based HIV-1 mutation assay, including the polymerase chain reaction (PCR) [1], in-vitro transcription (IVT), labeling, fragmentation, hybridization, scanning and data analysis with performance equivalent to benchtop assays.

INTRODUCTION

Chemical and biological assays employing a series of operations such as mixing, reacting, separating, and sensing are critical in numerous fields including clinical diagnostics, biomedical research, analytical chemistry, industrial process control and environmental monitoring. Existing methods are based on macroscopic, bench scale reactions and operator intervention and are not amenable to reducing sample and reagent volumes, labor, processing time, and cost or improving reliability and portability. Miniaturizing, integrating and automating these assays will achieve these goals and result in broader applicability. Many investigators have reported miniaturizing analytical systems [2]. Single reactions such as DNA amplification by PCR in microfabricated chambers [3] and analysis steps such as capillary electrophoresis on glass substrates have been demonstrated [4] [5]. Systems in which mixing is followed by separation have also been demonstrated [6].

The elucidation of the DNA sequence of all human genes is creating new opportunities in genetic analysis with applications in clinical diagnostics, biomedical research, and pharmaceutical discovery. Genetic analysis typically requires nucleic acid extraction, amplification, labeling, and sensing. These complex assays are clear candidates for miniaturization and integration.

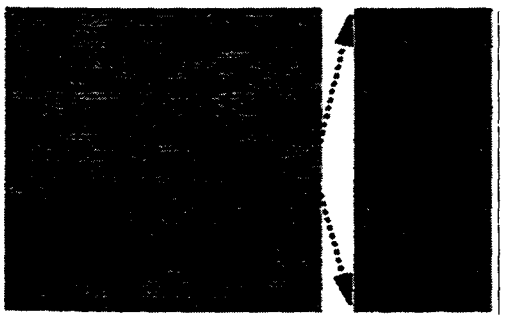


Figure 1. Fluorescence image of microfabricated 1.28 x 1.28 cm HIV-1 DNA-array chip after hybridization with test RNA sample. The array contains over 20,000 88 by 91 micron features.

Hybridization-based genetic analysis using high-density DNA probe arrays has been demonstrated [7] [8] [9] [10]. These microfabricated DNA arrays may contain more than one million different DNA probe regions and have sufficient complexity to analyze tens of thousands of nucleic acid bases. These arrays have been used to detect mutations associated with cystic fibrosis and hereditary cancers in human samples, and mutations in HIV putatively associated with drug resistance. Data from a DNA probe-array assay is shown in Figure 1.

The hybridization-based HIV-1 mutation assay is described in Figure 2. Currently, these manual steps are individually carried out at a laboratory bench in 20 to 1000 µL volumes. Any program to miniaturize and automate this assay must address biocompatability between reagents and the chamber walls, thermal isolation of sensitive reagents from the high temperature reaction chambers, and the quantitative transfer of fluids between chambers. This preliminary report describes a miniature cartridge to automatically perform these steps in 10 µl volumes.

REACTION MINIATURIZATION

We are following a natural progression from reactors to systems, first individually demonstrating each reaction in miniaturized chambers, and then combining these reactions into a system. This progression allows us to address reactor biocompatability before fluid movement, fluid control and thermal isolation. Reactors connected with channels were fabricated from polycarbonate.

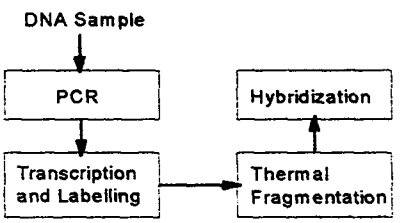


Figure 2. Process steps in the hybridization-based HIV-1 mutation assay. The temperature-cycled PCR reaction selectively amplifies the DNA region of the HIV-1 genome to be interrogated. T3 or T7 RNA Polymerase transcribes the DNA template yielding fluorescently labeled RNA. Thermal fragmentation randomly breaks the RNA into lengths appropriate for hybridization. The fragmented RNA is incubated with the DNA probe array so that fragments may bind (hybridize) with complimentary sequences on the array. After hybridization, the fragments that are not hybridized are washed away. The array is read by epifluorescence confocal scanning.

0-9640024-1-8 $3.00 © 1996 TRF

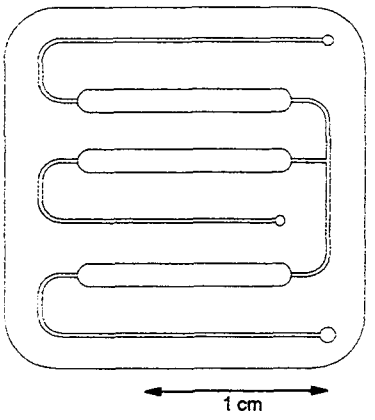


Figure 3. Top view of an array of polycarbonate reaction chambers. Dimensions are channel: 250 μm wide by 125 μm deep; reactors: 1.5 mm wide by 13 mm long by 125 to 500 μm deep, with reactor volumes from 2.5 to 10 μl.

Polycarbonate can be injection molded, ultrasonically welded, sterilized, manufactured at low-cost, and provides low thermal conductivity.

A top view of an array of polycarbonate reactors is shown in Figure 3. As shown in the cross section, Figure 4, the lower wall of the chamber is a thin sheet of polycarbonate. The reactor temperature is controlled by contacting a thermoelectric device with this thin wall (Figure 5).

PCR in a reaction chamber was performed using a standard biochemical and thermal cycling protocol [11]. A dedicated computer monitored and controlled the temperature as it was cycled between 55°C, 65°C and 94°C (Figure 6). Evaporation and bubble collapse were addressed by ensuring a clean reactor and bounding the reaction mixture with a hydrophobic

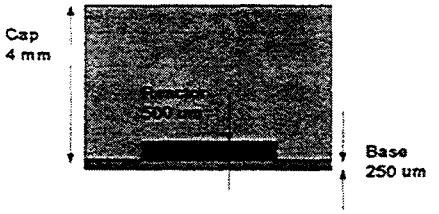


Figure 4. Cross section of a polycarbonate reactor.

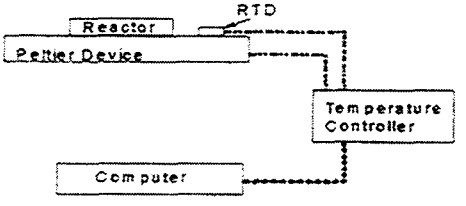


Figure 5. System for controlling reaction temperature with measurement using a resistive temperature device (RTD).

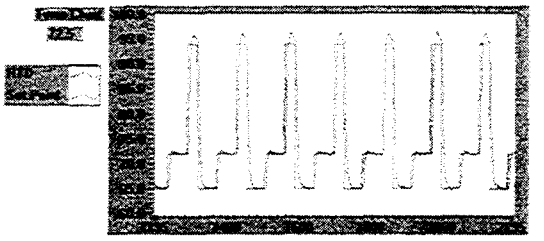


Figure 6. Labview-based controller for PCR reaction with target and actual temperature-time data for microreactor.

fluid. Thirty five thermal cycles resulted in billion-fold amplification in a 13μl sample, appropriate for subsequent reaction steps (Figure 7). The single narrow band on the gel and subsequent hybridization tests demonstrate excellent amplification fidelity.

Transcription and labeling were performed using a standard in-vitro transcription reaction to copy the DNA to RNA while incorporating bases with attached fluorescein labels [12]. Sensitivity to contamination by environmental RNA-destroying enzymes (RNAse), and denaturation of the transcription enzyme (RNA polymerase) were addressed by beginning with and maintaining sterile conditions, and slightly increasing the enzyme concentration. The reaction was carried out in the polymeric chamber in a 3μl volume with a 70% yield, compared to a control reaction (Figure 8).

The fragmentation reaction requires heating the labeled RNA to 94°C in the presence of Mg^{++} [13]. Temperature was controlled using the configuration shown for PCR. Evaporation was controlled by sealing the ends of the reaction chamber with a hydrophobic fluid. Purified labeled RNA was used to monitor fragmentation with agarose gel electrophoresis. Miniature and benchtop fragmentation gave equivalent

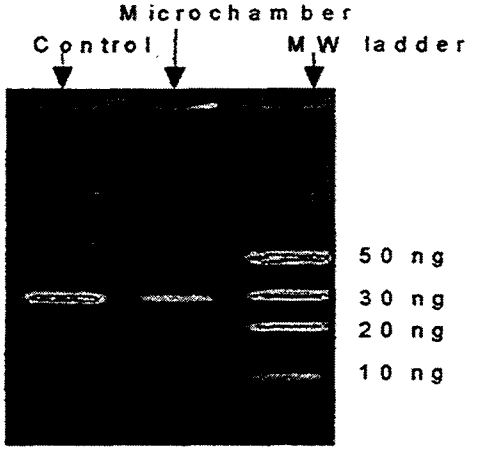


Figure 7. Agarose gel results for 35 cycle PCR amplification of 1100 base portion of HIV-1 from 2000 starting copies. Reaction in microchamber yielded 10^9 amplification, appropriate for subsequent reaction steps.

hybridization results.

SYSTEM INTEGRATION

Effectively combining the reactions described in Figure 2 into a single system requires fluid position control, liquid dosing, mixing, and thermal isolation between adjacent chambers. The RNA-polymerase enzyme would be denatured at the high temperature (94°C) used in the PCR and fragmentation reactions. Thermal isolation is accomplished by fabricating the reaction and storage chambers more than 10 mm apart in a thin polycarbonate reactor and by controlling the temperature in each region using thermoelectric devices.

In this assay the PCR product must be measured and mixed with an IVT cocktail stored in the same cartridge (Figure 9). The cartridge design and fluid-movement algorithm for this series of reactions are shown in Figures 3 and 10, respectively. Performing this series of reactions in the cartridge resulted in a 30% yield for the PCR reaction, and a 15% net yield through IVT, as compared to the control (Figure 11). This product concentration is appropriate for the subsequent hybridization reaction. In this case, less dilution of the fragmentation mixture prior to hybridization is required versus the standard benchtop protocol.

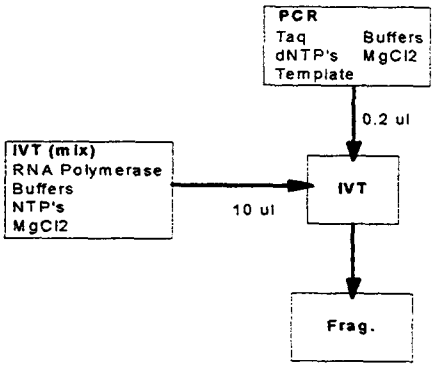


Figure 9. This HIV-1 target preparation protocol requires measurement and mixing of two liquid streams.

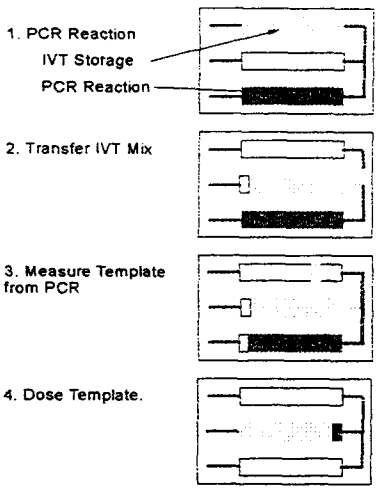


Figure 10. Series of fluid movements used to measure part of the PCR product and combine it with the IVT mixture. Oscillation at step 4 achieves mixing.

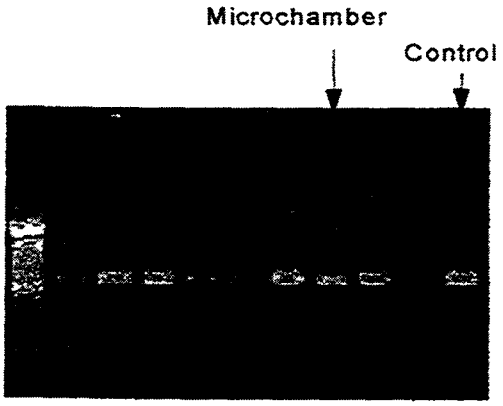


Figure 8. Agarose gel results indicating that the transcription reaction carried out in microchambers with a 3μl volume has 72% yield as compared to the control reaction performed in a conventional chamber.

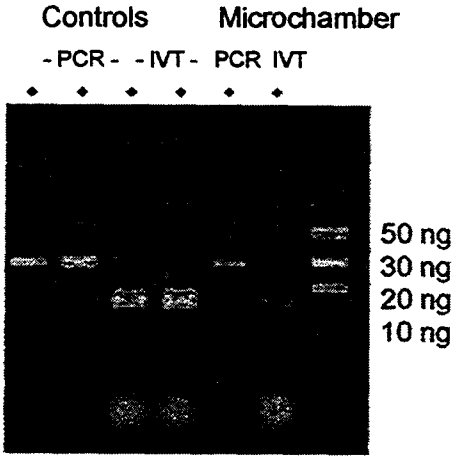


Figure 11. Agarose gel results for HIV-1 assay reaction sequence (Figure 2) demonstrated on a single device.

CONCLUSIONS

We have developed a genetic analysis system that integrates all of the preparative reactions necessary for a hybridization-based HIV-1 mutation assay with overall performance comparable to the bench-top protocol. Further optimization of this system will provide greater yields and improve sensitivity. This work comprises a crucial step towards a fully integrated genetic analysis system. We are following an evolutionary path, first developing disposable polymeric cassette-based reaction systems, and later, where appropriate, incorporating microfabricated elements. This approach has allowed us to quickly generate a working system and address materials compatibility issues in reaction volumes equal to or smaller than those published for silicon and glass devices.

The specific protocol we have demonstrated can be extended to other assays that employ hybridization-based sequence analysis by using different temperatures, other fluidic movement schemes, and other reagent combinations. Moreover, the components we have developed and the approach we describe can be applied to other genetic-analysis protocols and biochemical assays in general.

ACKNOWLEDGMENTS

The advice and assistance of scientists at Affymetrix were invaluable in performing these experiments. The authors would like to particularly thank Nila Shah, Maureen Cronin, Tom Gingeras, and Huu Trann for their patience, helpful advice, and for providing certain reagents. Richard Rava provided valuable scientific guidance and comments on this manuscript. This work was supported by NIST cooperative agreement 70NANB5H031.

REFERENCES

1. R.K. Saiki, et al., "Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia", *Science*, 230: 1350-1354 (1985).

2. *Proceedings of the μTAS '94 Workshop*, A. van den Berg and P. Bergveld, ed., Kluwer (1995).

3. M.A. Northrup, et al., *Technical Digesst of Transducers 1993, Seventh International Conference on Solid State Sensors and Actuators*, Yokohama, Japan, (1993).

4 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, 895-897 (1993).

5. A.T. Woolley and R.A. Mathies, "Ultra-high-speed DNA fragment separations using microfabricated capillary array electrophoresis chips", *Proc. Natl. Acad. Sci*, 91, pp. 11348-11352 (1994).

6. D.J. Harison, K. Fluiri, Z. Fan, and K. Seiler, "Integration of Analytical Systems and Incorporating Chemical Reactions and Electrophoretic Separation", *Proceedings of the μTAS '94 Workshop*, p.105 (1995).

7 S.P.A. Fodor, J.L. Read, M.C. Pirrung, L. Stryer, et al. "Light Directed, Spacially Addressable Parallel Chemical Synthesis", *Science*, 251, 4995, pp. 767-773 (1991).

8 S.P.A. Fodor, R.P. Rava, X.H.C. Huang, A.C. Pease, et al., "Multiplexed Biochemical Assays with Biological Chips", *Nature*, 364, 6347, pp. 555-556, (1993).

9 A.C. Pease, D. Solas, E.J. Sullivan, M.T. Cronin, et al.,"Light-Generated Oligonucleotide Arrays for Rapid DNA Sequence Analysis", *Proceedings of the National Academy of Sciences*, 91, 11, pp. 5022-5026 (1994).

10 R.J. Lipshutz, D. Morris, M. Chee, E. Hubbell, et al., "Using Oligonucleotide Probe Arrays to Access Genetic Diversity", *Biotechniques*, 19, 3, pp. 442-447 (1995).

11. Reaction mixture consists of 2000 molecules of template DNA, 0.3 units of TAQ polymerase, 1.5 mM $MgCl_2$, 0.2 mM dNTP's, 0.2 uM primers, and 1X Perkin Elmer PCR buffer. Thermal program includes: (1) an initial denature at 94°C for 60 seconds, (2) a denature at 94°C for 20 seconds, (3) an anneal at 65°C for 40 seconds, (4) an extend at 72°C for 50 seconds, (5) steps 2 through 4 repeated 35 times total, (6) a final extend at 72°C for 60 seconds.

12. Reaction mixture consists of 2.5 mM each of ATP, CTP, GTP, and 0.5 mM UTP, 0.25 mM Fluorescein-UTP, 8 mM MgCl2, 50 mM HEPES, 1X Promega Transcription Buffer, 10 mM DTT, 1 unit T3 RNA polymerase, 0.5 units RNAguard (Pharmacia). Thermal program consists of 37°C for 60 minutes.

13. Thermal program is 94°C for 60 minutes.

A HIGH-YIELD PROCESS FOR THREE-DIMENSIONAL MICROELECTRODE ARRAYS

Qing Bai and Kensall D. Wise

Center for Integrated Sensors and Circuits
Department of Electrical Engineering and Computer Science
The University of Michigan
Ann Arbor, MI 48109-2122

ABSTRACT

This paper presents a practical microassembly process for the realization of three-dimensional (3D) microelectrode arrays for recording and stimulation in the central nervous system (CNS). Several two-dimensional silicon probes are positioned in a surface platform and held parallel using micromachined spacers. Lead transfers between the probes and the platform are formed using gold beams on the probes that are attached to the platform using ultrasonic bonding with a wire-free wedge. The low-profile (150μm) outrigger design of the probes allows the bonding of fully-assembled high-density arrays. Arrays having 8x16-shanks on 200μm centers have been realized and used to record from cortical single units with interelectrode capacitances of less than 1fF, corresponding to on-chip crosstalk below 0.01 percent. Micromachined tools allow full array assembly in 30 minutes.

INTRODUCTION

One- and two-dimensional micromachined silicon probes now allow the realization of passive recording and stimulating arrays having sites lithographically controlled with spacings as small as a few microns on supporting shanks 10-15μm thick and less than 30μm wide [1]. Such probes are now being used successfully by many physiologists to explore new aspects of neural systems [2,3]. Probes having integrated on-chip site selection, low-noise recording amplifiers, and output multiplexers/buffers have allowed the monitoring of neural activity on as many as 8 parallel channels, selected from among 32 sites, over a single signal lead [4,5]. Similarly, active stimulating probes have been demonstrated that are capable of generating complex current waveforms (0 to ±127μA ±1μA) on as many as 8 parallel channels and supplying them to sites selected from 64 available on 1D or 2D probes [6]. Thus, the technology is in place to record or stimulate over many parallel channels selected from a still larger number of available sites and to do so using only a few external leads. Micromachined silicon ribbon cables [7] allow chronic signal transmission between the probes and a percutaneous plug with minimal tethering.

A key remaining element needed for detailed studies of neural networks and the implementation of neural prostheses to aid the handicapped is the ability to instrument three-dimensional volumes of tissue with site separations of 200-400μm while displacing no more than about 1% of the overall tissue volume. While some very novel 2D depth arrays have been reported [8] for cortical use, for fully 3D arrays the use of 2D silicon probes is attractive in order to take advantage of the considerable capabilities of these devices. The formation of such 3D arrays then requires the development of practical assembly techniques for submillimeter parts. Such techniques are also of interest for a variety of other sensor and micromachine applications.

In the creation of 3D arrays of 2D probes, the lead transfer structure between the probes and a surface platform holding the array is a major challenge. Since the probes and the platform are in perpendicular planes, the transfers must be made between orthogonal conductors. They must be robust, reliable, and formed with high-yield in a process that is reproducible in reasonable quantities and directly compatible with probe formation.

Three lead-transfer techniques have been explored. Hoogerwerf [9] employed a single subdural platform assembly, with nickel electroplated bridges between tabs on the probes and pads on the platform acting as lead transfers. Since electroplating was performed on a per-lead basis after assembly, the process was rather difficult to implement, especially for structures having large numbers of lead transfers. To allow higher throughput, an approach involving a two-platform structure (Figure 1) was next explored. In this design, each probe mounted in the subdural platform had its own ribbon cable, which was connected to a supradural platform using conventional ultrasonic bonding. Although easier to implement, the use of multiple cables complicated assembly and allowed appreciable tethering forces to be exerted on the implant as the brain moved within the cranial cavity. The number of probes and leads was also limited because of the space required for bonding at the second (supradural) platform.

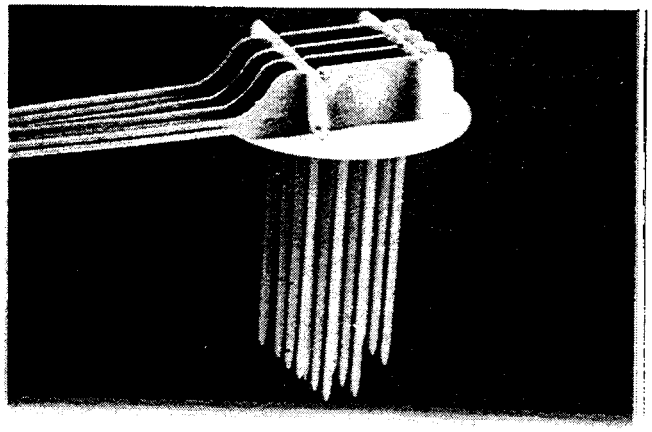
This paper describes a third approach, which uses gold electroplated beam leads on the probes to form the lead transfers. These gold tabs are bent at right angles to the probes and are attached to the platform using conventional thermocompression or ultrasonic bonding. The interconnects formed are robust physically and electrically.

ARRAY STRUCTURE AND FABRICATION

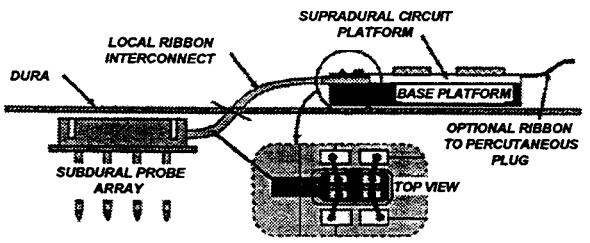
The overall structure of the 3D array is shown in Figure 2. It consists of a micromachined silicon platform, a number of 2D planar probes, and two spacer bars. The probes are inserted through slots in the platform and are held orthogonal to the platform by the spacers. All of these components are fabricated on the same wafer using only a slight modification of the basic probe process. The use of outriggers on the probes for the lead transfers permits the rim of the platform to be firmly supported in a simple jig during the lead bonding. The height of the flanges (outriggers) above the platform is about 150μm, which is much less than that of the central portion of the probe (1mm). This minimizes the effective thickness of the rim area and allows a conventional bonding tip to fit in the 200μm-wide space between the probes.

The gold beams [10] extend about 135μm below the flanges and are 100μm wide. The beams could be narrowed to 50μm to accommodate more leads if needed. They are typically electroplated to a thickness of 8-10μm, can be bent easily at right angles using a custom bending tool, and are attached to pads on the platform using ultrasonic bonding. Gold beam leads are also used at the end of the cable leading from the platform to

0-9640024-1-8 $3.00 © 1996 TRF



(a)



(b)

Figure 1: *An approach to 3D microassemblies using dedicated per-probe ribbon cables. (a) A 4x4-shank probe array implemented with this approach. (b) A schematic diagram of the approach.*

a percutaneous plug. These beams extend beyond the cable to bond to an insert in the base of the plug. Conventional wire bonds are then made between pads on this insert and pins on the plug. Since the cable mounts beam upside down on the plug insert, no twisting of the cable is necessary in this assembly, minimizing any tethering force on the tissue. The gold beams are fully compatible with micromachining, form high-quality physical and electrical contacts, and allow great flexibility in creating multilayer and 3D microstructures.

An 8-mask fabrication sequence is used to create the 2D probe components by defining the deep boron diffusion, shallow boron diffusion, polysilicon interconnects, contacts, iridium sites, field dielectrics, site protection, and gold beams.

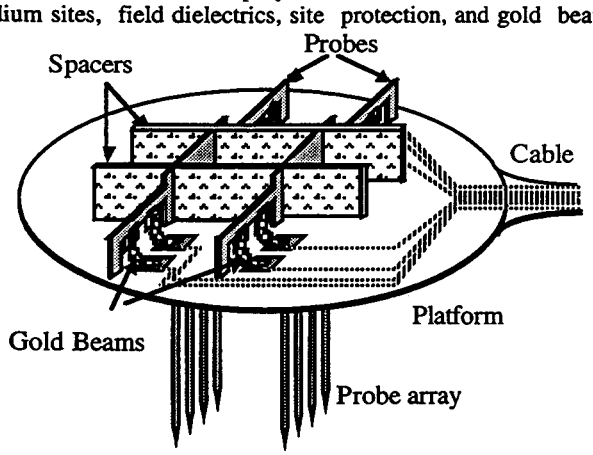


Figure 2: *A high-yield approach to 3D probe arrays. Lead transfers are implemented using beams that are bent at right angles and ultrasonically bonded to pads on the platform.*

The fabrication process starts with a deep boron diffusion, which forms an etch-stop to define the probe, the platform, and the spacers, setting their final thicknesses at 10-15μm. A shallow (2-3μm) boron diffusion is used next to form the silicon ribbon cable and the probe tips. Stress-compensated LPCVD dielectric layers (300nm/150nm/300nm $SiO_2/Si_3N_4/SiO_2$) are then deposited, followed by LPCVD polysilicon growth and phosphorus doping. Once the polysilicon interconnect is patterned using RIE, similar top dielectric layers are deposited. Contacts are defined using RIE, and Ti/Ir sites are then formed using sputtering and lift-off. The field dielectrics are now removed using RIE. This etch can be continued into the silicon to shape the probe sidewall if desired. A layer of photoresist is necessary to protect the Ir sites during electroplating. Cr/Au/Cr (50nm/ 300nm/50nm) is evaporated and used as blanket conductor during Au electroplating. A thick layer of photoresist is then used to define the beams and platform pads. The top Cr layer is removed to expose the Au seed layer for electroplating. After the gold beams/pads are plated, the photoresist is stripped, and the Au/Cr field layer is etched away. Finally, the wafer is thinned to approximately 100μm in a timed isotropic silicon etchant and is then given a final etch in EDP. After the components are sorted, they are ready to assemble into 3D probe arrays.

3D MICROASSEMBLY AND ASSOCIATED TOOLS

The procedure for assembling the 3D arrays is carried out in four steps: bending the gold beam leads on the outriggers of the probes, assembling the probe array, bonding the beams to the pads on the platform, and connecting the probe array to the percutaneous plug. The overall process is simple and reliable, especially when facilitated with the proper assembly tools. Complete array assembly takes about 30 minutes.

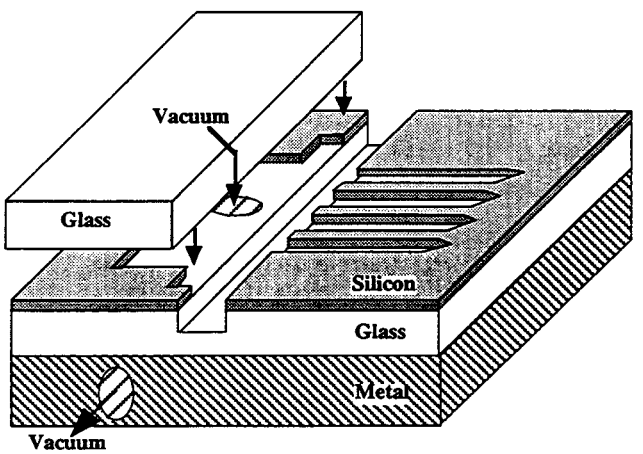


Figure 3: *Perspective view of a micromachined assembly fixture used for bending the beams on probes intended for use in 3D microassemblies.*

It is relatively easy to bend the gold beams if the probes can be held firmly. Because the probes are very small and quite thin, ordinary clamps either do not hold them firmly enough or can break them. One way to solve this problem is to attach the probes temporarily to a glass substrate, bend the beams over the edge of the glass, and then subsequently release them. This process is simple in principle but nonetheless tedious and time-consuming in the absence of a custom micromachined assembly tool.

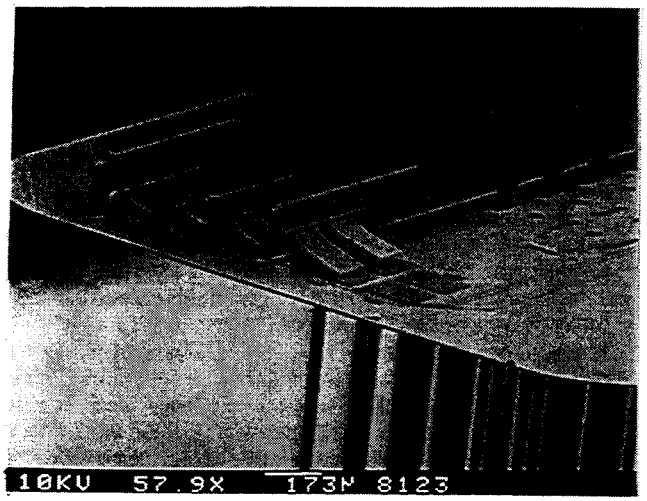


Figure 4: A partially-assembled probe array, showing the beam lead transfers before bonding.

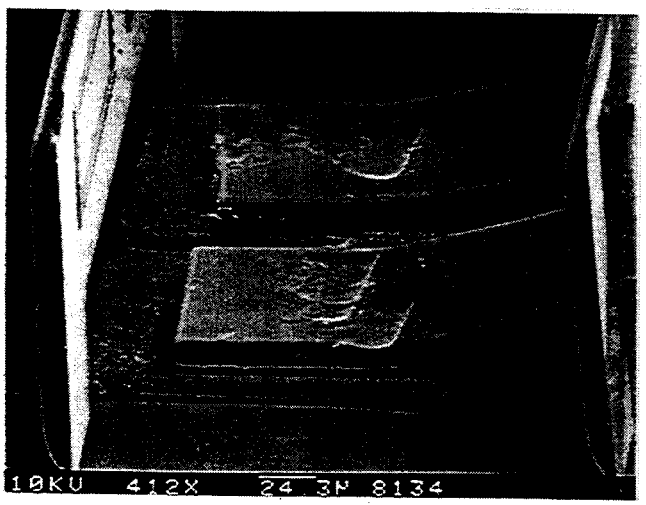


Figure 5: View of two ultrasonically-bonded lead transfers.

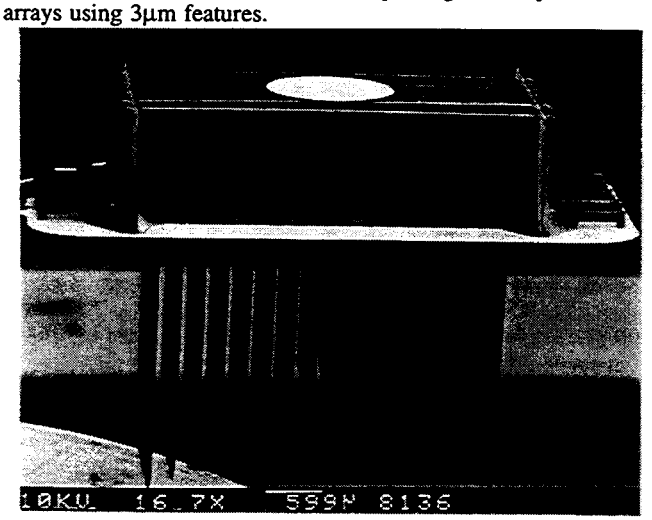
Figure 3 shows the schematic diagram of such a tool. A mold is lithographically defined to match the shape of the probes being used. A vacuum is applied to the mold fixture to hold the probe in it, and a piece of glass is used to clamp the probe firmly in position. Only one mask is needed for the formation of such a fixture. A deep boron diffusion is first used to define the mold recess. After drilling a vacuum access hole in a glass substrate, the glass is anodically bonded to the silicon mold wafer, and the wafer is then etched in EDP to remove all but the boron-doped areas. Thus, a probe mold is formed on the glass substrate. A groove 200µm wide and 200µm deep is next formed using a wafer saw. Such saw cuts produce a very clean edge. The mold is then glued to a metal support base having a corresponding vacuum hole to complete the fixture.

With this tool, the beams can be easily bent over into the groove in the glass using a special bending tool. The beams can be bent in either direction, forward or backward, and we have chosen to bend the beams backward under the probe so that direct gold-to-gold bonds can be made to the platform pads. A short overhang of dielectric under the beam prevents shorts to the probe substrate.

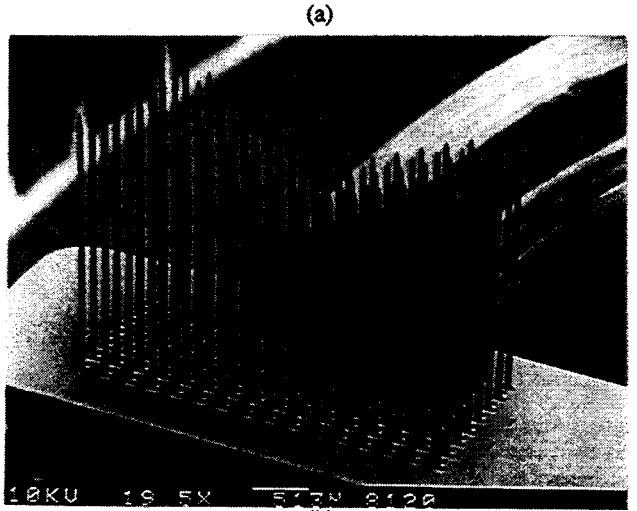
With the beams bent, the probes are inserted into slots in the platform, which is held on the mounting stage of a dedicated microassembly station equipped with two three-axis micromanipulators. Individual holes are used for each shank in order to give the center of the platform as much strength as possible. In addition, holes have been placed in the platform between probes to encourage tissue to grow through the platform to anchor it to the cortical surface.

It was initially very difficult to put the spacers in place when mounting the probes. The probes are not initially parallel to each other and tend to stick back to back when the spacer is put in. We therefore use a micromachined comb structure on a second manipulator to separate the probes one by one from the side. With the spacers in place, the fork is removed. The fork structure has 50µm spaces between the tines of the comb, with shanks that are tapered from one side to the other. It is very easy to use and replace. Figure 4 shows a partially assembled array of four probes before bonding.

Once the probe array is cemented together using a drop of Silastic, the beams are bonded to pads on the platform using conventional ultrasonic bonding with a wire-free wedge, as shown in Figure 5. Finally, the bonding areas are covered with Silastic to provide good insulation over the lead transfers. Figure 5 shows two views of a fully-assembled 3D array of eight probes and 128 shanks (an 8x16-shank array). The probe-to-probe separation is 200µm, which matches the shank-to-shank distance. The maximum shank length is 2.5mm. The height of the central part of the probe above the platform (1mm) mimics what will be required for CMOS multiplexing circuitry on active arrays using 3µm features.



(a)



(b)

Figure 6: SEM views of an 8x16-shank 8-probe array. The shanks are on 200µm centers. (a) Top view. (b) Bottom view.

EXPERIMENTAL RESULTS

For testing purposes, each of the probes in these arrays has four sites that are brought out to pads on the platform. The leads from the center three probes connect to a 12-pin percutaneous plug via an integrated ribbon cable. These arrays have been successfully tested in-vitro and in-vivo. The typical site impedance at 1kHz is about 2 MΩ before iridium activation, which represents worst case from the standpoint of signal transmission. An equivalent circuit of the probe-platform-cable assembly is shown in Figure 7. Without on-chip circuitry, the signal attenuation is less than 10% (set by the shunt capacitance of 18pF), with interchannel crosstalk (set by the connector and external leads) below 0.5%. Since the interchannel coupling on the assembly itself is less than 1fF, with on-chip signal amplification and buffering the crosstalk should be <0.01%, making simultaneous recording and stimulation possible. Typical neural recordings in guinea pig cortex are shown in Figure 8. A series of chronic implants are now being undertaken to assess the tissue reaction to such arrays.

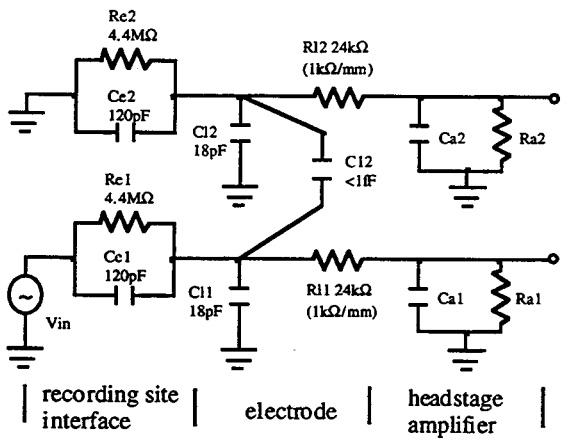


Figure 7: An equivalent circuit showing the various electrical elements that form the signal path in a typical probe assembly.

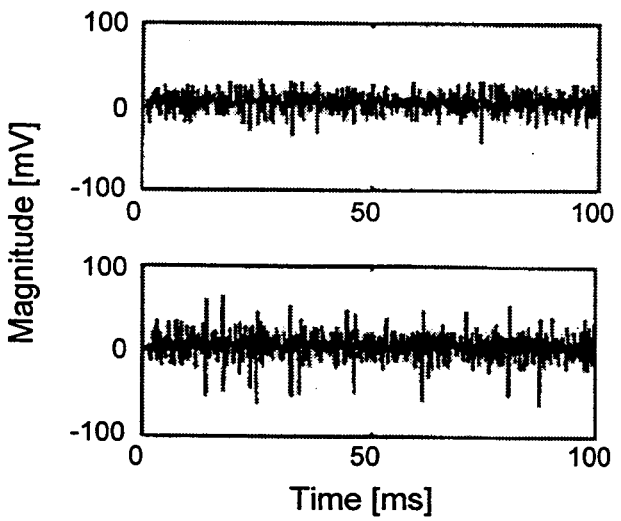


Figure 8: Simultaneous single-unit activity recorded from guinea pig cortex using an 8x8-shank 3D probe array.

CONCLUSIONS

Gold electroplated beams are being used to realize orthogonal lead transfers in forming 3D microelectrode arrays. This approach, implemented using simple micromachined assembly tools and fixtures, meets the needs of reliable array fabrication with high yield. It is easier to implement than previous approaches, is more reliable in terms of its interconnections, consumes less bonding area, and results in less tethering force on the tissue. The entire assembly procedure for an 8x16-shank array can be finished within half an hour. In further extending this approach to large numbers of sites, however, active multiplexing on the probes will be required. Using active multiplexing, arrays having 1024 sites on 200-400μm centers over 8-64mm^3 tissue volumes are being developed.

ACKNOWLEDGMENTS

The authors wish to thank Drs. F. T. Hambrecht and W. J. Heetderks of the Neural Prosthesis Program, National Institutes of Health (NINDS), for their guidance and support of this work under contract NIH-NINDS-NO1-NS-4-2303.

REFERENCE

1. K. Najafi, K. D. Wise, and T. Mochizuki, "A High-Yield IC-Compatible Multichannel Recording Array," *IEEE Trans. Electron Devices, 32*, pp. 1206-1211, July 1985.

2. A. Bragin, G. Jando, Z. Nadasdy, J. Hetke, K. Wise, and G. Buzsaki, "Gamma (40Hz-100Hz) Oscillation in the Hippocampus of the Behaving Rat," *J. Neuroscience, 15*, pp. 47-60, January 1995.

3. R. R. Carter and J. C. Houk, "Multiple Single-Unit Recordings from the CNS using Thin-Film Electrode Arrays," *IEEE Trans. Rehab. Engr., 1*, pp. 175-184, September 1993.

4. J. Ji and K. D. Wise, "An Implantable CMOS Circuit Interface for Multiplexed Microelectrode Recording Arrays," *IEEE J. Solid-State Circuits, 27*, pp. 433-443, March 1992.

5. J. L. Lund and K. D. Wise, "Chip-Level Encapsulation of Implantable CMOS Microelectrode Arrays," *Digest Solid-State Sensor and Actuator Workshop*, Hilton Head, S.C., pp. 29-32, June 1994.

6. C. Kim and K. D. Wise, "A 64-Site Multiplexed Low-Profile Neural Probe with On-Chip CMOS Circuitry," *Digest 1994 IEEE Symposium on VLSI Circuits*, Honolulu, June 1994.

7. J. F. Hetke, J. L. Lund, K. Najafi, K. D. Wise, and D. J. Anderson, "Silicon Ribbon Cables for Chronically-Implantable Microelectrode Arrays," *IEEE Trans. Biomed. Engr., 41*, pp. 314-321, April 1994.

8. P. K. Campbell, K. E. Jones, R. J. Huber, K. W. Horch, and R. A. Normann, "A Silicon-Based Three-Dimensional Neural Interface: Manufacturing Processes for an Intracortical Electrode Array," *IEEE Trans. Biomed. Engr., 38*, pp. 758-767, August 1991.

9. A. C. Hoogerwerf and K. D. Wise, "A Three-Dimensional Microelectrode Array for Chronic Neural Recording," *IEEE Trans. Biomed. Engr., 41*, pp. 1136-1146, December 1994.

10. M. P. Lepselter, "Beam-Lead Technology", *Bell System Technical Journal, 45*, pp. 233-254, February 1966.

PLASMA-ETCHED NEURAL PROBES

David T. Kewley[1], Matthew D. Hills[2], David A. Borkholder[2], Ion E. Opris[2], Nadim I. Maluf[2], Christopher W. Storment[2], James M. Bower[1], and Gregory T. A. Kovacs[2]

[1]Computation and Neural Systems, MC 216-76
California Institute of Technology
Pasadena, CA 91125

[2]Center for Integrated Systems, CISX 202
Stanford University
Stanford, CA 94305-4075

ABSTRACT

This paper presents a new method for microfabricating silicon-based neural probes that are designed for neurobiology research. Such probes provide unique capabilities to record high-resolution signals simultaneously from multiple, precisely defined locations within neural tissue. The fabrication process utilizes a plasma etch to define the probe outline, resulting in sharp tips and compatibility with standard CMOS processes. A low-noise amplifier array has been fabricated through the MOSIS service to complete a system that has been used in multiple successful physiological experiments.

INTRODUCTION

It is necessary to study the simultaneous functioning of many neurons in order to understand how the nervous system works [1]. Until now, most investigations into the functioning of groups of neurons have been conducted by studying one neuron at a time, then pooling the results into a summary that has been assumed to represent the activity of an entire population of cells. By using multiunit recording techniques, however, it has become clear that correlations among the activities of neurons often carry information that cannot be found by studying neurons separately. For instance, a recent study in the olfactory bulb focused on the considerable variability seen in the responses of individual mitral cells to odor stimulation [2]. With a single-neuron recording, this variability would have appeared random, but simultaneous recordings revealed it to be coordinated among several cells. In many other applications, it has been found to be crucial to study multiple cells simultaneously (e.g., [3,4]).

Of all the available techniques for studying neural function, it appears that only arrays of extracellular microelectrodes have the ability to record the activities of many individual neurons simultaneously, on a fast time scale, and in intact, behaving animals. Extracellular arrays are most commonly made from fine wires [2-4]. Wire arrays are attractive to the neurophysiologist because they are dependable. They can be built in the laboratory using readily available materials, and they have been demonstrated by many researchers to be capable of acquiring high-quality data.

Micromachined neural probes readily offer several advantages over wire arrays, including: (1) individual shafts easily support multiple electrodes, which facilitates the study of local networks; (2) relative electrode locations are precise and repeatable; and (3) probe shafts displace less volume than wire bundles, causing less neural damage.

A longer-term advantage offered by silicon neural probes is the possibility of integrating circuits directly on the probes. Neurophysiologists are pushing the technology toward a greater number of electrodes per array [2-4], and are hoping to eventually use systems with hundreds of channels. As the number of electrodes increases, it becomes nontrivial to make connections to the electrodes. Active probes with integrated amplifiers and multiplexers will be necessary to keep the packaging problem tractable, particularly in mobile, chronically implanted animals.

Microfabricated neural probes were first developed in the late 1960s [5]. Since then, several groups have developed neural probes [6-13], including two that have integrated circuitry [6,13,14]. Almost all of these previous efforts have produced probes with cross-sectional dimensions near the electrodes greater than 100 μm, which is likely to result in significant damage to the nearby neurons. Only two [6,7] have produced probes with properties that appear to be suitable for use in basic neuroscience.

The defining step in any neural probe fabrication process is the probe shaping technique. Of the two previous high-quality probe fabrication efforts, one uses a diffused p++ etch stop together with an anisotropic wet etch to define the probe shape [6]. The other uses (110) oriented wafers and an anisotropic wet etch, stopping on the {111} planes, to define the shaft sidewalls [7]. Both processes require extra steps specifically to produce sharp tips, and both require significant CMOS process modifications to incorporate circuits.

Neurons communicate principally via transient (≈1 ms) electrical signals called action potentials. When recorded extracellularly, these events occur as potentials with typical amplitudes up to 500 μV, and have significant power in a frequency band from tens or hundreds of Hz up to several kHz. A field potential is another, lower-frequency, signal that is also often of interest to neurobiologists. These signals typically have significant power from below one Hz up to hundreds of Hz, and can range in amplitude up to a few mV.

In order to record from a relatively intact neural network, an electrode array must be designed to minimize the damage it inflicts when it is inserted into the brain. The electrodes should be constructed of biocompatible materials, particularly if they will be implanted chronically. The electrode shafts should be straight, smooth, and stiff, and have sharp tips and small cross-sections. The shafts should be aligned with the insertion device so that they are inserted straight into the tissue [15]. These geometrical and mechanical constraints are particularly important when recording from the vertebrate brain, which is wrapped in a membrane called the pia mater. If the electrode dimples (locally depresses) the brain during penetration of the pia, the superficial neural tissue will probably be damaged.

0-9640024-1-8 $3.00 © 1996 TRF

SYSTEM DESIGN

The goal of this work has been to develop a complete neural probe system and to make it available to the research community. To facilitate this development, the probe and circuit fabrication have been pursued independently, with care taken to preserve the ability to integrate the two in the future. This has meant keeping all processing temperatures at or below 300°C.

Probe Design. The current probes are fork shaped, having one or more parallel shafts projecting from a base (Fig. 1). Current designs support 16 iridium electrodes placed near the tips of the shafts, with a single reference electrode near the midpoint of the middle shaft.

The electrical properties are primarily determined by the electrode characteristics [16]. The electrode design is based upon previous research [18].

Plasma etching provides the flexibility of defining arbitrary probe outlines. A second important feature is the low curvature in the probe shafts. Najafi, et al., accomplished this by depositing alternating layers of insulators with opposite stresses [6]. In the present design, a low-stress nitride has been used, eliminating the need for stress cancellation. Other advantages of the nitride are that it is deposited at low temperatures (300°C), it is very conformal, and it has an extremely low pinhole density.

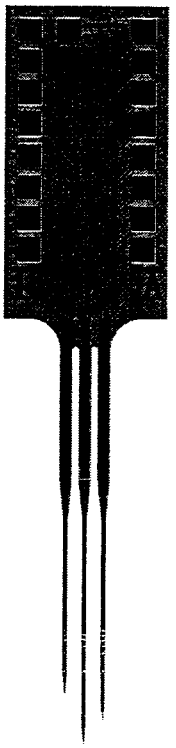


Figure 1. Plan view of a probe.

The use of two metal layers—one for traces, the other for the electrodes—minimizes the shaft cross-section. Another benefit of this approach is the flexibility to place the electrodes at arbitrary positions on the shaft, such as in tetrode configurations. This capability, together with the plasma-based probe shaping technique, has facilitated the design of fifteen different probes, each intended for a particular physiological experiment or engineering test.

Circuit Design. An 18-channel preamplifier–filter was implemented on the MOSIS 2 µm low-noise analog process (Fig. 2). Each channel is amplified with respect to the reference electrode input. The input range is sufficient to accommodate the range of signals expected (10 µV–5 mV). The passband gain of 150 is provided in three stages. The output buffer, B, was designed to drive a 10 nF load. The output is fed back into A3 to set the DC output level to an on-chip reference value. The low pass cutoff is fixed at 35 kHz by the shunt capacitor between A3 and B. The coupling capacitors in combination with M1–M4, which are biased in the sub-threshold regime, create the 2-pole high pass filter. The *hpf* signal is derived from a chip input, allowing the cutoff to vary between 1 Hz and 10 kHz.

A prototype board was made to mount the preamplifier-array chip directly to the probe carrier printed circuit board. A separate signal-conditioning amplifier was built using commercial components. It provides a gain of 2.8 to match signal levels to the input range on the data acquisition system. This amplifier also provides a two-pole rolloff at 3 kHz for anti-aliasing.

PROBE FABRICATION

Starting with standard 4-inch, (100)-oriented, p-type wafers, the probe fabrication process proceeds in three phases (Fig. 3):

1. Thin-film deposition (electrodes, bond pads, and insulated interconnect traces)
2. Probe-shaping etches
3. Finishing steps

Thin Film Deposition. Initially, a 1 µm insulation layer of silicon nitride was deposited in a PECVD system (STS-Electrotech) using dual-frequency RF excitation of the plasma. Process parameters were adjusted to yield low stress (typically <100 MPa tensile).

The first metal layer was then deposited and patterned using a liftoff process. After defining the liftoff structure (1 µm of photoresist over 500 nm of polysilicon), 25 nm of titanium (to improve adhesion) and 500 nm of gold were sequentially deposited by electron-beam evaporation. The liftoff was accomplished by dissolving the photoresist in acetone, and the polysilicon was plasma etched with 50 sccm SF_6 and 50 sccm C_2ClF_5 at 260 mW/cm^2 and 150 mTorr.

The second layer of silicon nitride, 0.7 µm thick, was deposited using the PECVD process. Contact holes to the bond pads and for gold–iridium contacts were etched in the nitride using a plasma with 50 sccm SF_6 and 50 sccm C_2ClF_5 at 270 mW/cm^2 and 150 mTorr.

Another liftoff process was used to pattern the iridium as it was deposited, using a bilayer of photoresist over aluminum. A 75 nm layer of silicon nitride was deposited to avoid the formation of unwanted intermetallics between the aluminum and the gold exposed in the contact cuts. A 500 nm thick layer of aluminum was sputter deposited, followed by 1 µm layer of spun-on photoresist that was then patterned. The aluminum was etched using a standard wet etch until the undercut was 1.5–2.0 µm. Finally, the thin nitride layer was removed from the contact holes with the nitride plasma etch.

The wafers were placed in a custom magnetron sputtering system to deposit 300 nm of iridium after a 30 nm tacking layer of chromium. To lift off the unwanted Cr/Ir, the wafers were

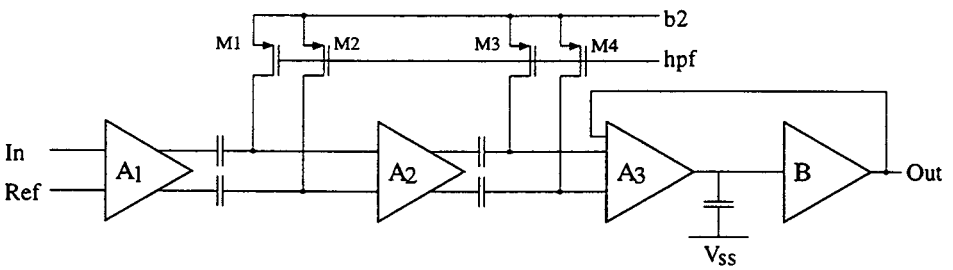


Figure 2. Schematic of one channel of the preamplifier circuit showing the filtering elements.

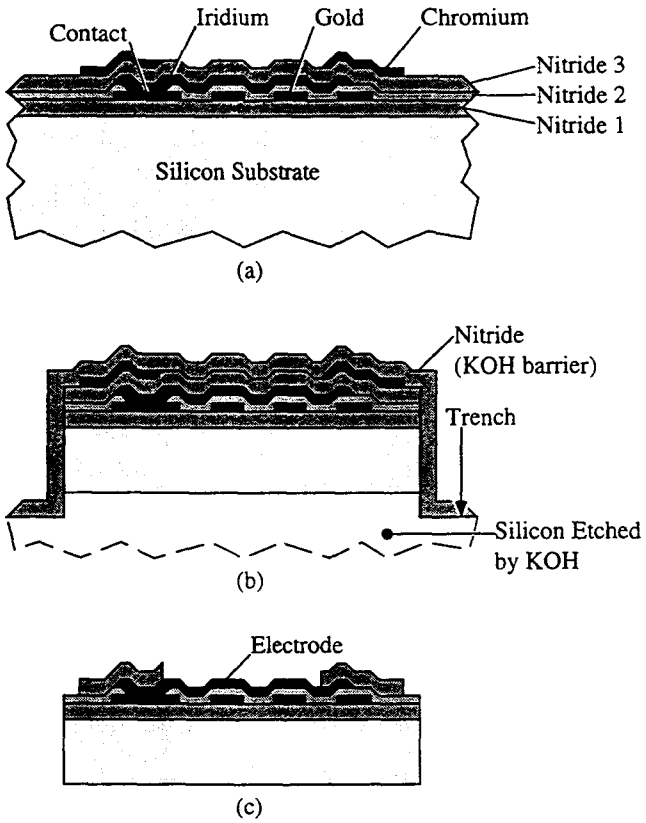


Figure 3. Schematic representation of the probe fabrication process flow, seen in a cross-section through an electrode site (not to scale). Parts (a), (b), and (c) show the end of the three major processing phases discussed in the text.

placed in a fresh solution of sulfuric acid and hydrogen peroxide. Any remaining Cr/Ir was then removed by a high-pressure water stream in a mask cleaner.

Silicon nitride was deposited to a thickness of 1.0 µm as the passivation layer for the probe. A 30 nm layer of chromium was deposited and patterned to later serve as the nitride etch mask. Fig. 3(a) depicts the cross-section at this stage.

Probe Shaping. A 15 µm layer of AZ-P4620 photoresist (Shipley) was spun on to the wafers and patterned as a mask for the front-side plasma etches. The nitride over the trenches was etched using RIE with 15 sccm CHF_3 and 4 sccm O_2 at 80 mW/cm^2 (sufficient for a -400 V bias) and 50 mTorr. The underlying silicon was plasma etched to a depth of 30–40 µm in 100 sccm SF_6 and 70 sccm C_2ClF_5 at 260 mW/cm^2 and 100 mTorr.

After the photoresist was stripped, a 1 µm layer of silicon nitride was deposited to protect the front of the wafer during the back side KOH etch. The wafer was diced in order to enhance the etching uniformity, then the dice were placed in a 15% KOH solution at 80°C. This solution thinned the silicon substrate from the back side at a rate of 1 µm/min. A visual end condition was used, as the nitride in the trenches became transparent when the underlying silicon had been removed. After a few minutes overetch, the final probe thickness was ≈25 µm. See Fig. 3(b).

Finishing Steps. An RIE etch was used to strip the KOH barrier nitride. As the etch continued, portions of the passivation nitride layer were etched with the chromium layer

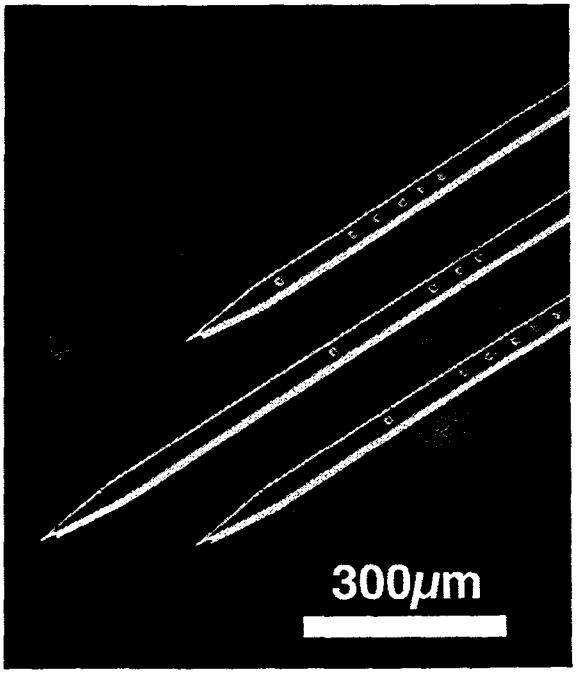


Figure 4. SEM of the tip region of a three-shaft probe, with 16 electrode sites: six on the left and right shafts, and four on the center shaft. The shafts are staggered in length.

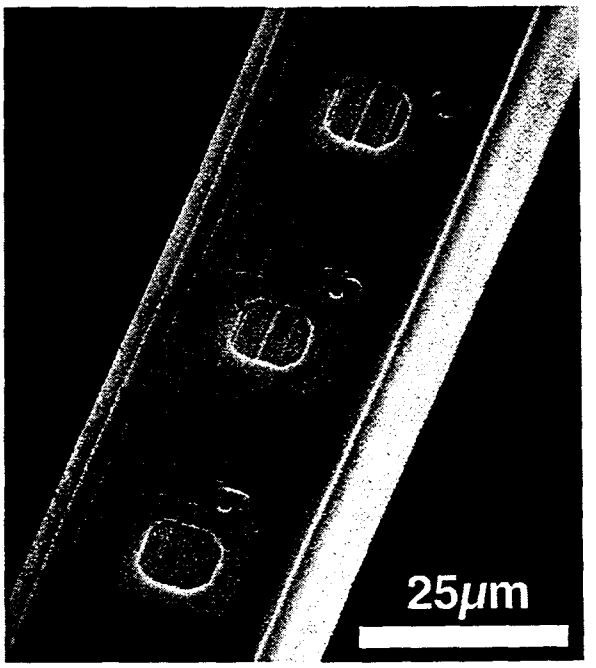


Figure 5. SEM of the top three electrode sites on a six-site shaft. Note the interconnect traces running beneath the electrodes, and the small dimples where the metal-to-metal contacts are made.

serving as a mask, exposing the electrode sites and the bond pads. Finally the chromium itself was removed by immersion in a chromium etching solution. See Fig. 3(c).

The completed probes were mounted on PC boards, with the shafts projecting off the boards. After wire-bond connections were made between the probe and the boards, the

probe base and wire bonds were encapsulated in an epoxy (87-GT, Epoxy Technology, Inc).

EXPERIMENTAL RESULTS

Bench-Top Testing. SEMs of the neural probes are shown in Figs 4-6. Probe tip sharpness was evaluated in a SEM, which shows a tip radius of curvature of less than 0.25 μm. Preliminary impedance measurements indicate an electrode capacitance of 40 pF with a shunt capacitance of 20 pF.

In gain measurements across 72 amplifier channels, midband gain was found to be 43.5±0.5 dB. Total harmonic distortion was found to be less than 1% with a 5 mVpp input and less than 0.25% with a 1 mVpp input. The input-referred noise is 8 μV rms for a bandwidth between 10 Hertz and 10 kHz. Operating at ±2.5 V supplies, the entire chip dissipates 10 mW.

Neural Recordings. The subjects for the physiological tests of the probe system were 90–180 day old female albino Sprague-Dawley rats, weighing about 250 g. During an experiment, an animal was anesthetized with ketamine, xylazine, and acepromazine, together with a small initial dose of pentobarbital. Surgery was performed to allow access to a portion of the right somatosensory neocortex, which responds to touch stimuli. The exposed brain was continuously bathed in a warm saline solution. The animal was euthanized with an overdose of pentobarbital at the end of the experiment.

To make a recording, a probe assembly was mounted on a motorized micromanipulator, positioned over the exposed brain, and slowly inserted. The deepest electrode on one of the probe shafts was monitored during insertion using both an oscilloscope and an audio system. Recordings were made at locations where several of the electrodes were close enough to neurons to pick up clear neural activity. Signal quality on each channel was comparable to that obtained with a conventional extracellular microelectrode. See Fig. 7 for an example of the data recorded.

The neural recordings demonstrate that the probes and electronics work, and that they are capable of recording high-quality data. These tests have also uncovered some areas where more work is needed. Despite the sharp tips, the probes still cause some dimpling, especially in experiments conducted in the rat cerebellum, where the pia is relatively tough. In one experiment, the pia was removed and a probe was inserted into the cerebellum with no dimpling, verifying that the pia is the major source of resistance to electrode entry.

Furthermore, although the probes have recorded action potentials in all preparations tried, some of which did not yield high-quality recordings. Besides neocortex, olfactory bulb slices (mitral cells) and deep portions of cerebellum (Golgi cells and deep Purkinje cells) have produced recordings with good signal-to-noise ratios. In the superficial cerebellum, the neonatal rat brainstem nuclei, and the locust antennal lobe, the signal quality obtained with probes has been worse than that typically obtained with conventional single electrodes.

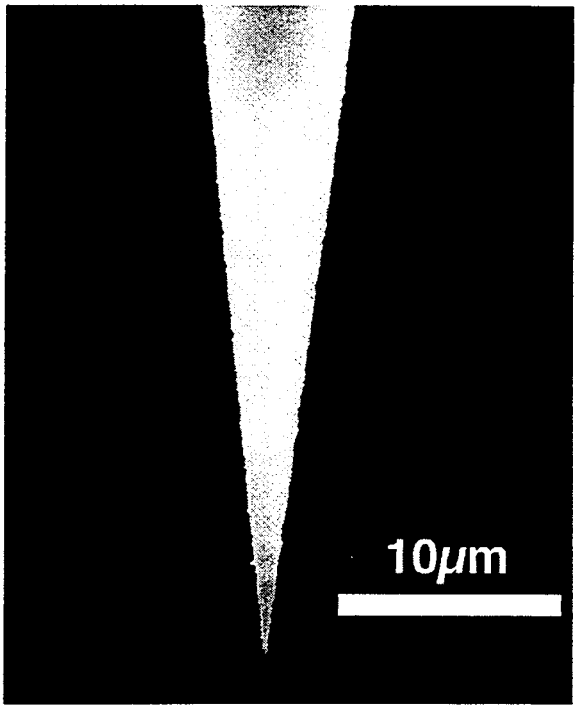


Figure 6. SEM of a probe tip (top view), showing the sharpness yielded by the dry etch process.

Another problem is that the probes are light sensitive. Photogenerated carriers in the probe substrate appear to couple capacitively into the gold traces via depletion regions formed at the surface of the silicon.

CONCLUSION

The positive results obtained so far indicate that the basic probe design is sound. It is believed that the probes are damaging nearby neurons during insertion, possibly because of a misalignment of the probe shafts to the insertion axis and snagging of the tissue by an abrupt edge in the nitride near the tips. The shaft misalignment is visually apparent and is probably due to stresses in the encapsulating adhesive lifting the probe shafts off the carrier PC board. New probe mounting methods are being explored to solve the adhesive stress problem, and the mask set will be redesigned to eliminate the nitride edge. The light sensitivity can be eliminated by placing a metal layer under the interconnect traces to provide electrical shielding from the substrate.

ACKNOWLEDGMENTS

This work has been funded by the National Science Foundation under grant # PC147129-1. Additional support was provided by Medtronic, Inc. Karim Elaagouby, Mike Wehr, and Nick Mellen have contributed to this research by providing additional experiment preparations.

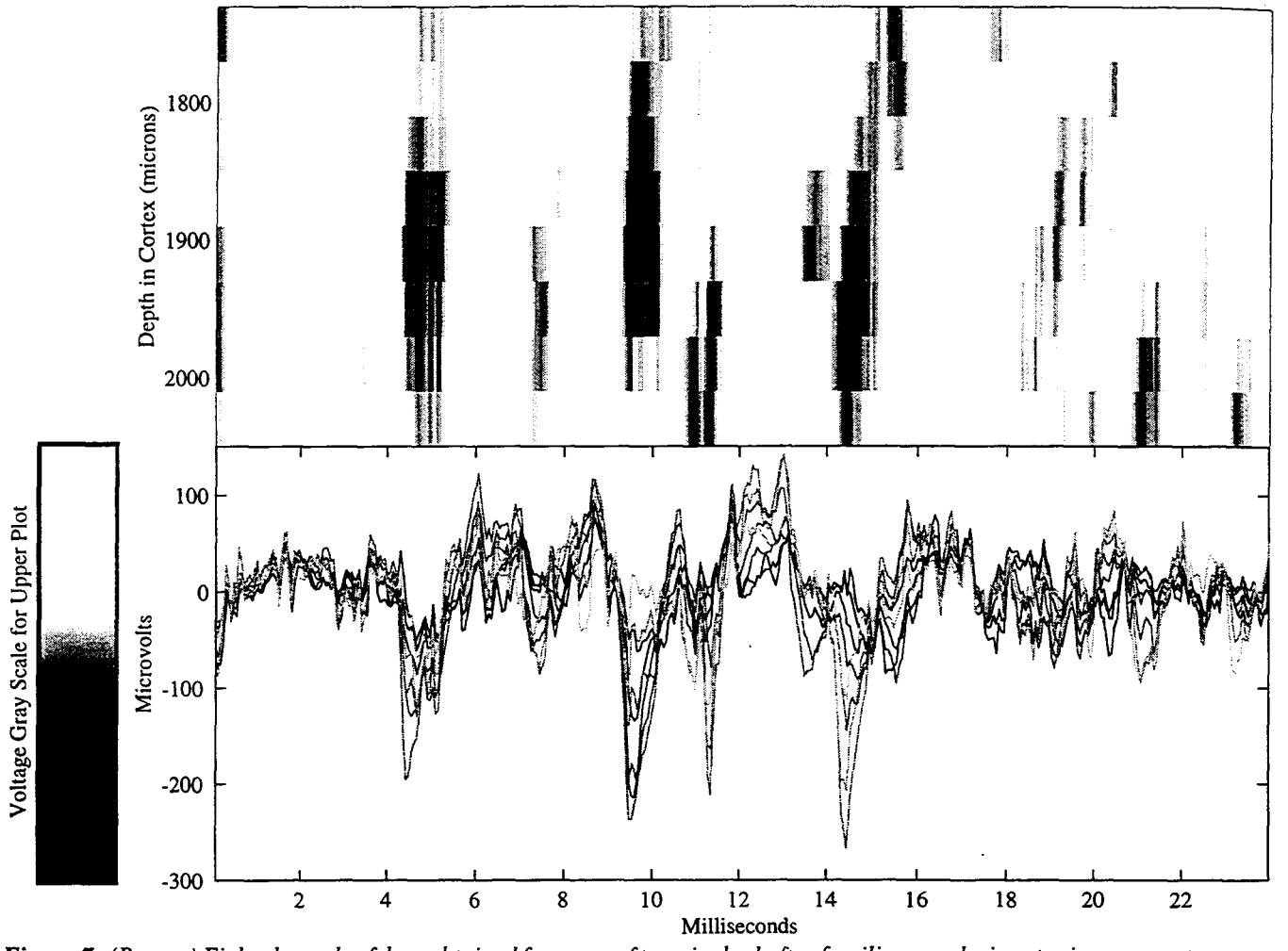


Figure 7. *(Bottom) Eight channels of data obtained from one of two single shafts of a silicon probe in rat primary somatosensory cerebral cortex. The electrode sites were spaced vertically on the probe needle at 40 μm intervals, covering 280 μm of cortical depth. The data presented shows a spontaneous burst of activity recorded in an anesthetized rat. (Top) The same data plotted to show spiking with respect to time and the depth into the cortex. Darker regions indicate increased neural activity.*

REFERENCES

1. S. A. Deadwyler and R. E. Hampson, "Ensemble Activity and Behavior: What's the Code?", *Science, 270,* 1316 (1995).

2. U. S. Bhalla and J. M. Bower, "Multi-Day Recordings from Olfactory Bulb Neurons in Awake Freely Moving Rats: Spatial and Temporal Patterns in Odorant Responses", *Journal of Computational Neuroscience* (in press).

3. M. A. Wilson and B. L. McNaughton, "Reactivation of Hippocampal Ensemble Memories During Sleep", *Science, 265,* 676 (1994).

4. M. A. L. Nicolelis, L. A. Baccala, R. C. S. Lin, and J. K. Chapin, "Sensorimotor Encoding by Synchronous Neural Ensemble Activity at Multiple Levels of the Somatosensory System", *Science, 268,* 1353 (1995).

5. K. D. Wise, J. B. Angell, and A. Starr, "An Integrated-Circuit Approach to Extracellular Microelectrodes", *IEEE Transactions on Bio-Medical Engineering, BME-17,* 238 (1970).

6. K. Najafi, K. D. Wise, and T. Mochizuki, "A High-Yield IC-Compatible Multichannel Recording Array", *IEEE Transactions on Electron Devices, ED-32,* 1206 (1985).

7. L. D. Clark, "System for Chronic Neural Signal Transduction Processing and Control", Ph.D. dissertation, Massachusetts Institute of Technology (1990).

8. N. A. Blum, B. G. Carkhuff, H. K. Charles, Jr., R. L. Edwards, and R. A. Meyer, "Multisite Microprobes for Neural Recordings", *IEEE Transactions on Biomedical Engineering, 38,* 68 (1991).

9. M. Kuperstein and H. Eichenbaum, "Unit Activity, Evoked Potentials and Slow Waves in the Rat Hippocampus and Olfactory Bulb Recorded with a 24-Channel Microelectrode", *Neuroscience, 15,* 703 (1985).

10. E. Peeters, B. Puers, W. Sansen, J. Gybels, and P. DeSutter, "A Two-Wire, Digital Output Multichannel Microprobe for Recording Single-Unit Neural Activity", *Sensors and Actuators B, 4,* 217 (1991).

11. R. S. Pickard, P. Wall, and M. Ubeid, "Recording Neural Activity in the Honeybee Brain with Micromachined Silicon Sensors", *Sensors and Actuators B, 1*, 460 (1990).

12. O. J. Prohaska, F. Pacha, P. Pfundner, and H. Petsche, "A 16-Fold Semi-Microelectrode for Intracortical Recording of Field Potentials", *Electroencephalography and Clinical Neurophysiology, 47*, 629 (1979).

13. K. Takahashi and M. Tadayuki, "Integration of Multi-Microelectrode and Interface Circuits by Silicon Planar and Three-Dimensional Fabrication Technology", *Sensors and Actuators, 5*, 89 (1984).

14. R. S. Pickard, "A Review of Printed Circuit Microelectrodes and Their Production", *Journal of Neuroscience Methods, 1*, 301 (1979).

15. D. J. Edell, V. V. Toi, V. M. McNeil, and L. D. Clark, "Factors Influencing the Biocompatibility of Insertable Silicon Microshafts in Cerebral Cortex", *IEEE Transactions on Biomedical Engineering, 39*, 635 (1992).

16. G. T.A. Kovacs, "Microelectrode Models for Neural Interfaces," pp 121-165 in D.A. Stenger, T.M. McKenna (Eds) (1994), "Enabling Technologies for Cultured Neural Networks," Academic Press, San Diego.

17. C. M. Gray, P. E. Maldonado, M. Wilson, and B. McNaughton, "Tetrodes Markedly Improve the Reliability and Yield of Multiple Single-Unit Isolation from Multiunit Recordings in Cat Striate Cortex", *Journal of Neuroscience Methods, 63*, 43 (1995).

18. G. T. A. Kovacs, C. W. Storment, M. Halks-Miller, C. R. Belczynski, Jr., C. DellaSantina, E. R. Lewis, and N. I. Maluf, "Silicon-Substrate Microelectrode Arrays for Parallel Recording of Neural Activity in Peripheral and Cranial Nerves", *IEEE Transactions on Biomedical Engineering, 41*, 567 (1994).

19. J. Ji and K. D. Wise, "An Implantable CMOS Circuit Interface for Multiplexed Microelectrode Recording Arrays", *IEEE Journal of Solid State Circuits, 27*, 433 (1992).

HIGH-DENSITY ARRAYS OF VALVES AND INTERCONNECTS FOR LIQUID SWITCHING

Luc Bousse*, Emiel Dijkstra, and Olivier Guenat*****
Molecular Devices Corp.
1311 Orleans Drive, Sunnyvale, CA 94025

ABSTRACT

We describe a process for the fabrication of high-density arrays of valves designed for switching liquids. To make small valves capable of opening and closing large orifices for liquid flow requires a flexible movable membrane. We use silicone rubber, because it is elastomeric, makes good seals, and can be spin-coated. Flow channels are defined by a sacrificial photoresist. The behavior of individual valves is measured, and a model describing the flow is presented. Valve arrays at densities up to 16 valves/cm^2 have been demonstrated.

INTRODUCTION

Previously, we have described an eight-channel chip with 32 sensing spots for the measurement of the metabolic responses of living cells [1]. To improve time response and reduce the sample volume, we need an array of eight four-way valves on a chip that can be directly interfaced with the sensing chip. The type of interface that minimizes dead volume is to stack the valve chip and the sensor chip directly on top of each other. This defines the valve density required; since a four-way valve can be implemented with four one-way valves, we need 32 valves in the area of the chip, which is 5.3 cm^2. This defines the required valve density as at least 6 valves/cm^2; however, the constraint is even greater since the placement is not arbitrary. A density of 10 to 20 valves/cm^2 is desirable, and a smallest dimension of less than 1 mm.

Most microfabricated valves described previously work only with gases and are 5–8 mm on a side for a single valve, meaning maximum densities of 4 valves/cm^2. The aim of this work is to develop a technology for making high density arrays of valves for liquids. This must include an interconnection method for the individual components. Although the design specifications follow from a specific application, the process and devices presented here are not limited to this application, but can be used for a wide variety of purposes.

CHOICE OF MEMBRANE MATERIAL

To make valves smaller than 1 mm, yet maintain openings of about 100 μm for liquid flow requires a movable membrane material much more flexible than glass or silicon and capable of large deformations. For instance, to deflect a membrane 500 μm wide by 100 μm in the middle, an elongation of about 15% is needed (assuming the deflected shape is an ellipse). To avoid large forces in causing these displacements, a low Young's modulus is needed. One material with a low Young's modulus is polyimide (E is 3×10^3 MPa, as opposed to 1.9×10^5 MPa for silicon), which has been used for micropump membranes [2]. However, polyimide is not capable of large elongations. A class of materials with even lower modulus values are the elastomers, such as silicone rubber, which have also been used for microvalve membranes [3]. We chose silicone rubber because it has a very low modulus, is an elastomer which forms very good seals, and can be spin coated and cured on silicon wafers. Most types of mechanical actuators can be adapted to our valves, but for our application the most convenient and flexible is pneumatic actuation.

PROCESS DESCRIPTION

Given that we desire an elastomeric membrane, it is desirable to make these membranes a permanent part of the valve structure. Handling loose membranes to place in some assembly is difficult, and cannot be done on a wafer scale. Therefore, we spin-coat the uncured polymer on the silicon wafer, and then cure it at 150°C. By using an addition-cure two-component rubber, this forms a very tight bond between the rubber and the oxidized silicon. To form flow channels between the rubber and the silicon, a sacrificial layer is needed to avoid the bonding of rubber to the substrate in some areas. The simplest type of sacrificial layer is photolithographically patterned photoresist. Subsequent removal of the sacrificial layer will then create the regions where flow can occur. To provide access to the flow channel, holes have to be made through the wafer. To do this, we used anisotropic etching, which can terminate gently on the oxide layer on the other of the wafer. This overall process is similar to that described by Shoji [4] for devices in which the membrane is negative photoresist. The detailed description of the process steps is :

1. Starting material: 100 mm diameter low-doped silicon wafers, double-sided polished, 500 μm thick.
2. Wet oxidation to form a 1 μm thick oxide on both sides.

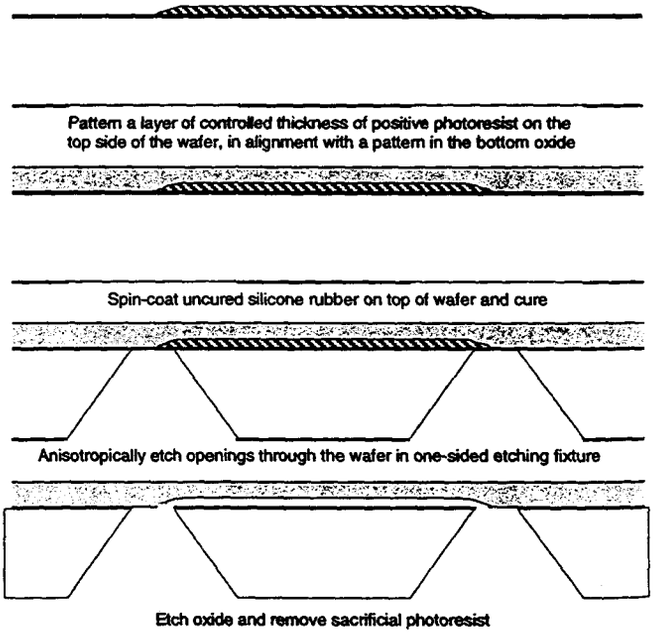


Figure 1. Process schematic

0-9640024-1-8 $3.00 © 1996 TRF

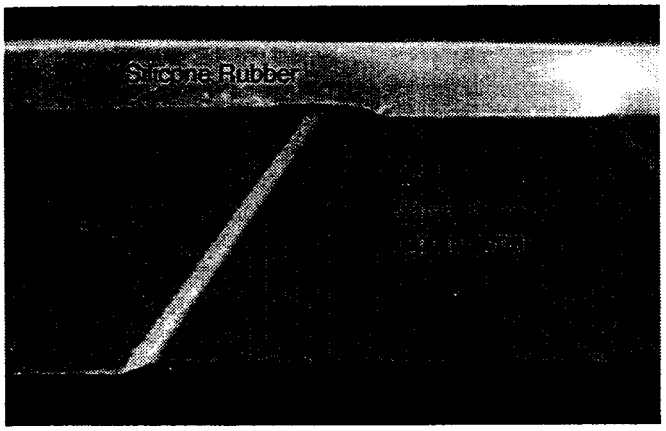


Figure 2. SEM close-up of the cross-section of a valve structure.

3. Pattern backside oxide with the interconnection channel pattern.
4. Anisotropically etch interconnection channel to desired depth.
5. Strip oxide and regrow it.
6. Photolithographically define openings in the oxide on the backside, for the holes through the wafer.
7. Spin-coat sacrificial photoresist on the frontside of the wafer, in alignment with the hole pattern on the backside.
8. Spin-coat uncured silicone rubber (Dow-Corning 96-083) on the frontside, on top of the photoresist, and cure (1 hour at 150°C).
9. Etch holes in the silicon from the bottom in an anisotropic etchant. This is done in a one-sided silicon etch fixture, so that the rubber on the frontside of the wafer is not damaged by the etchant. This etch step self-terminates when the oxide on the other side of the wafer is reached.
10. Etch the oxide windows in buffered HF, which also removes all the oxide exposed on the backside.
11. Remove the sacrificial photoresist layer in a resist removing solution.
12. Wafer dicing.

Fig. 1 shows the key steps in this process, but without including the interconnection channel. Fig. 2 shows an SEM close-up of the cross-section of an actual structure. The top geometry is defined in Fig. 3.

A SIMPLE MODEL TO DESCRIBE VALVE OPERATION

In this section, we will develop an analytical model for the flow Q in a valve structure such as shown in Fig. 1, as a function of the applied pressures. Those pressures are the driving pressure on the inlet side P_{dr}, and the pneumatic actuation pressure applied to close the valve P_{cl}. The pressure at the outlet is assumed to be zero, without loss of generality. The dimensions of the flow channel are typically 2000 μm in length, 500 to 1500 μm in width, and 0 to 100 μm in height, depending on the applied pressure. It is natural to model the flow in such a long and thin channel as the laminar flow between two parallel plates, which is given by:

$$Q = \frac{1}{12\mu}\frac{WH^3}{L}P_{dr} \tag{1}$$

where the dimensions W and L are defined in Fig. 3, μ is the dynamic viscosity of the fluid, and H is the height of the flow channel. This height depends on the displacement of the silicone rubber membrane due to the net pressure acting on it, namely $\Delta P = P_{dr} - P_{cl}$. The displacement of a rectangular membrane of width W and length L is given by [5]:

$$\Delta P = \frac{C_1\sigma t}{a^2}d + \frac{C_2 E t}{a^4}d^3 \tag{2}$$

where $2a$ is the smaller of W and L, E is Young's modulus, σ the residual stress, t the thickness of the membrane, and C_1 and C_2 are constants dependent on the ratio W/L and Poisson's ratio v. In the case of an elastomeric membrane, Young's modulus is very low, and thus the linear term is dominant (see measurements below). We will therefore assume that the membrane moves linearly with applied pressure, and we will call the proportionality constant α:

$$H = \alpha\,(P_{dr} - P_{cl}) \tag{3}$$

This equation neglects the contribution to H due to the sacrificial layer itself, which is usually much smaller than the displacement caused by pressure. Substituting this equation into (1) we obtain:

$$Q = \frac{1}{12\mu}\frac{W}{L}\alpha^3(P_{dr} - P_{cl})^3 P_{dr} \tag{4}$$

When $P_{dr} >> P_{cl}$ this can be approximated as:

$$Q = \frac{1}{12\mu}\frac{W}{L}\alpha^3(P_{dr} - P_{cl})^4 \tag{5}$$

although this approximation turns out to be useful in a wider range of pressures as well. Both equations (4) and (5) predict that flow will increase very rapidly with driving pressure, due to the flexibility of the membrane. When Q is high, however, the pressure loss in the openings through the wafer, and in the fittings used to access the valve is no longer negligible. This must be modeled by a flow resistance in series with the valve itself. This situation is similar to modeling the current-voltage characteristics of a semiconductor p-n junction: at high currents the effect of the series resistance of the semiconductor becomes dominant. If we call the externally measured pressure P_{ext}, then we can write:

$$P_{ext} = RQ + P_{dr} \tag{6}$$

where R is the flow resistance in series with the valve. The value of the approximation of equation (5) is that it is now possible to invert that equation to obtain:

$$P_{ext} = RQ + P_{cl} + \left(\frac{12\mu}{(W/L)\alpha^3}\right)^{1/4} Q^{1/4} \tag{7}$$

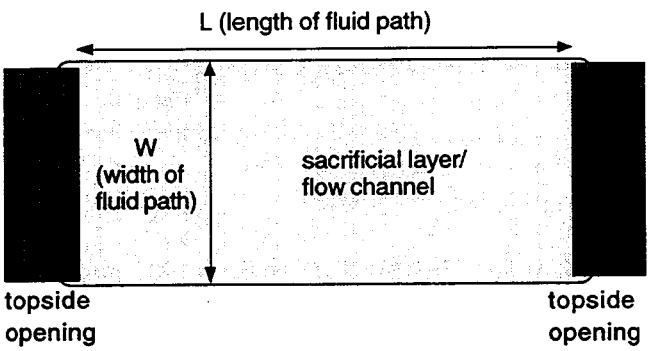


Figure 3. Definition of top geometry.

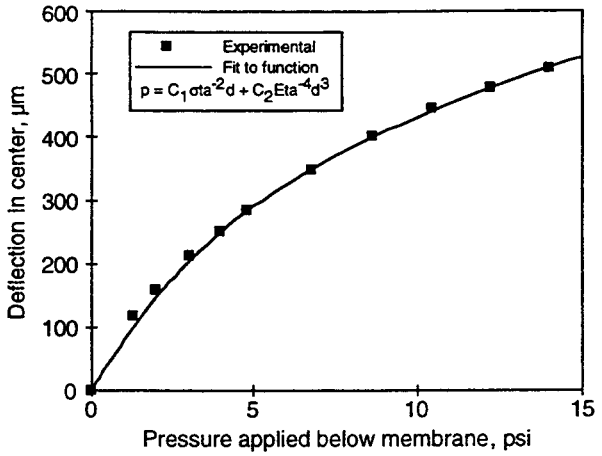


Figure 4. Experimental and theoretical load-deflection curves.

This equation models the entire family of curves of Q as a function of P_{ext} and P_{dr}, with two adjustable parameters, namely α and R.

MEASUREMENTS ON INDIVIDUAL VALVES

The first type of measurements we will present here are the load-deflection curves of the membranes as a pressure is applied underneath. These measurements, although they do not directly test the valve function, are useful on two levels. First, they allow the characterization of the silicone rubber membranes, in particular the residual tensile stress that keeps them stretched out. Second, by directly measuring the slope α used in Eq. (3), it is an important verification of the model presented in the previous section. Although we made some special-purpose square test structures similar to those used by Senturia *et al.* [6] for polyimide characterization, it turned out that a valve structure can be used as well, since pressure can be applied to the entire membrane through the two holes through the wafer. In that way, the same device can be used for measuring α with two different methods.

The device was tested in a fixture that allowed free observation of the top of the membrane, and the application of pressure at the bottom. Pressures were measured with solid-state pressure transducers, and the deflection height of the center of the membrane was measured with the calibrated focus of a microscope. The thickness of the silicone rubber was measured by looking at the diced cross-section at the edge of the chip with a calibrated microscope. Fig. 4 shows the data on a device with $W = 1239$ μm, $L = 2428$ μm, and membrane thickness 58 μm. A good fit with the function of Eq. (2) is obtained, and the linear term corresponds to $\alpha_{id} = 81.8$ μm/psi. The linear and cubic coefficients can be used together with Eq. (2) to give Young's modulus $E = 0.87$ MPa and the residual stress $\sigma = 0.28$ MPa. This corresponds to a quite large residual stress of 32%. We encountered values of E from 0.7 to 1.0 MPa, and σ from 0.23 to 0.28 MPa, while testing a range of samples with thicknesses from 29 to 107 μm. The thickness did not appear to affect either E or σ significantly.

The next type of test is to measure how well these structures work as valves. A different test fixture was used for this, one that included an external gasket to separate the two flow openings. Driving pressure was applied by pressurizing a vessel filled with deionized water. All tests were done at room temperature with filtered DI water. Flow was measured gravimetrically for high flow rates, and by bubble movement in a calibrated capillary for low flow rates. Fig. 5 shows data obtained on the same structure as used for Fig. 4, namely flow as a function of driving pressure, with closing pressure as a parameter. Each curve has an initial domain where Q increases very rapidly with pressure, and then the dependence becomes linear as the external flow resistances become dominant. The value of α found by fitting these curves to Eq. (7) is 65.3 μm/psi. As expected, this is somewhat smaller than α_{id}, since the channel height in Eq. (1) is an average, not the maximum found in the middle of the membrane. Taking account of this, the two values of α agree very well.

Eq. (7) predicts that applying a closing pressure shifts the curves by the same amount as the closing pressure. The dotted curves in Fig. (5) show how well that model works when higher closing pressures are applied. As expected, when P_{cl} is 2 psi or greater, the agreement with data becomes poor. There are two possible reasons for this: first, the approximation leading to Eq. (5) is not valid anymore. Second, Eq.(3) is not entirely correct, since the closing pressure is applied uniformly to the top of the membrane, while the driving pressure is only applied to one edge of the membrane. A simple way of modeling this is to assume that the average value of the pressure under the membrane in the channel is not P_{dr} but rather P_{dr}/β. This leads to Eq. (7) being modified to:

$$P_{ext} = RQ + \beta P_{cl} + \beta\left(\frac{12\mu}{(W/L)\alpha^3}\right)^{1/4} Q^{1/4} \tag{8}$$

The continuous line in Fig. (5) represents this model with three adjustable parameters, where β=1.2. Clearly, adding this parameter improves the fit of the model considerably.

Another way of looking at valve characteristics is as a function of closing pressure, for a fixed driving pressure. This shows how well a valve can shut off flow, and is best seen in logarithmic coordinates. Fig. (6) shows data in this way for three different valves that have the same W and L, but different thicknesses of the sacrificial layer. Clearly, thinner sacrificial layers produce much lower leakage flows. This is probably due to leakage at the edge of the flow channel, which does not entirely collapse when closing pressure is applied. This can be avoided by extending the sacrificial layer beyond the top openings of the flow holes.

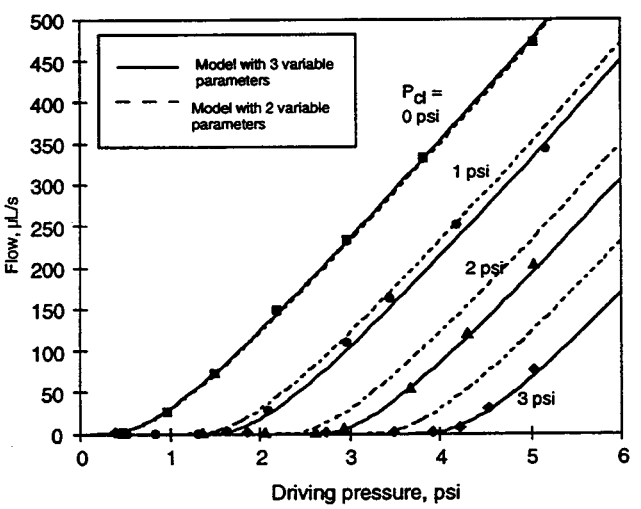


Figure 5. Flow as a function of driving pressure with closing pressure as a parameter.

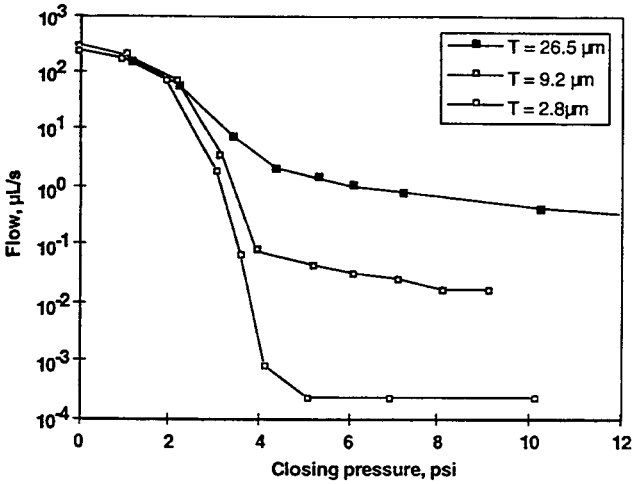


Figure 6. Flow as a function of closing pressure for a fixed 4 psi driving pressure for three sacrificial layer thicknesses.

TESTING OF INTEGRATED VALVE ARRAYS

The true advantages of this integrated valve technology are only realized when use is made of the fluid interconnect layer at the bottom of the wafer. This allows valves to be interconnected in any pattern that is allowable in two dimensions. In particular, for the applications we were pursuing, the most useful structure is a four way valve, made with four individual valves, but with inputs and outputs combined, as shown in Fig. (7). In addition to testing the individual valves that make up this array, we have tested the ability of the array to switch flow from one path to another. The setup of that experiment and its results are shown in Fig. 8. Two of the valves are constantly held closed, while 2 psi driving pressure is applied to the input located between the two other valves. As the closing pressure is gradually switched from one valve to the other, the flow also is switched. Note that in Fig. 8 the flow on/flow off ratio is 10^6; this is due in part to the improved layout seen in Fig. 7, where the sacrificial layer extends beyond the top openings. The density of the structure of Fig. 7 is about 16 valves/cm^2; the densest structures we have fabricated are a set of three four way valves in a 1 cm^2 chip, for a density of 12 valves/cm^2.

DISCUSSION AND CONCLUSION

Comparing this work to the recently reported work by Vieider *et al.*, [3] who also used silicone rubber membranes, there are several important differences. In this paper, the flow path and the movable membrane are integrated in a single die; in Ref. [3] two chips must be aligned in an external fixture. The flow path is much shorter in [3], leading to a valve with a rapid transition between open and closed states; the valves described here have a longer flow path, which lends itself more to accurate flow control.

The integration of the movable membrane in the flow chip is achieved with a sacrificial layer; one of the problems we encountered after removing that layer is, not surprisingly, stiction. In most cases, acetone or isopropanol were needed to initially unstick membranes. More work will be needed to devise methods to overcome this problem. There are many further directions into which this work can be taken. One of them is to combine these structures with other actuation methods. Another is to integrate the flow input/output manifolds, and the closing pressure manifold, so that valve chips need no further external fixtures to operate.

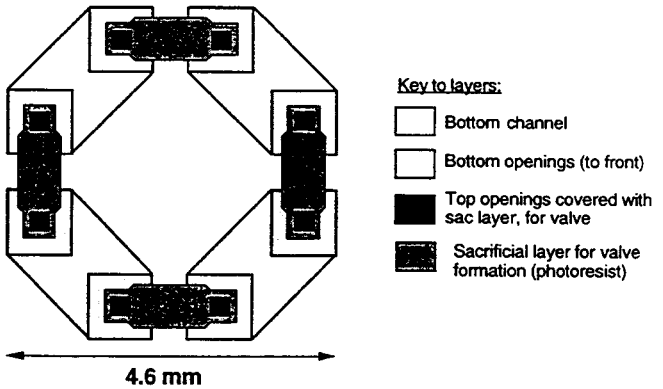


Figure 7. Layout of a symmetrical four-way valve array, made with four one-way valves. The size is 4.6 mm on a side. Other, non-symmetrical, layouts that minimize the size along one dimension have also been fabricated.

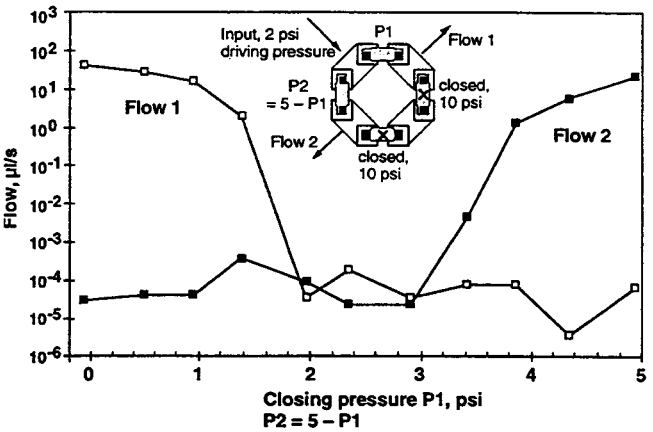


Figure 8. Test results of the valve array shown in Figure 7

REFERENCES

1. L. Bousse, "Whole cell biosensors", *Proceedings of Transducers 95*, Vol. 2, Stockholm (1995) p. 483.

2. W.K. Schomburg, B. Büstgens, J Fahrenberg, and D. Maas, "Components for microfluidic handling modules", *Proceedings of the µTAS '94 Workshop*, Kluwer (1995) p. 255.

3. C. Vieider, O. Öhman, and H. Elderstig, "A pneumatically actuated micro valve with a silicone rubber membrane for integration with fluid handling systems", *Proceedings of Transducers 95*, Vol. 2, Stockholm (1995), p. 284.

4. S. Shoiji, B. Van der Schoot, N. de Rooij, and M. Esashi, "Smallest dead volume microvalves for integrated chemical analyzing systems", *Proceedings of Transducers 91*, San Francisco (1991), p. 1052.

5. O. Tabata, K. Kawahata, S. Sugiyama, and I. Igarashi, "Mechanical property measurements of thin films using load deflection of composite rectangular membranes", *Sensors and Actuators, 20*, 135 (1989).

6. M.G. Allen, M. Mehregany, R.T. Howe, and S.D Senturia, "Microfabricated structures for the *in situ* measurement of residual stress, Young's modulus, and ultimate strain of thin films", *Appl. Phys. Lett., 51*, 241 (1987).

Present affiliation:

*: Caliper Technologies Corp., 3180 Porter Dr., Palo Alto, CA 94304.
**: Twente University, Enschede, The Netherlands.
***: Institute of Microtechnology, Univ. of Neuchatel, Switzerland.

Ultrasonically Driven Silicon Atomizer and Pump

Amit Lal and Richard M. White

University of California at Berkeley
Berkeley Sensor & Actuator Center, Dept. of Electrical Engineering and Computer Sciences
497 Cory Hall, Berkeley, California 94720-1770 USA

ABSTRACT

High amplitude ultrasound, generated by a micromachined silicon resonant transducer, was used to atomize and pump liquids. The needle-shaped device made by bonding two silicon horns, is driven by lead zirconate titanate (PZT) piezoelectric plates. Water was atomized continuously at feed rates of 2.4 ml/min into a mist with mean drop size of 25 μm at a 72 kHz resonance frequency. An isopropanol mist was ignited to produce a spray combustion flame. When the vibrating tip is immersed in fluids, streaming forces cause the fluid to be pumped at rates of 10 ml/min and generate pressure as high as 0.75 psi (5200 Pa.)

INTRODUCTION

In this paper, we present further results on using a micromachined silicon ultrasonic resonant transducer which can generate high amplitude ultrasound (ultrasonic tip velocities as high as 23 m/s = 51 miles/hour = 83 km/hour). Earlier [1,2] we used this silicon transducer to perform ultrasonic surgery on a human cataractous eye lens. While performing surgery, we noticed pumping and atomizing surrounding the tissue. A literature survey revealed that high strength titanium-alloy based ultrasonic transducers have been used before to pump and atomize liquids[13,14]. In particular, fuel atomization has been of practical significance for fuel combustion systems [12,13].

Silicon based micro-fluidic systems are predicted to have applications in chemistry, biology and medicine. Examples include drug delivery, analysis and processing. Many of these structures consist of silicon wafer sandwiches, where fluids have to be delivered and removed from micrometer size orifices. The pump presented in this paper can be made small enough to remove, deliver and atomize fluids from micro-fluidic systems (Figure 1.)

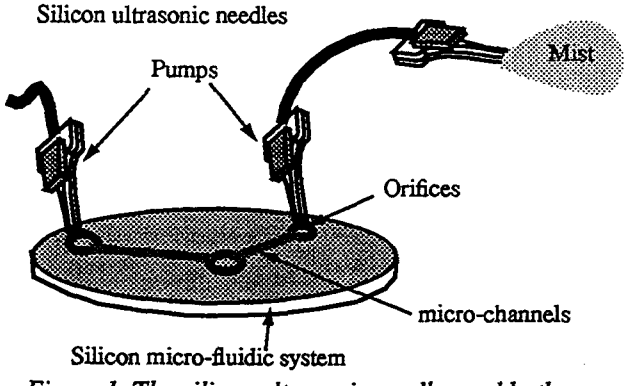


Figure 1. The silicon ultrasonic needle used both as a pump and an atomizer.

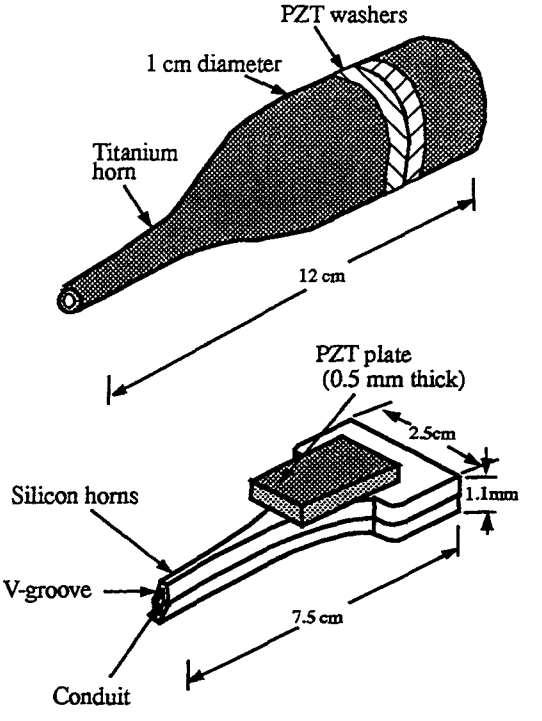


Figure 2. Conventional metal transducer and the silicon transducer

A conventional metal-based resonance transducer (Figure 2, top) is assembled by sandwiching PZT plates between a high strength titanium alloy horn amplifying section and a back reflecting mass. The PZT plates drive the structure in its longitudinal resonance mode. The resonators are designed such that some multiple number (usually 1-3) of half-wavelengths of acoustic wave fit along the length of the transducer. The horn focuses the ultrasonic energy to amplify the longitudinal motion of the tip. The PZT plates are placed at the displacement node of the resonator.

The silicon resonator tool (Figure 2, bottom) consists of two bonded silicon horn shaped plates. A channel is formed by V-grooves on each plate. A PZT plate is bonded at the nodal point of the structure on one or both sides of the device. The d_{31} coefficient

0-9640024-1-8 $3.00 © 1996 TRF

Table 1. Ultrasonic properties of selected ultrasonic materials [5,6,7,15]

Property	Y	Strength	Density	$S_m c$	Acoustic Loss	Acoustic Impedance
Material	(GPa)	(GPa)	(Kg/m^3)	(m/s)	1/Q	$Kg/m^4/s$
Silicon	190	7 [15]	2300	335	$\sim 10^{-5}$	661
Diamond	1035	53	3500	881	$\sim 10^{-5}$	1903
Si_3N_4	385	14	3100	405	$\sim 10^{-5}$	1092
SiO_2	73	8.4	2500	622	$\sim 10^{-5}$	427
SiC	700	21	3200	444	$\sim 10^{-5}$	1496
Titanium alloy (C-120-AV)	110	0.9	4430	41	$\sim 3 \times 10^{-4}$	698
Carbon steel	200	2.1	7900	53	$\sim 10^{-4}$	1256
PZT 4	85	0.024	7600	.93	$\sim 2.5 \times 10^{-3}$	803
PZT 5	67.5	0.0276	7700	1.21	$\sim 1.54 * 10^{-2}$	720

causes the PZT to stretch and contract in the longitudinal direction when the applied electric field is in the thickness direction.

Silicon transducers have several significant over their metal counterparts: The tip displacements can be higher, and the losses and costs can be lower. The maximum tip velocity or displacement in an ultrasonic resonator is proportional to $S_m c$ [4]. Here S_m is the fracture strain and c is the speed of sound in the resonator material. $S_m c$ for silicon is times higher than titanium alloy (Table 1, highlighted sections). Silicon can be driven eight time higher in amplitude than titanium alloys before failure. Since effects like pumping and atomization are due to high acoustic intensity, which is proportional to the square of the ultrasonic amplitude, silicon transducers can have 64 times higher pumping/atomization rates.

A problem experienced with the conventional metal transducers is heating at high ultrasonic amplitudes. Since silicon's dislocation density is very low (1 dislocation/cm3), its internal Q is an order of magnitude higher than metal alloys. Hence, a silicon transducer will generate much less heat and have a higher device quality factor.

The conventional metal transducers are also costly (>$3000) because each is machined serially using expensive hard-alloy machining techniques. Micromachined silicon transducer can be inexpensively batch-fabricated to tighter tolerances. Furthermore, any strength reducing surface flaws can be eliminated by oxidation to recover high fracture strength of 7 GPa [15]. Low cost, due to batch processing, will allow parallel operation of several silicon transducers for a fault tolerant system. Integrating stress, flow, and temperature sensors in the silicon device will permit operation with closed-loop control.

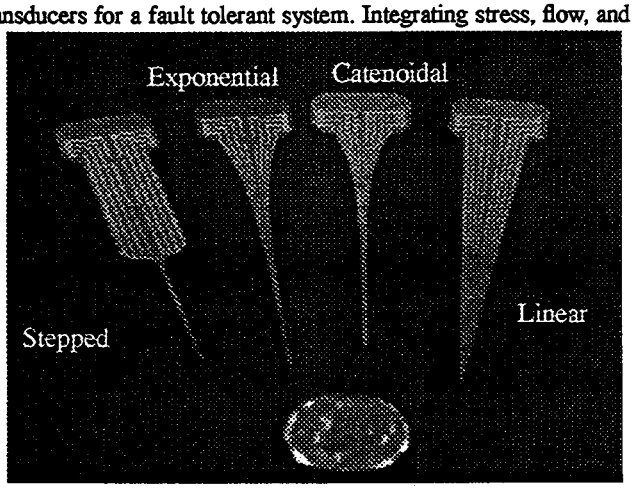


Figure 3. Silicon horns of various shapes (with a 10 cent coin)

FABRICATION

The fabrication of the silicon transducer starts with a silicon nitride (0.3 μm thick) coated four-inch silicon wafer (<100> orientation, 0.55 mm thick). Using photolithography and front-back alignment, the V-groove and the horn shapes are defined in a spun-on photoresist. A standard nitride plasma etch was used to remove nitride from exposed areas. KOH anisotropic etch is used to etch silicon from front and back simultaneously, until the horn structures are suspended in the wafer by the thin silicon nitride membrane alone. The silicon horns (Figure 3) are broken free by a gentle push as the membranes are quite weak. After piranha and HF oxide removal, these horns are thermally oxidized (0.5 μm oxide) to passivate the silicon surface. Then a 1 μm thick LPCVD deposited low-stress nitride is deposited conformably on all the horn surfaces to increase impact resistance of the final needle structure. Figure 5shows the different shapes of horns fabricated.

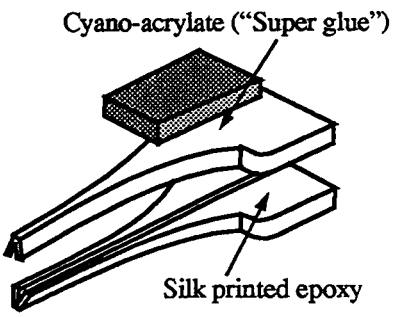


Figure 4. Assembly of the silicon ultrasonic device

To form the structure (Figure 4), a thin layer of epoxy is screen printed on a set of horns. They are then aligned together on top of each other by using a piano wire inserted in the conduit formed by the two V-grooves. The structure is then heat treated at 70° C for 1 hour for best epoxy strength. A brass grommet is also epoxy glued at the shank side of the needle structure for the tubing connection. In the future, the two silicon horns will be joined with stronger bonding techniques like fusion, anodic or silicon-gold eutectic bonding. The ultimate strength of the bond will also determine the ultimate peak displacement amplitudes achievable with the silicon resonator.

PZT-4 plate (1 mm thick) is bonded at the nodal point of the ultrasonic resonator. Both super-glue (cyano-acrylate) and epoxy were tried. The former permitted obtaining a thinner bond and bet-

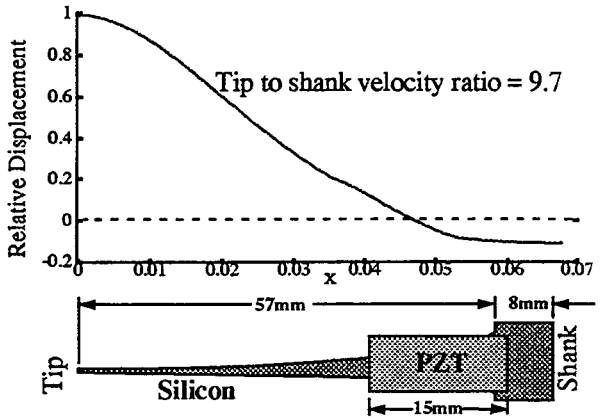


Figure 5. Numerically computed relative longitudinal displacement along the length of the silicon device.

ter acoustic coupling. After bonding, lead wires are soldered to the two faces of the PZT plate.

DEVICE MODEL

The usual balance between elastic and inertial forces on each infinitesimal section of the structure was used to model the steady state resonance of the silicon resonator. Because the thickness and width are much less than the length of the transducer, a one-dimensional longitudinal wave propagation model can be used. Webster's horn equation [2] was solved numerically in the horn section as the cross-sectional areas varies along the length. Stress-free boundary conditions were applied to the PZT/silicon composite horn. The elastic effects of the "super glue" and the epoxy were not included in the model because of their negligible thicknesses compared to those of the silicon and PZT. An amplitude enhancement factor of 10 was measured under an optical microscope. This is in good agreement with the ratio of 9.6 obtained from the numerically computed displacement profile shown in Figure 5.

MEASURED RESPONSE

Figure 6 shows the electrical impedance measured using a HP 4195 impedance analyzer. The plot reveals a predicted transverse mode at 55 kHz and the desired $\lambda/2$ mode at 72 kHz.The full wavelength resonance is 95 kHz, followed by the first PZT half-wavelength resonance at 123 kHz. The impedance of the device

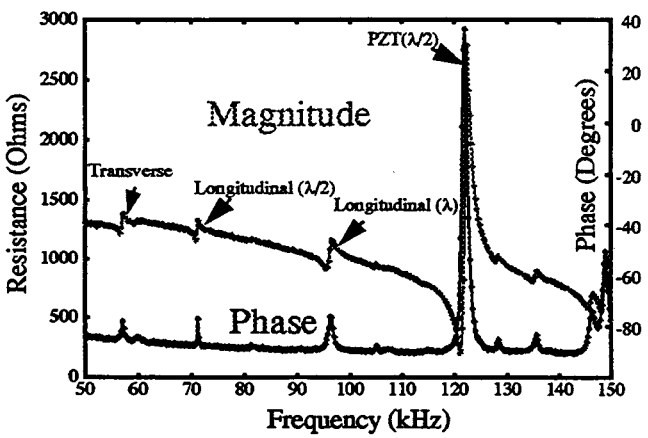


Figure 6. Electrical impedance vs. frequency

vs. frequency can be used to ascertain the quality factors, coupling coefficients, and the electrical model parameters for the device at a particular resonance [1]. The smaller in phase at resonance for the $\lambda/2$ mode as compared to PZT mode indicates that coupling between the silicon and the PZT can be further improved by better bonding methods.

The maximum longitudinal displacement measured in air, when the transducer held at the nodal point by a X-acto blade holder was 100 μm at a 200 Vpp drive at 72 kHz (The amplitude achieved was limited by the obtainable amplifier output.). This vibration amplitude implies a tip velocity of 23 m/s. In the future, we want to drive silicon to near its theoretical $S_m c$ velocity of 335 m/s by using better coupling and multiple horn amplifying sections.

ATOMIZATION

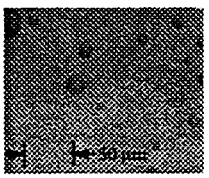


Figure 7. Droplets observed optically

The high longitudinal vibration amplitude was used to atomize and pump fluids through the channel of the device. A fluid mist cone, one to five inches in length and one inch in diameter, was generated by the silicon atomizer. This ultrasonic atomization occurs in a two-stage process. First, capillary waves form which are ejected once their amplitude becomes high enough [8]. Figure 7 shows droplets collected from the atomized mist on a hydrophobic plastic plate. Drop sizes were measured by measuring the drop size optically and multiplying by 0.70 [16]. The drop sizes measured were 20-35 μm. This range agrees with a mean diameter of 25 μm obtained from an semi-empirical expression [17].

$$D = 0.34\left(8\pi\frac{T}{\rho f^2}\right)^{\frac{1}{3}}$$

Here T is the fluid surface tension, ρ is the fluid density and f is the acoustical frequency.

APPLICATION TO FUEL INJECTION

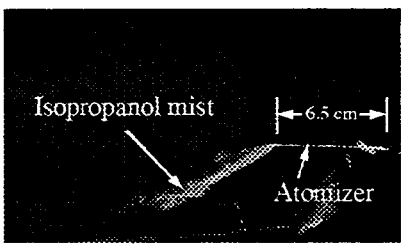


Figure 8a. Alcohol atomization

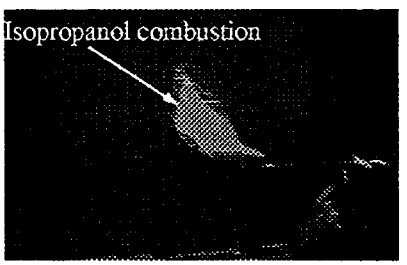


Figure 8b. Flamethrower.

The rate at which fuel burns can be greatly enhanced by combusting the fuel in drop size from 30-60 μm [8] due to optimal mixing of air and fuel possible. Furthermore, more complete combustion and hence higher fuel burning efficiencies are possible. Most conventional fuel spray atomizers accelerate the fuel into turbulence, forcing it to break into drops. This requires high power pumps and moving equipment. Ultrasonic atomization eliminates needs for these pumps and inherently produces a low-velocity mist [11] that results in less wear. In order to investigate ultrasonic fuel combustion, we atomized isopropanol into a mist with the silicon device (Figure 8a.) Then we ignited the mist to measure the rate of combustion (Figure 8b.)

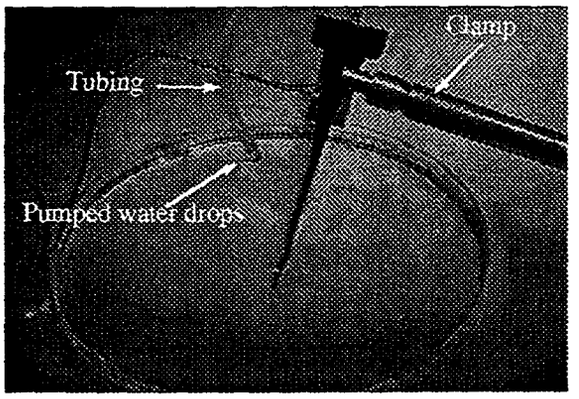


Figure 9. Evidence of pumping water against gravity

Continuous combustion at fuel feed rates of 2.4 ml/min was observed. The evidence of complete combustion was the absence of alcohol droplets leaving the flame.

PUMPING

The driven silicon device pumped water (Figure 9) when it was inserted in a petri-dish containing the water. Pumping is caused by forces generated by vortices induced by acoustic streaming (Figure 10.) The outward vortices are formed due to shear coupling at the outside walls of the device. The direction of these vortices is determined by the gradient of the motion along the horns and the stress-free boundary condition on the liquid surface. The outward flowing vortices induce a flow into the orifice. Zero-flow pump pressures as high as 0.75 psi was measured. Flow rate for a pressure differential of five inches of water was 10 ml/min.

Using acoustic streaming theory, Jackson [9,10] calculated the size and shape of streaming vortices generated by a solid bar vibrating longitudinally in a liquid. It was shown that the vortex sizes scale with the dimensions of the vibrating bar unless the dimensions approach that of the shear viscous layer. Below this dimension, streaming inside the shear viscous layer becomes important. This length is $\sqrt{\frac{2\nu}{\omega}}$.

Here ν is the kinematic viscosity of the fluid and ω is the radial frequency of operation. This layer is 2 μm thick for water at 72 kHz.Hence, this pumping mechanism will work for needle orifice diameters two or three times this length. Using micromechanical techniques, it is possible to make a silicon pump whose outer diameter is 10 μm.

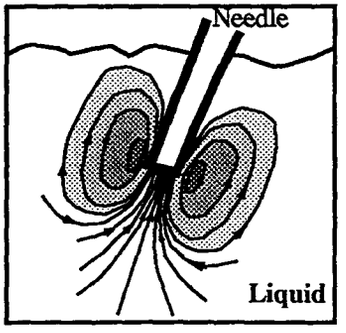


Figure 10. Vortex flow pattern at vibrating tip.

CONCLUSIONS

Silicon's material properties make it superior to titanium alloys as an ultrasonic transducer material. It's high ultimate particle velocity, $S_m c$, and low acoustic loss make it ideal for high-amplitude ultrasonic applications. Silicon devices made by micromachining techniques, creates a new regime of high amplitude miniature ultrasound actuators. Batch fabrication, tight tolerances, and integration of sensors will increase the utility of the ultrasonic device.

We have used the silicon ultrasonic resonator to atomize and pump liquids such as water and isopropanol. The atomization at 72 kHz produced drop sizes varying from 25-35 um with fluid rates as high as 2.5 ml/min. We atomized isopropanol and ignited the mist to produce a sustained butane-torch like flame. This device could be used as a fuel injector in a small aircraft combustion engine. Arrays of these devices could be used to produce higher feed rates and fault-tolerant systems. The same device has been used to pump fluids via acoustic streaming. Zero-flow pump pressure of 0.75 psi was measured. Flow of 10ml/min was obtained at a pressure differential of 5 inches of water.

In the future, we will integrate stress, displacement, temperature and flow sensors to actively control the rate of pumping and atomization. Better electrical-to-mechanical energy transduction is expected from better bonding of PZT to silicon horns, and PZT bonded on both top and bottom of device. Our goal is to develop a technology to remove, deliver and atomize fluids from micro-fluidic systems.

REFERENCES

[1] Lal, A., White, R. M., "Silicon micromachined ultrasonic microcutter," IEEE Ultrasonic Symposium 1994, Nov 1-4, Cannes, France.

[2] Lal, A., White R. M., "Silicon microfabricated horns for power ultrasonics.", Transducers 95, July 24-27, Stockholm, Sweden.

[3] Lal, A., White, R. M., "Micromachined silicon ultrasonic pump," IEEE Ultrasonics Symposium 1995, Nov. 7-10, Seattle, USA.

[4] Hunt, F.V., "Stress and strain limits on the attainable velocity in mechanical vibration," *J. Acous. Soc. Am.*, vol. 32 No. 9, pp. 1123-1128, Sept. 1960.

[5] Ashby, M. F. *Materials Selection In Mechanical Design*, Pergamon, 1992.

[6] Petersen, K.E., "Silicon as a mechanical material," *Proc. IEEE*, vol. 70, no. 5, pp. 420-457, May 1982.

[7] C. J. Smithels, Ed., *Metals Reference Handbook*, London, England: Butterworths, 1976.

[8] Williams, A., *Combustion of Liquid Fuel Sprays*, Butterworth, 1990.

[9] Jackson, F. J., Nyborg, W. L., "Sonically-induced micro-streaming near a plane boundary. I. The sonic generator and the associated acoustic field," *J. Acous. Soc. Am.*, vol. 32 No 10, pp 1243-1250, Oct. 1960.

[10] Jackson, F. J., "Sonically-induced micro-streaming near a plane boundary. II. Acoustic streaming field," *J. Acous. Soc. Am.*, vol. 32 No 11, pp 1387-1395, Nov. 1960.

[11] Singh, A., Mehregany, M., "Micromachined silicon fuel atomizers for Gas Turbine Engines," MEMS 95 proceedings, pp 473-478, San Diego, 1995.

[12] Bruetsch, R. I., Hamady, F. J., "Cold starting an alcohol-fueled engine with ultrasonic fuel atomization," EPA technical report EPA/AA/TDG/93-02a, 1993.

[13] Lierke, E. G., "Ultrasonic atomizer incorporating a self-acting liquid supply," *Ultrasonics*, October 1967, pp 214-218.

[14] Wang, W., Carter, R., "PZT ultrasonic pump," *Ultrasonics*, March 1986, pp 105-106.

[15] Ericson, F., Schweitz, Jan-Ake, "Micromechanical fracture strength of silicon", *J. Appl. Phys.*, 68(11), 1 December 1990, pp 5840-5844.

[16] Nelson, P. C., *The Shape of Small Water Drops...*, Dissertation, Univ. Calif. Berkeley, 1945.

[17] Lang, R. J., "Ultrasonic atomization of liquids," *J. Acous. Soc. Am.*, vol. 34 No. 1, pp. 6-8, Sept. 1962.

ACKNOWLEDGEMENTS

This research was supported by the Berkeley Sensor & Actuator Center (BSAC), an NSF/Industry/University Cooperative Research Center, funded by the National Science Foundation and a consortium of industrial corporations and government agencies. Research and fabrication of test structures was carried out in the Berkeley Micro-Fabrication Facility.

MICROMACHINED ACOUSTIC-WAVE LIQUID EJECTOR

Xu Zhu, Eddie Tran, Weichih Wang, Eun Sok Kim and Soo Young Lee*

Department of Electrical Engineering
University of Hawaii at Manna
Honolulu, HI 96822

*Tektronix, Inc.
Beaverton, OR 97077

ABSTRACT

This paper describes a micromachined Acoustic-Wave Liquid Ejector (AWLE) array on silicon wafer for ink jet printing and microfluidic systems. The AWLE focuses acoustic waves (generated by piezoelectric ZnO thin film) through a constructive interference without any lens. Initial results show that the AWLE ejects ink liquid of 20 μm diameter size when driven by a 260 MHz sinusoidal wave for 100 μsec.

INTRODUCTION

Affordable, high quality and reliable printers are in great demand with the booming of personal computer industry. Currently, two most dominant printing methods are laser printing and ink jet printing. Laser printing produces high quality images. However, the technological complexity in laser printing has resulted in an expensive machine. In contrast, printers using ink ejection are relatively inexpensive due to their simple printing process.

Various methods have been employed to eject ink droplets for ink-jet printers. Among them are techniques using physical vibration through piezoelectric actuation [1,2] and thermal induced actuation [3]. Non-vibration based printers use electrostatics [4], vapor bubbles [5] and focused acoustic waves [6]. Most of these printers eject ink droplets through small nozzles onto a paper to form an image. The small nozzles are difficult to construct with good uniformity, and also tend to be clogged. Ink ejection without the use of small nozzles is thus desired.

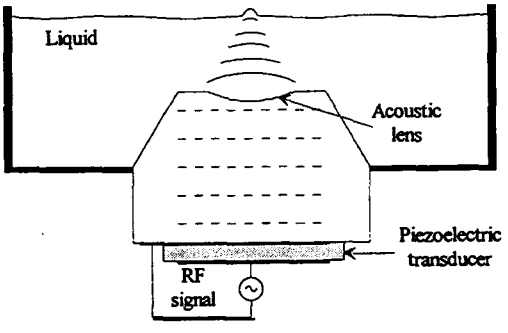


***Figure 1.** Liquid ejector using acoustic lens to focus acoustic waves.*

Nozzleless liquid ejection has been shown with a high intensity focused acoustic beam [7]. A high intensity burst of acoustic energy focused to a diffraction-limited spot at a free-liquid surface (Fig. 1) can result in droplet ejection from the surface. Liquid ejection using a focused acoustic beam is capable of ejecting liquid droplets as small as a few μm in diameter which are stable in size and directionality. This technique, when applied to ink jet printing, can provide excellent quality printing with good reliability.

The liquid ejector shown in Fig. 1 requires an acoustic lens to focus the acoustic waves. Additional processing steps required to construct any physical lens in an IC process are not desirable. Spherical acoustic lens would be especially difficult to implement using planar IC processing. Thus, we came up with a self-focusing planar transducer.

THEORY

Based on the concept of constructive interference of acoustic waves, we have developed a self-focusing liquid ejector shown in Fig. 2. When the transducer is excited with a burst of RF signal,

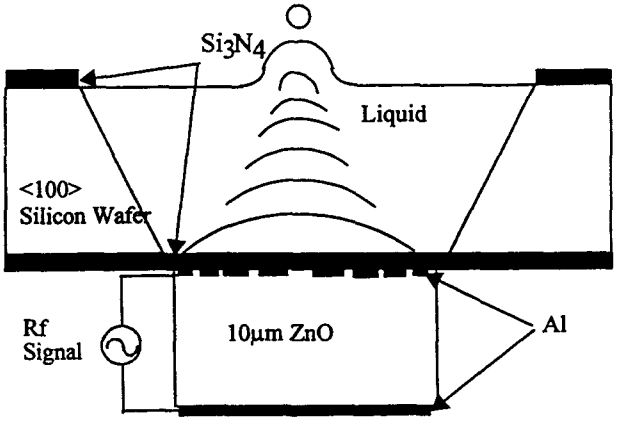


***Figure 2.** Lensless liquid ejector using a constructive interference of aocustic waves.*

it generates acoustic waves that propagate in the liquid toward the liquid/air interface. If the electrodes of the transducer are properly designed, the acoustic waves will add in-phase at the focal point.

Our lensless design borrows its concept from optical Fresnel lens which blocks certain area of light to get intensity enhancement. We make only certain areas of the piezoelectric film generate acoustic waves, which arrive at a focal point in phase. The other areas (that would have generated waves with a phase difference of π at the focal point) are designed not to generate any acoustic wave.

0-9640024-1-8 $3.00

This is what we call Half-Wave-Band sources. Figure 3 shows the acoustic wave transducer with several annular sources. The acoustic waves generated by the successive annular sources are designed to arrive at the focal point (f) with finite delays (equal to multiples of the wavelength) by ensuring that the radii r_n satisfy the following relation [7]:

$$\sqrt{r_n^2+f^2}-f=\frac{n\lambda_w}{4}. \qquad (1)$$

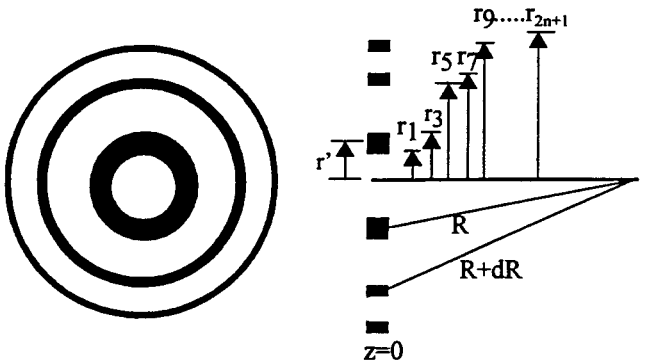


Figure 3. Focusing acoustic waves by half-wave-band sources.

Solving Eq. 1 for r_n results in $r_n=\sqrt{\frac{n\lambda_w}{2}\left(f+\frac{n\lambda_w}{8}\right)}$, where n=1,3,5,....2n+1..; λ_w is the wave length of acoustic wave in liquid.

If we use an isotropic medium, such as water, where the particle displacement u associated with the longitudinal waves can be defined as the gradient of a potential $u=\nabla\phi$, the potential ϕ at any point in the water for a half-wave-band source shown in Fig. 3 is obtained as

$$\phi(x,y,z)=-\frac{1}{2\pi}\int_s u_z(x',y',0)\frac{e^{-jkR}}{R}ds' \qquad (2)$$

where $k=\frac{2\pi}{\lambda}$; R is the distance between a half-wave-band-source and the focal point; s' is the surface area of the transducer at $z=0$; and s is the active surface area at $z>0$. The potential along the z-axis due to n numbers of half-wave-band sources can be expressed as

$$\phi_z=-u_0\left[\int_{r_1}^{r_3}\frac{e^{-jkR}}{R}r'dr'+\int_{r_5}^{r_7}\frac{e^{-jkR}}{R}r'dr'+\ldots\right]$$

$$=\frac{u_o}{jk}\left(e^{-jk\sqrt{r_3^2+z^2}}-e^{-jk\sqrt{r_1^2+z^2}}+e^{-jk\sqrt{r_7^2+z^2}}\right.$$

$$\left.-e^{-jk\sqrt{r_5^2+z^2}}+\ldots\right) \qquad (3)$$

where $u_z(r,0)=u_0$ is assumed to be uniform over the transducer and zero outside it.

In water, the pressure p is isotropic, and is equal to $p=\omega^2\rho\phi$ where ρ is the liquid density. The average intensity of an acoustic beam is

$$I=\frac{|p^2|}{2\rho V_a}=\frac{\omega^4\rho|\phi^2|}{2V_a} \qquad (4)$$

where V_a is the acoustic wave velocity in liquid. Thus the intensity along the z-axis is

$$I(z)=I_0\left|\left(e^{-jk\sqrt{r_3^2+z^2}}-e^{-jk\sqrt{r_1^2+z^2}}+e^{-jk\sqrt{r_7^2+z^2}}\right.\right.$$

$$\left.\left.-e^{-jk\sqrt{r_5^2+z^2}}+\ldots\right)\right| \qquad (5)$$

where I_0, the intensity at the transduce (i.e., at $z=0$), is equal to

$$I_0=\frac{\rho V_a\omega^2u_0^2}{2}=\frac{p_0^2}{2\rho V_a}. \qquad (6)$$

At the focal point, $\sqrt{r_1^2+z^2}=f+\frac{\lambda}{4}$; $\sqrt{r_3^2+z^2}=f+\frac{3\lambda}{4}$; ...., in which case Eq. 5 simplifies to

$$_f=I_0\left|\left[e^{-j2kf}(2nj)^2\right]\right|=4n^2I_0. \qquad (7)$$

Therefore, the half-wave-band sources effectively increase the intensity at the focal point by a factor of $4n^2$.

DEVICE STRUCTURE AND FABRICATION

We use silicon bulk micromachining to produce a 1 x 1 mm^2 membrane and a liquid cavity in the AWLE (Fig. 2). Low-stress LPCVD silicon nitride (0.5 μm thick) is used to support 10 μm thick piezoelectric ZnO film and 0.5 μm thick Al electrodes. The segmented bottom electrode and the ZnO film act as half-wave band sources as shown in the photograph of the fabricated transducer (Fig.4).

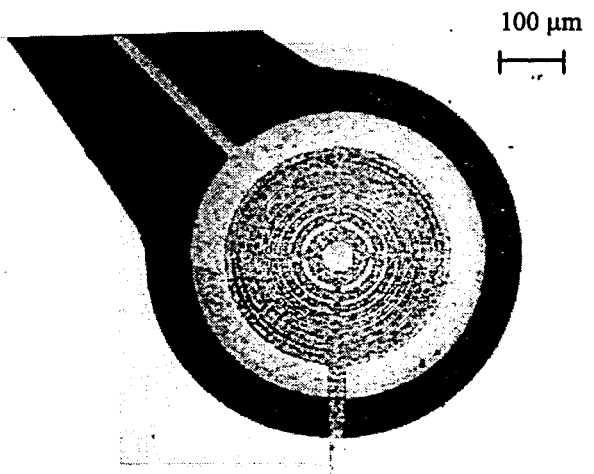


***Fig. 4** Top View photograph of lensless acoustic-wave liquid ejector. Bottom Al electrode and ZnO are patterned with annular ringsand a tailed circle, respectively. Lightly shaded square area represents the 1x1 mm² thin membrane.*

An array of transducers is fabricated on a silicon wafer with the following 4-mask process. After depositing and patterning 0.5 μm thick LPCVD low-stress silicon nitride, we remove silicon with KOH to form 1 x 1 mm^2 membranes, on which 0.3 μm thick Al is evaporated and patterned for an annular bottom electrode. Then a 10 μm thick ZnO is sputter-deposited in Ar/O_2 environment from a ZnO target, followed by 0.5 μm thick Al evaporation and

delineation for a top electrode.

RESULTS AND DISCUSSION

The fabricated devices are tested with the set-up shown in Figure 5. A continuous sinusoidal wave (in the range 100 - 400

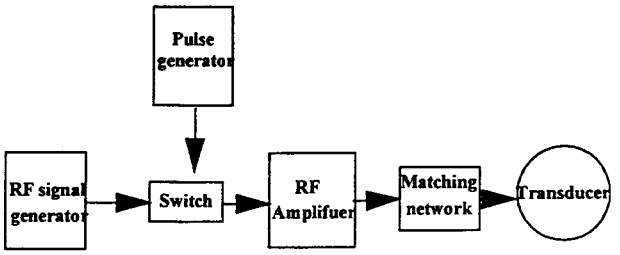


Figure 5. Schematic representation of the test apparatus for liquid ejection.

MHz) from a signal generator is fed into a high speed switch which is turned on/off with a modulating pulse duration set at, in this case, 100 μsec. The RF power amplifier then amplifies the RF burst to 10 Watts level before feeding it to the transducer.

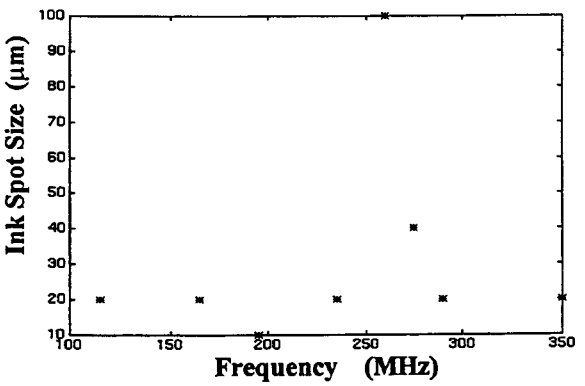


Figure 6. Experimental data on spot size of ejected ink versus driving RF frequency.

We use color methanol mixed with tin solution as the liquid for (1) low surface tension for an ejection with a relatively small power and (2) low boiling point for a relatively fast change in the liquid level so that the exact focal plane can gradually be reached without any elaborate apparatus. The ejected droplets are detected by a piece of a paper placed over the chamber area. The droplet formation is made possible by a series of high power RF pulse (with single-frequency sinusoidal waves) fed to the piezoelectric transducer which in term generates focused acoustic waves at the focal plane. Initial results show that ink spots roughly around 20 - 100 μm in diameter have been ejected around the thickness-mode resonant frequency (260 MHz) of the piezoelectric transducer (Figure 6). The largest ink spot of 100 μm is observed at the resonant frequency (260 MHz) where the output acoustic wave is at ts maximum. The smallest pulse duration for any visible liquid ejection is 100 μsec during which a sinusoidal wave ranging from 100 - 400 MHz is applied to the transducer. The power out of the power amplifier is kept constant at 10 Watts.

When the pulse duration is much longer (e.g., >100 msec), sporadic ink ejection is observed due to thermal evaporation and also large thermal-induced deflection of the membrane. The flexural deflection of the membrane occurs due to thermal expansion differences in the films employed for the transducer as the membrane temperature gets high. The temperature on the transducer is measured to be a function of the sinusoidal frequency and the pulse duration. When the transducer is driven continuously by a sinusoidal wave in a frequency range 100-400 MHz, the temperature measured above the liquid chamber over the transducer appears to peak as the frequency approaches the thickness-mode resonant frequency.

CONCLUSION

An array of self-focusing acoustic beam transducers using a 0.5 μm thick low-stress silicon-nitride membrane and a 10 μm thick piezoelectric ZnO film has been fabricated on a silicon wafer, and shown to eject liquid of 20 μm diameter size with 100 μsec application of a 260 MHz sinusoidal wave. This technique of nozzleless liquid ejection may advantageously be applied to ink jet printing and microfluidic systems.

ACKNOWLEDGEMENT

The authors wish to thank Yongli Huang of UH and Dr. Babur Hadimioglu of Zerox PARC for helpful discussion. This work was initiated by a grant from Tektronix, Inc. Partial supports were provided by Advanced Research Projects Agency under contract DABT63-95-C-0039 and NSF CAREER Award #ECS-9501698

REFERENCE

1. D.B. Bogy and F.E. Talke, "Experimental and Theoretical Study of Wave Propagation Phenomena in Drop-on-Demand Ink Jet Devices," IBM Journal of Research and Development, Vol. 28, No. 3, May 1984, pp. 314-321.
2. F. C.Lee, "PZT Printing Applications, Technologies, New Devices," Ultrasonic Symposium, 1988, pp. 693-697.
3. S. Hirata, Y. Ishii, H. Matoba and T. Inui, "An Ink-Jet Head Using Diaphragm Microactuator," MEMS Workshop conference, San Diego, CA. ,1996, pp. 418-423.
4. T. Yamada, Y. Matsuda, E. Yoshino, S. Sagae, Y.A. Ono and T. Doli, "high-Resolution Full-Color Printer by Microdot Ink-Jet Printing Method," Proceeding of the SID, Vol. 25/1, 1984, pp. 53-58
5. J. Chen and K.D. Wise, "A High-Resolution Silicon Monolithic Nozzle Array for Inkjet Printing," Transducers '95, Eurosensors IX, Stockholm, Sweden, June, 1995, pp. 321-324
6. S.A. Elrod, B. Hadimooglu, B. T. Khuri-Yakub, E.G. Rawson, E. Richley, C.F Quate, N.N. Mansour, and T.S. Lundgren, "Nozzleless Droplet Formation with FOcused Acoustic Beams," Journal Applied Physics, 65 (9), 1989, pp. 3441-34447.
7. K. Yamada and H. Shimizu, "Planar-Structure Focusing Lens for Acoustic Microscope," Journal of Acoustic Society Japan. (e) 12, 1991, pp.123-129.

Surface Micromachined Z-Axis Vibratory Rate Gyroscope

William A. Clark and Roger T. Howe
Berkeley Sensor & Actuator Center
Department of Electrical Engineering and Computer Sciences
University of California at Berkeley
Berkeley, CA 94720-1770

Roberto Horowitz
Department of Mechanical Engineering
University of California at Berkeley
Berkeley, CA 94720-1740

ABSTRACT

We have designed and tested the first polysilicon surface micromachined vibratory rate gyroscope that is sensitive to angular rotation about an axis normal to the plane of the silicon chip, the Z-axis. The Analog Devices, Inc. BiMEMS technology was used to fabricate the gyroscope, which has as a demonstrated minimum detectable signal of 1 deg/sec/$Hz^{1/2}$ with performance projected to improve to 0.1 deg/sec/$Hz^{1/2}$ in a second-generation BiMEMS design.

Our design makes use of electrostatics to improve performance considerably. Electrostatic frequency tuning of the sense mode is used to enhance sensitivity. Quadrature error due to asymmetries in the vibrating structure is nulled by detecting the resulting motion and applying electrostatic forces by modifying voltages on the interdigitated sense comb fingers to cancel error precisely.

INTRODUCTION

Multi-axis sensors are highly desirable for inertial sensing of motion in three dimensions. Integration of three axis rotation rate sensors coupled with three axis accelerometers on a single chip would result in a monolithic, six degree-of-freedom inertial measurement system capable of measuring all possible translations and orientations. The gyroscope described here is sensitive to Z-axis angular rate and complements another polysilicon surface micromachined rate gyroscope that is sensitive to rotations about the X and Y axes [1], thereby enabling single chip, three axis rate sensing.

The Z-axis gyroscope superficially resembles recently reported electrostatic comb-drive, lateral vibratory gyroscopes that sense rotations about an axis in the plane of the silicon chip[2][3][4][5]. In contrast to these other rate gyroscopes that exhibit out-of-plane deflections in response to an input Coriolis signal, the sense mode in the Z-axis gyroscope is an in-plane deflection that is orthogonal to the driven mode. Among the several benefits from this design approach are the use of differential interdigitated capacitors for both sense and force balance. As a result, a closed-loop Z-axis gyroscope can be realized in a single level of structural polysilicon.

GYROSCOPE DESIGN

Generating and Sensing Coriolis Acceleration

The operation of the Z-axis vibratory rate gyroscope is best illustrated by the spring-mass system shown in Fig. 1. As in all vibratory rate gyroscopes, a Coriolis acceleration is generated by a oscillating proof mass in a rotating reference frame. The proof mass is oscillated along the X-axis (the driven mode), the reference frame rotates about the Z-axis, and the Coriolis acceleration is detected as deflections along the Y-axis (the sense mode).

The microelectromechanical sense element of the surface micromachined Z-axis vibratory rate gyro is also shown in Fig. 1. The structure comprises three major elements: the suspension sys-

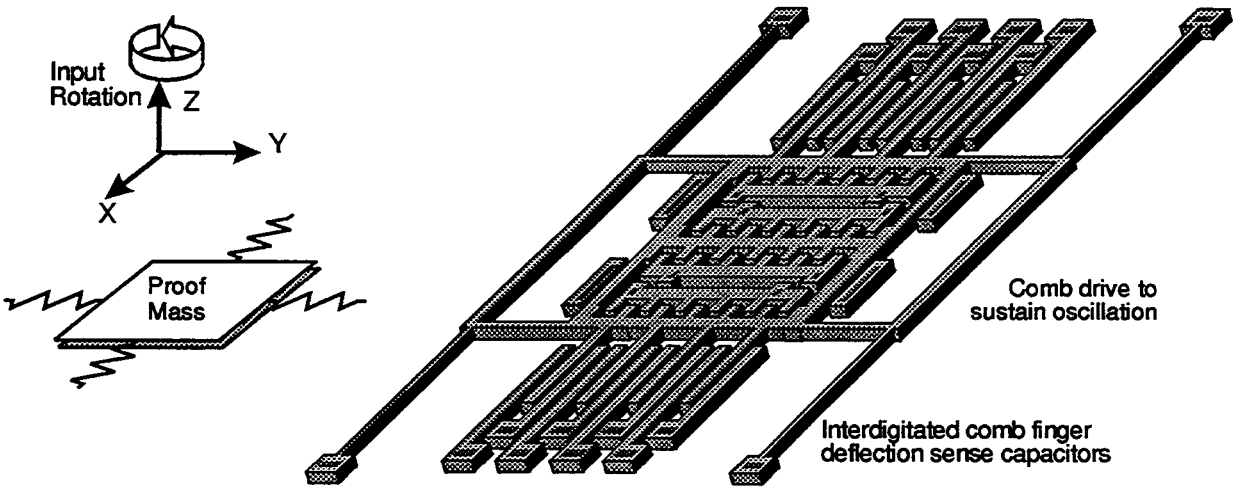


Figure 1. *The spring-mass system at left, illustrates the basic operation of the Z-axis vibratory rate gyroscope. Coriolis acceleration results from the velocity of the oscillating proof mass and the rotation rate of the reference frame. The figure on the right represents the surface micromachined Z-axis vibratory rate gyro. The proof mass is made to vibrate along the X-axis using comb drives. Any deflections that result from Coriolis acceleration are detected in the sense mode using interdigitated comb fingers*

0-9640024-1-8 $3.00 © 1996 TRF

Solid-State Sensor and Actuator Workshop
Hilton Head, South Carolina, June 2-6, 1996

tem, the comb drive used to sustain oscillation[9], and the interdigitated comb fingers used both to detect deflections along the Y-axis and to apply electrostatic force balance.

The oscillation of the gyroscope and the magnitude of the acceleration are given by Eq. 1. As a point of reference, for a gyroscope with an oscillation amplitude, $X_0 = 1$ μm, oscillation frequency, $\omega_x = 20$ kHz, and an input rotation rate, $\Omega_z = 1$ deg/sec, the Coriolis acceleration has a magnitude of 0.45 milliG's.

$$\begin{aligned} x(t) &= X_0 \sin\omega_x t \\ \ddot{y}_{Coriolis} &= 2 \cdot \Omega_z(t) \times \dot{x}(t) \\ &= 2 \cdot \Omega_z(t) \cdot X_0 \cdot \omega_x \cdot \cos\omega_x t \end{aligned} \tag{1}$$

Notice that the Coriolis acceleration is an amplitude-modulated signal where the carrier frequency is the oscillation frequency and the rotation rate modulates the amplitude. The resulting Coriolis acceleration is a dual sideband signal centered on the oscillation frequency. Since the Y-axis accelerations are proportional to velocity, the motion of the proof mass is elliptical as shown in Fig. 2.

Mechanical Design Considerations

The suspension is designed to be compliant along the X- and Y-axes and as rigid as possible to other modes. In particular, the suspension must be rigid in rotation about the Z-axis since small rotations will degrade the performance of the gyroscope and large deflections result in comb fingers crashing. The H-suspension shown in provides the necessary translational compliance while rotational rigidity is provided by thickening two trusses as shown in Fig. 1.

Quadrature Error

Now consider one effect of manufacturing flaws. Suppose that the proof oscillates along an axis that is not exactly parallel to the X-axis as shown in Fig. 2. This implies that there is some small fraction, ε, of the oscillation lies along the Y-axis. If we differentiate this displacement twice, we arrive at the acceleration due to off-axis oscillation as shown in Eq. 2. This acceleration term is referred to as quadrature error.

$$\begin{aligned} y_{Quadrature} &= -\varepsilon \cdot x(t) \\ \ddot{y}_{Quadrature} &= \varepsilon \cdot X_0 \cdot \omega_x^2 \cdot \sin\omega_x t \end{aligned} \tag{2}$$

Note the similarity between the quadrature error and the Coriolis acceleration: both are sinusoidal signals centered at the frequency of oscillation. However, they can be distinguished by their phase relative to the driven oscillation.

Quadrature errors can be quite large. The ratio of quadrature error and Coriolis acceleration is in Eq. 3. Using the previous example of an input rotation rate of 1 deg/sec and an oscillation frequency of 20 kHz, we find for the quadrature error to be as small as the Coriolis acceleration, the oscillation direction must be true to a factor of 1 part in 3.6 million. Due to manufacturing flaws and other imbalances[8], the quadrature error will be considerably larger than this and must be dealt with accordingly.

$$\frac{\ddot{y}_{Coriolis}}{\ddot{y}_{Quadrature}} = \frac{2 \cdot \Omega_z(t) \cdot X_0 \cdot \omega_x}{\varepsilon \cdot X_0 \cdot \omega_x^2} = \frac{2}{\varepsilon}\frac{\Omega_z(t)}{\omega_x} \tag{3}$$

Electrostatic Quadrature Error Nulling

Recall that quadrature error is a signal that is directly proportional to position of the proof mass. To cancel this error signal without effecting the Coriolis signal, we must apply a balancing force that is *exactly* proportional to position. This is achieved using the interdigitated position sense fingers. As the proof mass oscillates, the position sense capacitors change proportionately. A slight modification, ΔV, of the DC bias voltage, V_{DC}, applied to the comb fingers as shown in Fig. 2 results in a net force directly proportional to position of the proof mass as given by Eq. 4.

$$\begin{aligned} F_y &= \frac{C_{overlap}}{y_0} x(t)\{(V_{DC}+\Delta V)^2 - (V_{DC}-\Delta V)^2\} \\ &= \frac{2C_{overlap}}{y_0} V_{DC}\Delta V \cdot x(t) \end{aligned} \tag{4}$$

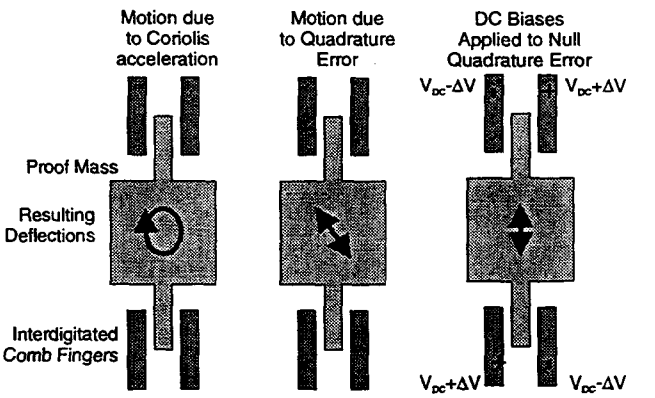


Figure 2. *The motions caused by Coriolis acceleration and quadrature error are presented above. The right most figure shows our method used to null quadrature error. We use slight changes in the DC bias on the comb fingers to generate a force that is exactly proportional to position.*

Matching Modes for Improved Sensitivity

Because the spring-mass system illustrated in Fig. 1 is a high Q second order system, the response of the proof mass to Coriolis acceleration can be enhanced. From Eq. 1, we know that Coriolis acceleration is a signal centered around the oscillation frequency. Therefore the relative values of the oscillation frequency and the Y-axis resonant frequency, ω_y, have a dramatic effect on the response of the sense mode and hence the sensor sensitivity. If we match the Y-axis resonant frequency to the oscillation frequency, we can get a gain of Q in the system response[6]. However, the bandwidth of the sensor is then limited to ω_y/Q. Given that system is run in a vacuum and Q is typically greater than 10,000, the system bandwidth for $\omega_x = \omega_y$ will be only a few Hertz.

For larger bandwidth and increased sensitivity, the Z-axis gyro is operated with a slight mismatch in resonant frequency[5]. The system response from Coriolis acceleration to Y-axis displacement is given by Eq. 5.

$$\frac{y}{\Omega_z} = \frac{2X_0\omega_x \sin\omega_x t}{\omega_y^2 + \frac{j\omega_x\omega_y}{Q} - \omega_x^2} \approx \frac{X_0}{\omega_x}\frac{\omega_x}{\Delta\omega}\sin\omega_x t \tag{5}$$

Provided that there is a means to tune the Y-axis resonant frequency, it is desirable to operate the gyroscope with a 5-10% frequency mismatch, $\Delta\omega/\omega_x$, yielding a gain of 5-10 in sensitivity.

Reducing the mismatch would increase the sensitivity but is probably difficult to maintain for any length of time.

Tuning of the Y-axis resonant frequency to improve sensitivity is accomplished using electrostatic "negative" springs[6]. When a DC voltage is established between the moving structure and the comb fingers, an attractive force is applied to the structure. Using a first-order model for the parallel-plate capacitors, the resulting linearized spring force is given by Eq. 6.

$$F_y = \left\{\frac{1}{(y_0 - y)^2} - \frac{1}{(y_0 + y)^2}\right\}\frac{y_0 C_S V_{DC}^2}{2}$$

$$k_e = \frac{2C_S}{y_0} V_{DC}^2 \qquad (6)$$

$$\omega_y = \sqrt{\frac{k_y + k_e}{M}}$$

Capacitive Deflection Sensing

Interdigitated comb fingers, as shown in Fig. 1, are used to sense deflection. The comb fingers are arranged in a capacitor bridge, any deflection of the proof mass results in measurable changes the relative size of the capacitors.

There are two common amplifier configurations used for capacitive position sensing as shown in Fig. 3. The obvious difference between these two circuits is that the integrator offers the flexibility of trading bandwidth for sensitivity by varying the integrating capacitor size. The equivalent input referred noise (measured as a deflection of a sense capacitor plate) is the same for both configurations. However, the integrator offers the advantage of lower distortion. Some of the parasitic capacitors involved in MEMS are nonlinear which will vary the gain of a buffer and thereby result in distortion. Furthermore, the practice of bootstrapping parasitic capacitance is a form of positive feedback which exacerbates amplifier distortion. Because the integrator uses a fixed, linear capacitor, distortion is kept to a minimum.

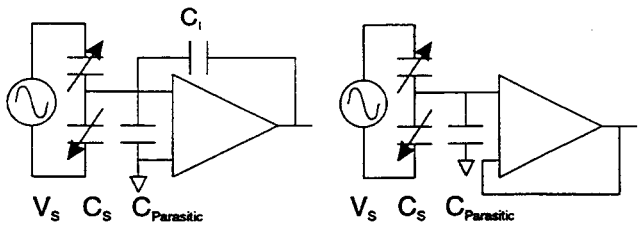


Figure 3. Capacitive position sensing using a voltage buffer and an integrator.

The integrator presents a difficult biasing problem. Ideally, biasing can be done by placing a very large resistor in parallel with the integrating capacitor as in Fig. 4. Any practical implementation of a large resistor results in considerable parasitic capacitance. A diode can also be used in parallel with the integrating capacitance, however, that adds a nonlinear capacitance to the integrating capacitance resulting in distortion. We have solved the biasing problem using subthreshold MOS devices. In subthreshold, the MOS devices exhibit extremely low transconductances and no source-to-drain capacitance. The use of subthreshold MOS has resulted in a well behaved integrator operable down to 1 kHz with a 50 F integrating capacitor without additional noise or distortion attributable to the bias circuitry.

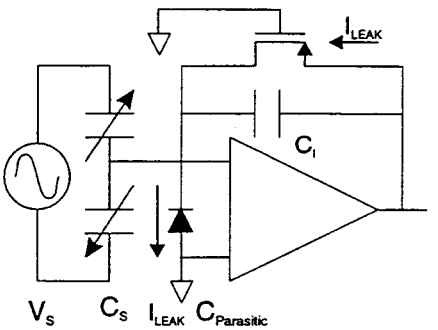


Figure 4. Integrator biasing using subthreshold MOS.

Signal Processing

At the core of the gyroscope signal conditioning is a phase-locked-loop (PLL) and several synchronous demodulators.The PLL is used to get robust, digital, signals that are in phase with the proof mass position and velocity. The phase accuracy of the PLL is critical to the operation of the gyroscope. Phase error will result in cross talk between Coriolis and quadrature signals. For the development in Eq. 7, we assume a phase error of $\theta_n(t)$.

Synchronous demodulators are used to extract the rotation rate information from the amplitude modulated Coriolis acceleration signal while rejecting the quadrature error. Signals from the PLL are mixed with measurements of the proof mass displacement as shown in Fig. 5 and Eq. 7.

$$y(t) = \left[\frac{X_0}{\omega_x}\frac{\omega_x}{\Delta\omega}\Omega\cos(\omega t) + Y_{Quadrature}\sin(\omega t)\right]$$

$$\Omega_{est} = 2\frac{\omega_x \Delta\omega}{X_0 \omega_x} y(t)\cos(\omega_x t + \theta_n(t)) \qquad (7)$$

$$= \Omega + \frac{\omega_x \Delta\omega}{X_0 \omega_x} Y_{Quadrature}\theta_n(t) + \text{High freq. terms}$$

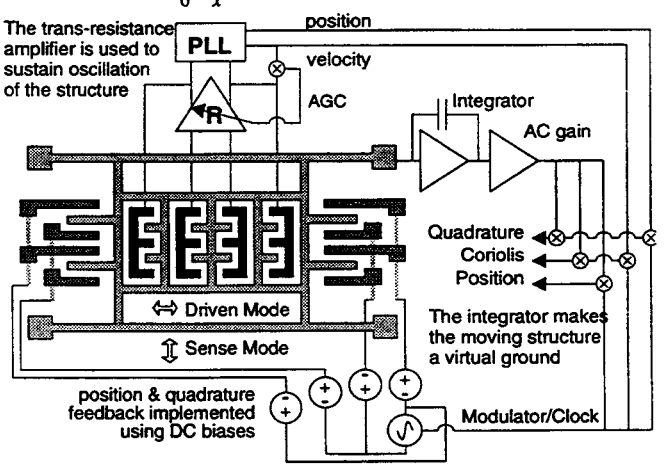


Figure 5. A block diagram of our Z-axis gyroscope and the associated signal processing elements. Shown above are the position sensing elements, the oscillation amp, and the demodulation required to operate the gyroscope.

Noise Sources

There are a number of possible noise sources. Principle among these are: Brownian noise, op-amp noise in the integrator, and PLL phase noise. Brownian noise represents the fundamental limit[10]

and is given by Eq. 8. As an example, consider a gyroscope with mass, $M = 0.2\mu g$, oscillation amplitude, $X_0 = 1\ \mu m$, a quality factor, $Q = 10{,}000$, a bandwidth, $BW = 100$ Hz, and resonant frequencies of 20 kHz. For this example, we find a Brownian noise floor, $\Omega_{nB} = 0.06$ deg/sec. Because the gyroscope is run in vacuum and has a high Q, Brownian noise is not the dominant noise source.

$$\Omega_{nB} = \sqrt{\frac{kT\omega_y}{MQ\omega_x^2}\frac{BW}{X_0^2}} \tag{8}$$

As seen in Eq. 7, phase noise in the PLL, represented by $\theta_n(t)$, can cause variable cross talk between quadrature and Coriolis signals. In our design, this effect is kept to a minimum due to the low phase noise in the PLL and nulled quadrature error.

The integrator op-amp noise is the dominant noise source. The input-referred noise of the op-amp noise is a function of the total capacitance attached to the summing node of the integrator, C_T, and is given by Eq. 9. The electronic noise is at best $\Omega_{nE} = 0.08$ deg/sec for a modest CMOS process with $f_T = 250$ MHz, a frequency mismatch, $\Delta\omega = 1$ kHz, a sense voltage, $V_S = 1$ volt, a sense capacitance, $C_S = 100$ fF, and the oscillation amplitude and finger spacing, $X_0 = y_0 = 1\ \mu m$.

$$\begin{aligned}\Omega_{nE} &= \frac{C_T}{C_s}\frac{y_0}{X_0}\frac{v_n}{V_s}\Delta\omega\sqrt{2BW} \\ &= \frac{y_0}{X_0}\frac{\Delta\omega}{V_s}\sqrt{\frac{32kTBW}{3\pi C_S f_T}} \qquad \text{Optimal for } C_{Parasitic} = 0\end{aligned} \tag{9}$$

EXPERIMENTAL RESULTS

Fig. 6 is a micrograph of the gyroscope sense element as implemented in the Analog Devices, Inc. BiMEMS process. Included on this die are the trans-resistance amplifier and the integrator. The remaining electronics was implemented off chip. The mechanical sensing element is about 1 mm across. A second, much improved, gyroscope has been implemented in the BiMEMS process.1

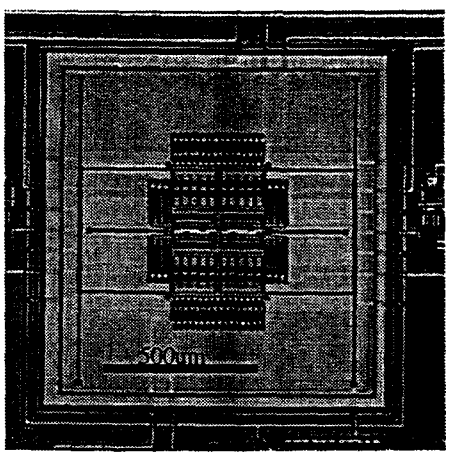


Figure 6. Die photo of Z-axis vibratory rate gyro. This structure makes extensive use of interdigitated comb fingers for both detection, excitation, and control.

Resonant Frequency Tuning

Fig. 7 plots the measured mechanical resonant frequencies as a function of the DC bias voltage applied to the position sense fingers. As expected, the Y-axis resonant frequency drops with increase bias voltage and the oscillation frequency of the proof mass remains constant at 12 kHz. Also shown are the resonant frequencies of the vertical and tipping modes that increase with DC bias. This increase is due to electrostatic levitation effects[11].

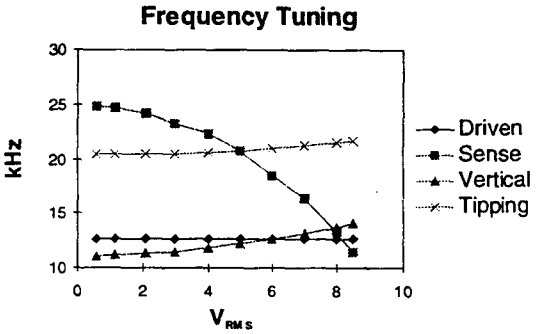


Figure 7. Frequency tuning as measured in the gyroscope. A DC bias combined with the modulation voltage establishes an RMS voltage between the comb fingers and the proof mass. The resulting electrostatic springs reduce the resonant frequency of the sense modes, raise the resonant frequencies of the out-of-plane modes, and leaves the driven mode unaffected.

Quadrature Nulling

This gyro has a means to null quadrature error. Measurements of the quadrature error and rotation rate signal in Fig. 8 demonstrate that the quadrature error signal can be controlled independently of the Coriolis signal.

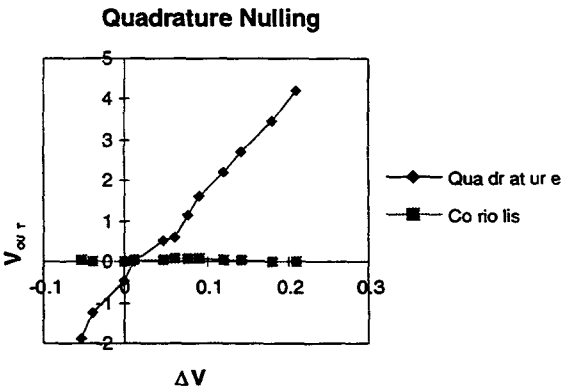


Figure 8. This figure demonstrates the effect of the quadrature control scheme shown in Fig. 2 on measured rotation rate. Both the quadrature and rotation rate signals are plotted for a zero input rotation rate as the quadrature error control level is adjusted.

Rotation Rate Sensing

An initial characterization of the gyroscope response is shown in Fig. 9, which is the measured output in response to a 1 Hz, 5 deg/sec sine wave. The gyroscope is being operated with an oscillation frequency of 12 kHz and a sense resonant frequency of about 12.5 kHz. The noise floor for this version of the Z-axis vibratory rate gyro is 1 deg/sec/$Hz^{1/2}$.

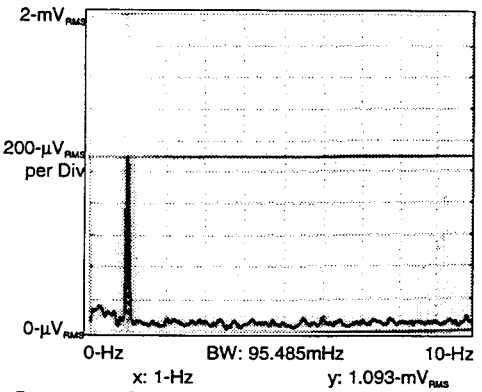


Figure 9. Response of gyro to a 1 Hz 5 deg/sec sine wave. Notice the low distortion and the noise floor that corresponds to 1 deg/sec/$Hz^{1/2}$.

CONCLUSIONS

We have designed, fabricated, and tested a polysilicon surface-micromachined Z-axis vibratory rate gyroscope that can be used in combination with other polysilicon surface micromachined X/Y gyroscopes to yield a 3 axis gyroscope on a single chip. An alternative approach to sensing Z-axis angular rate is the resonant ring gyroscope[6]. Our design takes full advantage of an existing single structural layer MEMS process and the available electrostatic forces to yield a gyroscope with performance suitable for many real world applications.

Future research includes working toward more integrated versions of this gyroscope and working with others[1] to create a three axis gyroscope and a full six degree-of-freedom inertial navigation system.

ACKNOWLEDGEMENTS

This research was supported by the DARPA funded *i*MEMS program, as a subcontract the Berkeley Sensor & Actuator Center (BSAC) from Analog Devices, Inc. The authors thank Dr. Richard S. Payne and the Advanced Accelerometer Team for design support and for access to the BiMEMS technology. The Berkeley Sensor & Actuator Center is an NSF/Industry/University Cooperative Research Center.

REFERENCES

1. T. Juneau and A.P. Pisano, "Micromachined Dual Input Axis Angular Rate Sensor," *Solid-State Sensor and Actuator Workshop*, Hilton Head, SC, June 1996.

2. J. Bernstein et al, "A Micromachined Comb-Drive Tuning Fork Gyroscope," *Digest IEEE/ASME Micro ElectroMechanical Systems (MEMS) Workshop*, Ft. Lauderdale, FL, February 1993, pp 143-148.

3. P.B. Ljung, et al, "Micromachined Two Input Axis Angular Rate Sensor," *ASME Dynamic Systems and Control Division*, San Francisco, CA, November 1995, pp 957-962.

4. K. Tanaka et al, "A Micromachined Vibrating Gyroscope," *Proceedings IEEE MicroElectroMechanical Systems 1995*, Amsterdam, Netherlands, January 1995, pp 278-81.

5. J. Choi, et al, "Silicon Resonant Angular Rate Sensor by Reactive Ion Etching," *Technical Digest of the 13th Sensor Symposium*, 1995, pp 177-180.

6. M.W. Putty and Khalil Najafi, "A Micromachined Vibrating Ring Gyroscope," *Solid-State Sensor and Actuator Workshop*, Hilton Head, SC, June 1994, pp 213-220.

7. R.S. Payne, et al, "Surface Micromachining: From Vision to Reality," *IEEE International Solid-State Circuits Conference*, San Francisco, CA, February 1995, pp 164-165.

8. Howe, R.T., "Recent Advances in Surface Micromachining," *Technical Digest of the 13th Sensor Symposium*, Japan, 1995, pp. 1-8.

9. C.T.-C. Nguyen and R.T. Howe, "CMOS Micromechanical Resonator Oscillator" *IEEE International Electronic Devices Meeting*, Washington, D.C., December 1993, pp. 127-134.

10. T.B. Gabrielson, "Mechanical-Thermal Noise in Micromachined Acoustic and Vibration Sensors," *IEEE Transactions on Electronic Devices*, vol. 40, no. 5, May 1993, pp. 903-909.

11. W.C. Tang and R.T. Howe, "Electrostatic Comb Drive Levitation and Control Method," *IEEE/ASME Journal of MicroElectroMechanical Systems*, vol. 1, no. 4 December 1992, pp. 170-178,.

SILICON BULK MICROMACHINED VIBRATORY GYROSCOPE

Tony K. Tang, Roman C. Gutierrez, Jaroslava Z. Wilcox, Christopher Stell, Vatche Vorperian, Robert Calvet, Wen J. Li, Indrani Charkaborty, Randall Bartman
Jet Propulsion Laboratory
MicroDevices Laboratory
California Institute of Technology
4800 Oak Grove, Pasasena, CA 91109-8099

William J. Kaiser
Department of Electrical Engineering
University of California, LA, CA 90024-1594

ABSTRACT

This paper reports on design, modeling, fabrication, and characterization of a novel silicon bulk micromachined vibratory rate gyroscope designed for microspacecraft applications. This new microgyroscope consists of a silicon four leaf clover structure with a post attached to the center. The whole structure is suspended by four thin silicon cantilevers. This device is electrostatically actuated and detects Coriolis induced motions of the leaves capacitively. A prototype of this microgyroscope has a rotation responstivity (scale factor) of 10.4mV/deg/sec with scale factor nonlinearity of <1%, and a minimum detectable noise equivalent rotation rate of 90 deg/hr at an integration time of 1 second. The bias stability of this microgyroscope is better than 29 deg/hr. The performance of this microgyroscope is limited by the electronic circuit noise and drift. Planned improvements in the fabrication and assembly of the microgyroscope will allow the use of Q-factor amplification to increase the sensitivity of the device by at least two to three orders of magnitude. This new vibratory microgyroscope may have applications in inertial navigation and attitude control. It offers potential advantages of almost unlimited operational life, high performance, extremely compact size, low power operation, and low cost.

INTRODUCTION

Future spacecraft and rovers will require high performance microgyroscopes that are compact and cost effective with low power consumption for attitude and maneuver control, tumble recovery of spacecraft, and stabilization and pointing of instruments such as cameras, antenna, detectors, and solar panels. Microgyroscopes with 0.1 to 10 deg/hr bias stability performance can also be used to supplement other external inertial reference systems for spacecraft such as the Global Positioning System (GPS), sun sensor, or star tracker by "filling in the gaps" during times the direct use of these systems is not possible [1]. For example, a sun sensor cannot be used during an eclipse, or during periods of several minutes between star sightings by a narrow field-of-view star tracker, or when a GPS navigation signal suffers short term dropouts due to inadequate antenna coverage or physical or electromagnetic interferences. These new microgyroscopes not only will be relevant for various extraterrestrial missions, such as planetary and interplanetary astrophysics research but will also have terrestrial applications in automotive navigation systems, autonomous control of underwater and land vehicles, platform stabilization, and smart guidance for munitions, missiles, and robotics.

This paper reports on the development of a novel silicon bulk micromachined vibratory gyroscope that is under development for microspacecraft attitude and navigation control applications. This new microgyroscope offers several features, including: 1) the use of simple bulk silicon technology which provides precise structures, increases yield, and reduces cost; 2) high rotation rate sensitivity due to the low frequency resonator design, differential mode of operation, and lock-in amplification detection; 3) low loss resonator to provide high quality (Q) factor for mechanical amplification. and 4) symmetric resonator design with degenerate modes. The design, fabrication, modeling, experimental characterization, and packaging of the microgyroscope for spaceflight experiment will be described in the next sections, followed by a conclusion.

DESIGN

In a vibratory gyroscope, the rotation sensing mechanical element is initially driven to oscillate in one mode (called input or drive mode [2,3,4]. Under rotation, Coriolis acceleration induces energy transfer from the input mode to another oscillatory mode (output or sense mode) causing the sensing element to precess. The best performance for the vibratory gyroscope is obtained when the two oscillatory modes (drive and sense) have the same resonant frequency with high quality (Q) factor. The response to the Coriolis acceleration is then amplified by the Q factor of the resonance resulting in improved sensor performance.

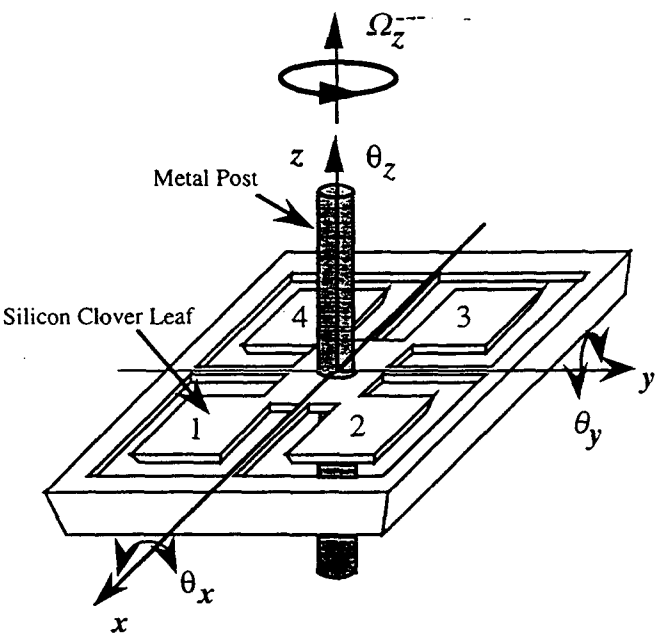


Fig. 1 Schematic of the clover-leaf microgyroscope.

The JPL/UCLA microgyroscope resembles a four leaf clover suspended by four thin wires from the gyroscope housing. The rim, the clover structure and the suspensions are made from single crystal silicon (Fig. 1). A metal post is rigidly attached through the center of the structure in the direction perpendicular to the plane of the clover leaves. The four suspensions provide mechanical support and restoring force for the harmonic motion of the whole structure. The clover leaf structure and the metal post have a combined mass of about 10^{-5} kg. The mechanical resonant frequencies of the microgyroscope vary according to design in the 300 to 500 Hz range. The low resonant frequency is due to the large mass of the post and soft suspensions. Since the response to the gyroscope is inversely proportional to the resonant frequency, low resonant frequency increases the device responsitivity.

The silicon clover-leaf structure with the post is bonded to a quartz base plate patterned with gold electrodes. The clover leaves provide large areas for electrostatic driving and capacitive sensing. The resonator is excited by applying an AC voltage to the capacitors associated with leaves 2 and 3, and this excites the rotation θ_x about the x axis. Since the resonator is symmetrical, θ_x and θ_y are degenerate with rocking-like displacement for the clover leaves. The degeneracy of the drive and sense modes allows Q-factor amplification of modal displacements. Q-factor amplification of the drive and sense motion reduces drive voltage and increases the sensitivity of the device respectively. The Q-factor of this symmetric resonant structure can be very high since it can be perfectly balanced with a fixed center of gravity, and the sum of forces and moments resulting from the vibration modes can be zero.

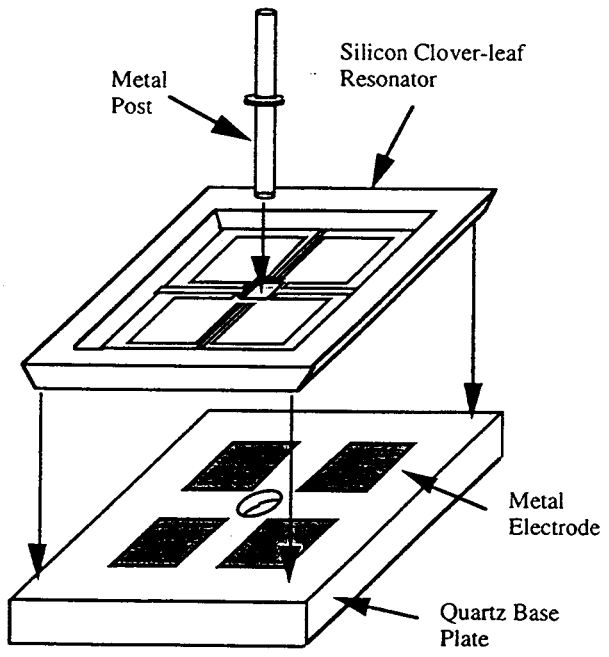


Fig. 2 Exploded view of the clover-leaf microgyroscope.

Since the post is rigidly attached to the clover-leaf structure, the movements of the clover leaves translate to movements of the post. When the leaves oscillate in the drive mode, the displacement of the post is near parallel to the clover leaf surface in the y-direction. When a rotation rate is applied about the z-axis, Coriolis force acts on the oscillating post causing it to displace in the orthogonal direction (i.e. x-direction). This post displacement is translated back into the rocking motion θ_y of the clover leaves. The post provides a large Coriolis coupling which transfers energy between the two orthogonal rocking modes.

FABRICATION

The structure of the silicon microgyroscope is shown in Fig.2. The microgyroscope consists of a quartz base plate patterned with metal electrodes, and the silicon clover-leaf resonator suspended from the thick silicon rim. The quartz base plate and the silicon rim are bonded together by using low shrinkage UV curable epoxy. The dimension of the silicon clover-leaf resonator is 7 mm x 7 mm, the dimension of each clover leaf is 1.1 mm x 1.1 mm, and the post is 500 microns in diameter and 5 mm in length. The quartz base plate is a square 1 cm x 1 cm and approximately 400 microns thick. The sensing capacitors of the microgyroscope are formed between Cr/Au electrodes on the quartz base plate and the silicon clover leaves. The gap between the base electrodes and the clover leaves is about 10 to 20 microns.

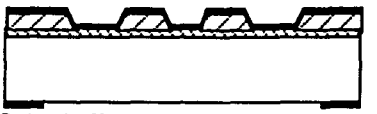
The fabrication of the silicon clover-leaf resonator requires only two masking and etching steps using bulk micromachining technology. The fabrication process sequence for the microgyroscope is shown in Fig.3a. Fabrication begins with the standard n-type <100> double side polished silicon wafer that has two silicon epilayers grown on top. The bottom layer is a 4 micron p+ (10^{20} boron doped) etch stop layer. The top layer is a 26 micron n-type silicon epilayer. A 1000Å layer of thermal silicon dioxide is grown on the wafer and patterned with the clover leaf pattern. This patterned oxide layer is used as a mask for EDP etching of the 26 micron silicon epilayer to the p+ etch-stop layer. All features of the clover-leaf pattern run along the crystal planes, resulting in very smooth shapes. This is due to the selectivity of the anisotropic EDP etch which automatically provides very precise dimensions and shape for the suspensions. This is very important for the microgyroscope performance since the mode shape and resonant frequencies are affected by any device asymmetries.

An SEM picture of the etched profiles of the clover leaf structure is shown in Fig. 4. After the wet etch the thermal oxide mask is stripped away in buffered HF. This is followed by a 3000Å thick thermal silicon dioxide. The silicon dioxide on back of the microgyroscope is then patterned and the device is again etched in EDP. This etch removes almost all the silicon substrate and stops at the p+ etch-stop layer. Next, the wafer is cleaved into individual devices and separated. The cleaved devices are then etched in 1:3:8 HNA (H_2O_2: HNO_3: CH_3COOH) until the p+ stop-etch layer is removed. Finally, the oxide layer is striped away in buffered HF. The removal of both the p+ layer and the silicon dioxide layers removes stress gradient which can cause warpage of the clover leaf structure.

The fabrication process of the quartz base plate consists of two lithography and etching steps as shown in Fig.3b. A hole is first drilled into the quartz base plates for the post. This is followed by the patterning and etching of the quartz base to the depth of 10 to 20 microns using HF. This provides the gap spacing that separates the silicon clover leaves and the electrodes on the quartz plate. Next, a Cr/Au metallization is evaporated onto the whole quartz plate, and is etched to give the electrodes patterns. Then, the quartz plate is diced into individual devices. The silicon resonator and the quartz base plate are then bonded together using UV curable epoxy. All devices are first tested for resonant frequencies and mode shapes.

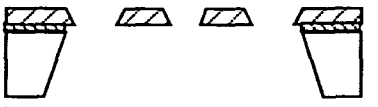
After these initial tests, a metal post is epoxied into the hole on the silicon resonator (Fig. 5).

1) 1st thermal oxide growth and patterning, topside silicon epilayer EDP etch.



2) Striped off 1st thermal oxide, 2nd oxide growth and patterning.

3) Backside EDP etch.



4) Backside HNA etch and oxide etch.

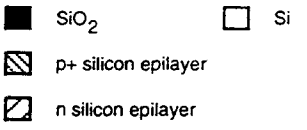
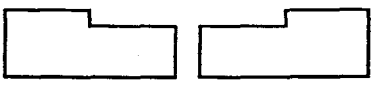
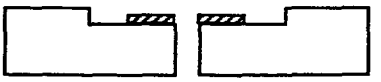


Fig. 3a Fabrication process of the silicon clover-leaf structure.



1) HF etching of the quartz base plate.



2) Metallization of the quartz base plate.

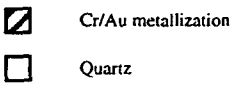


Fig. 3b Fabrication process of the quartz base plate.

CONTROL AND READOUT ELECTRONICS

In order to design the electronics for actuating and sensing the microgyroscope, an electrical model of the microgyroscope is derived and analyzed. The model described in this section gives a complete description of the frequency response of the microgyroscope. Important parasitic effects, such as interplate capacitances and mechanical coupling between the orthogonal modes under no rotation, are included in the model. Fig. 6 shows the basic electrical model which gives the excitation-to-sense transfer function, $\hat{v}_2/\hat{v}_1$ or $\hat{v}_3/\hat{v}_4$, of the microgyroscope taking one pair of opposite plates at a time. The mechanical resonance is represented by the series $L_\mu C_\mu R_\mu$ resonant circuit which is excited via the dependent sources $\delta\hat{v}_1$ and $\delta\hat{v}_2$ which communicates with the port voltage $\hat{v}_1$ and $\hat{v}_2$ respectively (δ represents the deflection/volt in the plates at zero frequency).

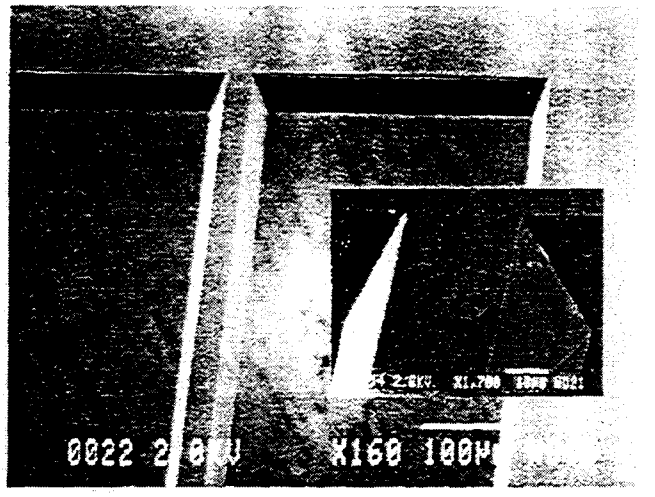


Fig. 4 SEM photograph of the corner between the suspension and the center of the clover-leaf structure demonstrating the sharpness and precision of the etching process.

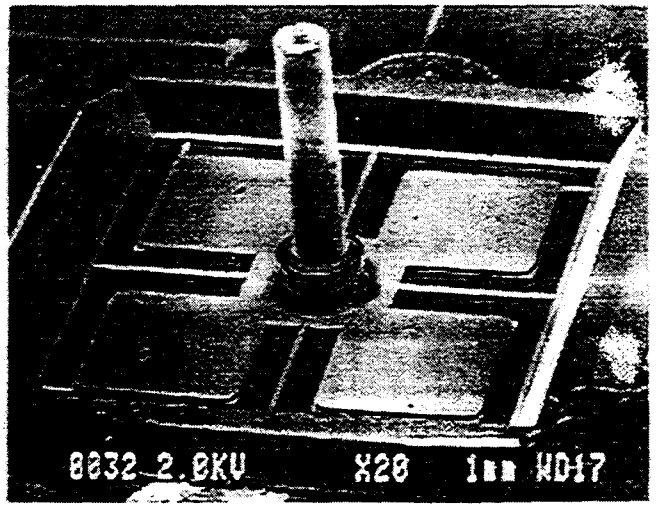


Fig. 5 SEM photograph of an assembled microgyroscope

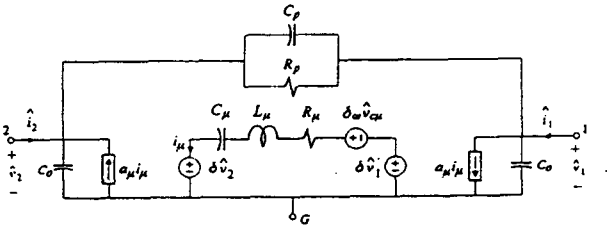


Fig. 6 Equivalent circuit model for an opposite pair of plates.

The effect of the mechanical motion on the electrical ports in turn is given by the dependent sources $a_\mu i_\mu$. The interplate

parasitic resistance and capacitance are given by R_p and C_p respectively. The dependent source $\delta_\omega \hat{v}_{c\mu}$ is the Coriolis induced term which is further explained in the next paragraph. The transfer function predicted by this model (with ω=0 and, hence, δ_ω=0) is shown in Fig.7 in which the notch at ω_n is mainly due to the parasitic capacitance C_p.

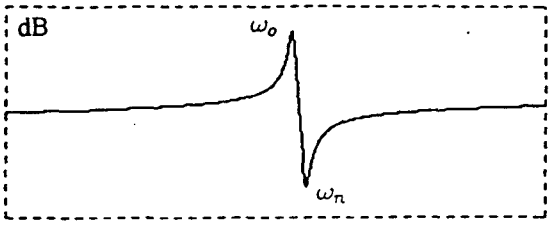


Fig. 7 Excitation-to-output transfer function obtained from the model in Fig. 6.

Fig. 8 shows the experimental transfer functions of each pair of opposite plates of a microgyroscope whose orthogonal resonant modes (drive and sense) are at 445 Hz and 435 Hz respectively. The departure of each of these transfer functions from the one shown in Fig. 7 for a single pair of plates is primarily caused by parasitic electrical and mechanical coupling between the orthogonal modes when the gyro is at rest and by the slight difference between the δ and a_μ for each pair of plates. These curves can now be easily predicted when a full model of the gyro is constructed by assembling together the model for each pair of plates given in Fig. 6 as shown in Fig. 9. The parasitic mechanical coupling between the drive and sense modes when ω=0 is modeled by $\delta_{ds}\hat{v}_{\mu d}$ and $\delta_{sd}\hat{v}_{\mu s}$ while the Coriolis term is given by $\delta_\omega \hat{v}_{\mu d}$. The predictions of this model (with $\omega = \delta_\omega$=0) are shown in Fig. 10 and are seen to match well with those in Fig. 8. Detailed analytical expressions derived from the circuit in Fig. 9 will be given in a future paper.

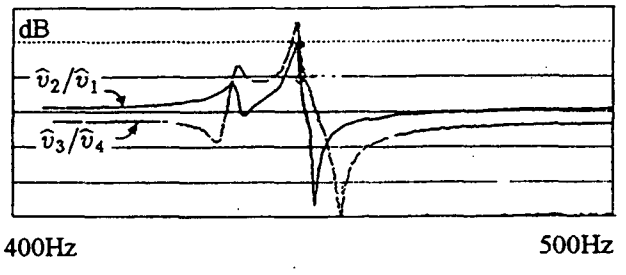


Fig. 8 Experimental plots for the excitation-to-output transfer function.

The actuation, or drive, circuit is realized by designing an oscillator (H(s)) around the microgyroscope which locks in onto the drive resonance mode as shown in Fig. 11. The purpose of summing the signals from both sense plates is to remove the differential signal between them and hence the response of the sense resonance from the feedback loop. The sense circuit, on the other hand, subtracts the signals from both sense plates in order to remove the common-mode drive signal. A lock-in amplifier is used in order to detect a differential signal induced due to rotation.

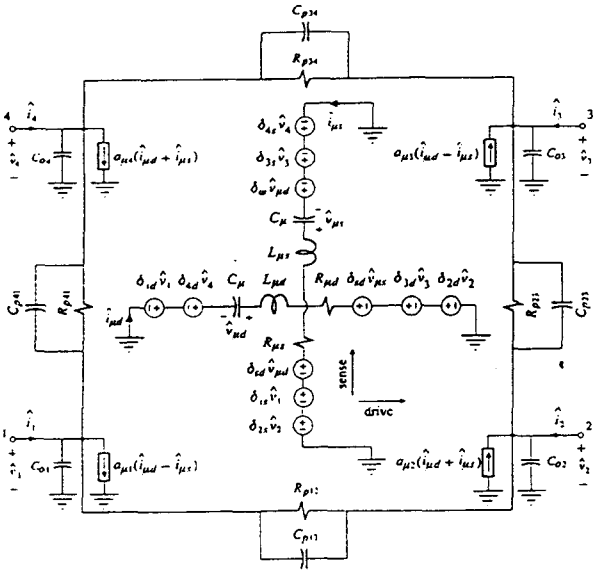


Fig. 9 Complete electrical model of the microgyroscope.

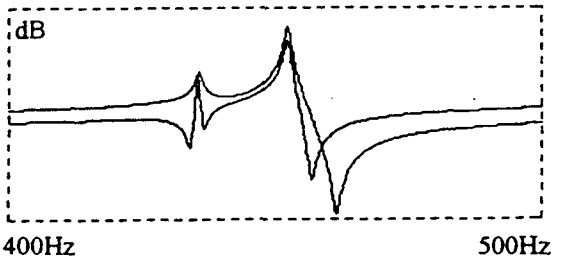


Fig. 10 Transfer function obtained from model in Fig. 9.

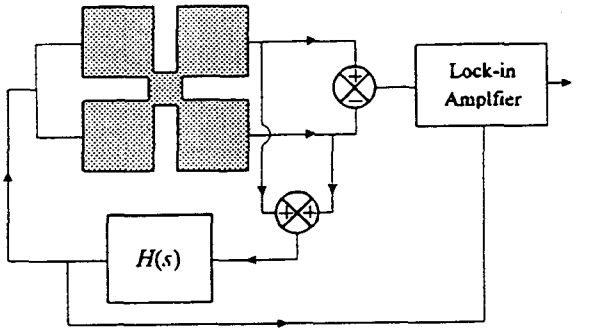


Fig. 11 Exciting and sensing of the microgyroscope.

MODELING

The Lagrange equations of motion are written for the angular displacements θ_x, θ_y, and θ_z measured in the coordinate system that rotates with the microgyroscope housing. The center O is assumed stationary, θ_z is rotation around the post (z) axis, and x and y are the spring axes. Thus, in this system, θ_x and θ_y represent measurable, near-vertical angular displacements. For spin about the z axis, θ_z remains decoupled from θ_x and θ_y, and for small θ_x, θ_y, and Ω_z/ω, the cloverleaf

displacements for a quad-symmetric structure driven by the application of the torque T to the leaves 2 and 3, are described by the following approximate equations of motion:

$$2I_{xx}\ddot{\theta}_x - 4I_x\Omega_z\dot{\theta}_y + \kappa_x\theta_x + \alpha_x\dot{\theta}_x \approx T(t) \qquad (1)$$

$$2I_{yy}\ddot{\theta}_y - 4I_z\Omega_z\dot{\theta}_x + \kappa_y\theta_y + \alpha_y\dot{\theta}_y \approx 0 \qquad (2)$$

where the terms proportional to α's and time derivatives $\dot{\theta}_x$ and $\dot{\theta}_y$ describe vibration damping, and T(t) is the driving torque. The "mass moment" is $I_z \equiv \sum_i z_i^2 (m_i/2)$ where z_i is the z component of $\vec{r}_i$ from O to any mass element of the moving structure. The other moments of inertia are :

$$I_{xx} \equiv I_z + I_y = \sum_i \left(z_i^2 + y_i^2\right)\left(m_i/2\right) \qquad (3)$$

$$I_{yy} \equiv I_z + I_x = \sum_i \left(z_i^2 + x_i^2\right)\left(m_i/2\right) \qquad (4)$$

For the structure with the post, M5 >> (M1, M2, M3, M4), and $I_{xx} \equiv I_{yy} \equiv I_z \equiv (M_5/2)L_5^2$, where L_5 is the arm length from O to the mass center of $M_5/2$ and $M_{1\text{-}4}$ are the masses of the clover leaves and M_5 is the mass of the post. The restoring torque parameters κ_x and κ_y relate to "spring" constants for linear displacements due to tension and twist, and to the device dimensions.

The homogeneous solutions to Eq. 1 and 2 are the normal modes. For small damping in the absence of platform rotation, the modes oscillate with the resonant frequencies $\omega_x \equiv \sqrt{\kappa_x/2I_{xx}}$ and $\omega_y \equiv \sqrt{\kappa_y/2I_{yy}}$ for θ_x and θ_y rotations, respectively. The application of AC voltage to the capacitors associated with leaves 2 and 3 (Fig. 1) drives the vibrational rotation about the x-axis (angular displacement θ_x). Upon application of input rotation Ω_z, the vibrational energy is transferred from the θ_x to the θ_y vibrational mode. We consider $T(t) = T_o \cos(\omega t)$ that drives θ_x, at $\omega >> \Omega_z$. Since $\theta_x >> \theta_y$, by neglecting the Coriolis term in Eq. 1, the driven θ_x is:

$$\theta_x \approx \theta_{xo}\sin\left(\omega t + \psi_x\right) \qquad (5)$$

where the amplitude

$$\theta_{xo} \approx \frac{T_o}{2I_{xx}} \frac{1}{\sqrt{\left(\omega^2 - \omega_x^2\right)^2 + \frac{\omega^2\omega_x^2}{Q_x^2}}} \rightarrow \frac{T_o}{2I_z}\frac{Q_x}{\omega_x^2} \rightarrow \frac{T_o}{\kappa_x}Q_x \qquad (6)$$

where the last two expressions on the right hand side apply for ω near ω_x, i.e., $|\omega - \omega_x| < \gamma_x$, $\gamma_x \equiv \alpha_x/2I_{xx} \equiv \omega_x/Q_x$. At times $t >> 1/\gamma_x$, after the application of Ω_z, and the transient effects have leveled-off, the spin-induced θ_y is given by the inhomogeneous solution of Eq.2:

$$\theta_y \approx \theta_{yo}\cos\left(\omega t + \psi_y\right) \qquad (7)$$

where

$$\theta_{yo} \approx \theta_{xo}\frac{I_z}{I_{yy}}\frac{2\Omega_z}{\omega_y}\frac{\omega\omega_y}{\sqrt{\left(\omega^2 - \omega_y^2\right)^2 + \frac{\omega^2\omega_y^2}{Q_y^2}}} \rightarrow$$

$$\rightarrow \frac{4I_{yy}\Omega_z\omega\theta_{xo}}{\kappa_y}Q_y \rightarrow \theta_{xo}\frac{2\Omega_z}{\omega_y}Q_y \qquad (8)$$

where the last two expressions on the right hand side are valid for a heavy post ($I_{yy} \equiv I_z$, and ω near ω_y)and Q_y is the Q-factor for the θ_y resonance. These results show that the degeneracy of the resonant frequency can enhance the spin-induced θ_y by the product of both the driven and sensing Q-factors. This degeneracy can be accomplished with a symmetric configuration of the cloverleaf and springs resonator design.

MEASUREMENT

In order to fully analyze a micromachined vibratory gyroscope, it is necessary to carry out a significant amount of testing. In this case, two separate test stations were assembled; the first test station was designed to be used to quantify the mechanical properties of the gyroscope, and the second to measure the performance of our device as a rotation sensor.

The first test station consists of a small vacuum chamber with electrical probes to electrically excite the microgyroscope, and a Polytech OFV 501 laser vibrometer (Doppler velocimeter) to accurately measure mechanical motion. Some of the mechanical properties of the gyroscopes measured with this setup are: resonance frequencies, Q-factors, mode shapes, hysteresis, snap-down voltage, symmetry (by exciting different electrodes), operational voltage, and survivability (by driving it with a very large amplitude signal).

The second test station uses a Triotech S347B rate table with a resolution of 0.01 degrees per second to rotate a micro-gyroscope inside a vacuum chamber. The motion of the microgyroscope is measured both electrically and with a Polytech OFV 501 laser vibrometer. Some of the performance characteristics measured using this apparatus are: the drive and up/down mode frequencies, the sense mode residual mechanical excitation (zero for a perfectly balanced gyro), the rotation responsitivity (mV/deg/sec), vibrational amplitude, linearity of the response, the noise characteristics of the gyroscope (Green chart, PSD, autocorrelation), bias offset (output of the lock-in without rotation), bias drift, response to rotation when sense axis is perpendicular to rotation, and response to rotation when excited in the up/down mode.

Several finished microgyroscopes were tested inside the test stations. The resonant frequency of the drive mode is at 452Hz, and the sense mode is at 447Hz for one of the devices. The Q-factor of the drive and sense mode is approximately 500. Both resonant modes are rocking modes with nodal lines along the spring directions. The separation between the resonant frequencies is primarily due to stress from the bonding of the silicon clover leaf structure to the quartz structure and to fabrication errors. The microgyroscope was driven at the drive resonance for rotation testing. The response of this device does not have significant Q-factor amplification of the sense mode displacement due to splitting of the drive and sense resonant frequency. The driving voltage of the microgyroscope is approximately 4V peak to peak. The snap-down voltage is approximately 20V. A plot of the output voltage vs. rotation rate is shown in Fig.12. From this figure, the measured responsitivity (scale factor) of the gyroscope is 10.4mV/(deg/sec) with a scale factor nonlinearity of <1% at an integration time of 1 second.

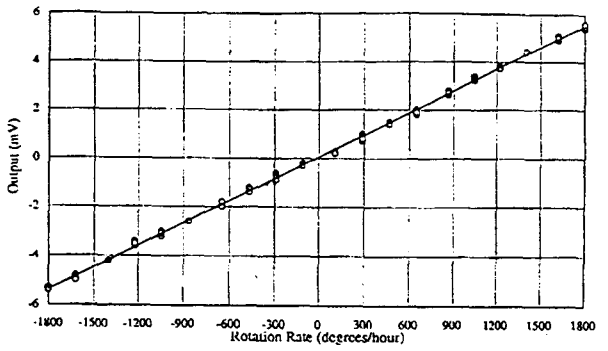


Fig. 12 Rotation rate responstivity of the clover-leaf micro-gyroscope.

The Allan deviation chart (Fig. 13) shows the lowest detectable noise equivalent rotation rate is about 90 deg/hr at an integration time of 1 seconds. The rise in the noise after the 15 second integration time is mostly due to drift in the electronics. Thus, the bias stability of this microgyroscope should be below 29 deg/hr. The performance of this micro-gyroscope is limited by the electronic circuit. Closed-loop circuitry is being developed to improve the sensitivity of this microgyroscope by lowering the noise and reducing the drift. Improvements in the fabrication and assembly of the microgyroscope will allow the use of Q-factor amplification to increase the sensitivity of the device by at least two to three orders of magnitude. This will allow the clover-leaf microgyroscope to be useful in attitude and navigation control applications.

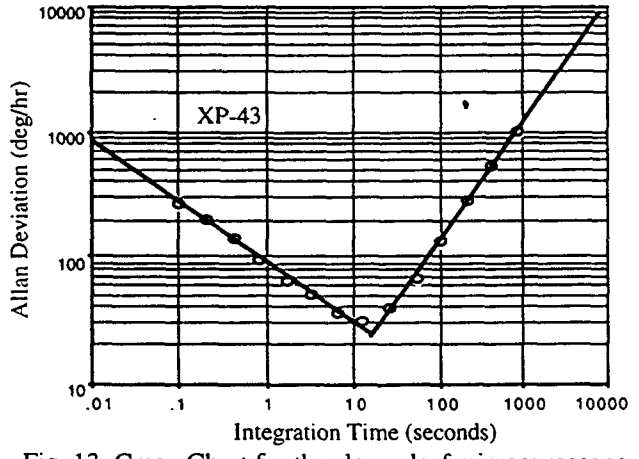


Fig. 13 Green Chart for the clover-leaf microgyroscope.

CONCLUSION AND FUTURE PLAN

This paper has reported on the design, modeling, fabrication, and characterization of a novel silicon bulk micromachined vibratory gyroscope designed for micro-spacecraft applications. A microgyroscope prototype that has been fabricated so far has a rotation responsitivity (scale factor) of 10.4mV/deg/sec with < 1% nonlinearity at an integration time of 1 second. The lowest detectable noise equivalent rotation rate is 90 deg/hr at an integration time of 1 second. The bias stability of this microgyroscope is at or better than 29 deg/hr. The performance of this microgyroscope is limited by the electronic circuit noise and drift, fabrication errors, and stress resulting from the assembly process. Improvements in these areas will further increase the sensitivity and improve the drift performance of this microgyroscope.

This JPL/UCLA clover-leaf microgyroscope is planned as a piggyback spaceflight experiment called the Commercial Hardware Micro-Gyro (CHμG) on the Mars Pathfinder spacecraft to study re-entry dynamics. Each microgyroscope is individually packaged in 2.5 x 3 x 0.4 inches stainless steel case. Four such packages are assembled onto the mother board which consists of a thermal plate and the digital electronics for communication with the spacecraft. The participation of the JPL/UCLA clover-leaf microgyroscopes in the CHuG space-flight experiment has provided valuable experiences in the integration, packaging, and space qualification of Micro-machined ElectroMechanical System (MEMS) microdevices for spaceflight missions.

The immediate goals for the JPL/UCLA microgyro effort in the next few years include 1) improve the microgyroscope design to achieve 1-10 deg/hr bias stability performance, 2) produce a three-axes Micro-Inertial Measurement Unit (MIMU) consisting of hermetically sealed microgyroscopes packaged with micro-accelerometers for spaceflight validations in spacecraft, 3) investigate the issues relevant to the space qualification of the microgyros, such as device lifetime, device packaging, and survivability during launch and in the space environment.

In conclusion, we presented a new vibratory micro-gyroscope that offers potential advantages of almost unlimited operational life, high performance, extremely compact size, low power operation, and low cost for applications in inertial navigation and attitude control of microspacecraft.

ACKNOWLEDGMENTS

The authors thank Dr. Linda Miller for encouragement and support. We give special thanks to Dr. Paul Maker and Richard Müller for the fabrication of e-beam photolithographic masks. The work described in this paper was performed by the Center for Space Microelectronics Technology, Jet Propulsion Laboratory, California Institute of Technology, and was sponsored by National Aeronautics and Space Administration, Office of Space Access and Techology.

REFERENCE

[1] G. E. Sevaston, L. Craymer, and W. Breckenridge, "GPS Based Attitude Determination for Spacecraft: System Engineering Design Study and Ground Testbed Results", 19th Annual AAS Guidance and Control Conference, February 7-11, 1996, Breckenridge, CO, AAS 96-16, pp. 1 - 20.

[2] M. W. Putty, and K. Najafi, "A Micromachined vibrating Ring Gyroscope", Technical Digest of the 1994 *Solid-State Sensor and Actuator Workshop*, Hilton Head Isl., SC, 6/13-16/94, Transducer Research Foundation, Cleveland (1994), pp. 213-220.

[3] J. Bernstein, S. Cho, A.T. King, A. Kourepenis, P. Maciel, M. Weinberg," A Micromachined Comb-Drive Tuning Fork Gyroscope," *Digest IEEE/ASME Micro ElectoMechanical Systems (MEMS) Workshop*, Ft. Lauderdale, Fl., pp. 143-148, February 1993.

[4] P. B. Ljung, T. N. Juneau, and A. P. Pisano, "Micromachined Two Input Axis Angular Rate Sensor", *Proceedings of the ASME Dynamic Systems and Control Division*, DSC-Vol. 57-2, 1995 IMECE/ASME, pp. 957-962.

OPTICALLY DRIVEN RESONANT MICROBEAM TEMPERATURE SENSORS FOR FIBER OPTIC NETWORKS

D. W. Burns, W. R. Herb, J. D. Zook and M. L. Wilson
Honeywell Technology Center
12001 State Highway 55
Plymouth, Minnesota 55441-4799

ABSTRACT

Fiber optic sensors combining micromachined polysilicon resonators with all-optical drive and sense techniques have been demonstrated for single-point and multi-point measurements of temperature using an optical fiber network. Sealed cavity resonators, varying from 122 μm to 324 μm in length, are integrally fabricated on a thin, single crystal silicon paddle with a thermal expansion mismatched film on the backside. Bending of the bimorph due to differential expansion with temperature alters the resonant frequency of the microbeam. The microbeam is driven photovoltaically with an embedded photodiode under the microbeam, and the resonance is detected by modulation of the reflected light as the microbeam vibrates. A vibration amplitude of 20 nm can modulate the reflected light by as much as 50%.

Each resonant microbeam is designed to operate within a prescribed frequency band, providing four to sixteen nodes of temperature sensing in a linear network with one microbeam at each node. Scanning, pulse-and-ringdown and self-resonant methods have been used to detect the microbeam frequencies. Individual devices have been demonstrated with a sensitivity up to 265 Hz/°C depending on paddle and bimorph thickness. Noise measurements on self-resonant microbeams have shown root Allan variances as low as 0.014 Hz, corresponding to a temperature resolution of 10^{-4} K. Hysteresis and repeatability measurements are currently underway. The devices have been packaged and operated successfully between 0°C and 200°C. Pressure sensors, vibration sensors, accelerometers, strain gages, acoustic emission sensors and other mechanical sensors can be configured using an appropriate silicon microstructure and attached to the optical fiber network.

INTRODUCTION

Honeywell is developing fiber optic sensors utilizing silicon micromachined, sealed cavity resonant microbeam technology initially developed at the University of Wisconsin [1] to meet the need for single-point and multi-point distributed optical sensor networks. An optical resonant microbeam sensor consists of a micromachined silicon chip attached to a multimode optical fiber. Light from an external source excites the beam into resonance, and vibrations of the microbeam modulate the light reflected back into the fiber which is detected by an external photodetector. We have previously demonstrated optical drive and sense using light from a semiconductor laser [2,3], and more recently reported the demonstration of optically excited self-resonance with unmodulated laser light [4]. Strain and pressure sensitivity have also been demonstrated, both at the wafer level and in pressure sensitive fixtures. Honeywell is currently teamed with the University of Wisconsin and Mobil Oil to develop a class of devices for fiber optic sensors and networks. These STORM (Strain Transduction by Optically Resonant Microbeams) devices are specifically aimed at precision temperature measurement for use in remote, hazardous or high EMI environments where advantages exist over inexpensive thermocouples and RTDs.

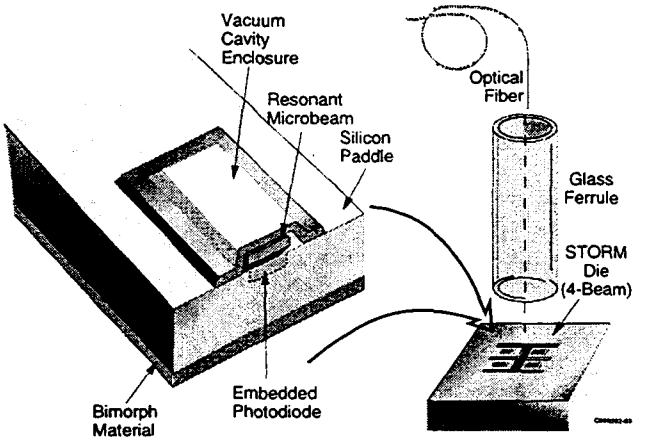


Figure 1. STORM temperature sensor die with ferrule and fiber. The STORM chip contains sealed cavity resonant microbeam strain sensors on a suspended bimorph of silicon and a thermally mismatched material. The optical fiber and glass ferrule are selectively aligned to a microbeam on the STORM temperature sensor die. Light transmitted over the fiber is used to excite the microbeam and detect its resonant frequency, which shifts appreciably with temperature.

The STORM temperature sensor consists of an optically excited, optically detected resonant microbeam on a silicon die which has been micromachined to form single crystal silicon "paddles" on which a thermal coefficient of expansion (TCE) mismatched material is deposited (see Figure 1).

The resonators are fabricated from thin films of fine-grain LPCVD polysilicon between 0.8 and 2.4 μm thick, and are enclosed in a vacuum cavity enclosure formed by a polysilicon shell. The gaps between the microbeam and enclosure, on the order of 0.5 μm, are accurately dimensioned based on design and modeling. The vacuum cavity allows the microbeam to vibrate freely at its resonant frequency, unaffected by the medium outside the integral enclosure. The silicon paddles isolate the microbeams from package-induced strains. A large shift in the resonant frequency of the microbeam occurs when strained by temperature-induced bending of the bimorph. An optical fiber with a 50 μm core is held in a simple glass ferrule which is fastened to the die. Incident light at a specified wavelength (ie. 780 nm) is partially transmitted through the shell and microbeam, and absorbed by the underlying photodiode. A photovoltage is generated which attracts the microbeam downwards. Vibrations of the resonant microbeam cause intensity modulation of reflected and transmitted light. The incident light is partially reflected from each surface (shell, microbeam and substrate) and collected by the same optical fiber. The microbeam vibration changes the modulation index for the reflected light. A 3db coupler is used in conjunction with a simple photodetector to measure the reflected light and determine the resonant frequency of the microbeam. Also, with proper choice of cavity and microbeam thickness, unmodulated light can be used to excite the microbeam to self-resonance [4].

0-9640024-1-8 $3.00 © 1996 TRF

MECHANICAL DESIGN

The design of the STORM temperature sensor involves electrical, optical, mechanical and structural considerations. The electrical considerations are minimal, requiring only that one side of the photodiode be ohmically connected to the resonant microbeam and that the photodiode has an appropriate carrier lifetime. Ohmic connection is obtained by implanting the microbeam with the same type (n or p) as the substrate it resides on, with the photodiode formed by implanting the opposite species prior to microbeam formation. The optical considerations are critical and are discussed in the next section. Mechanical considerations are made to maximize the signal within constraints imposed by the processing to avoid parasitic resonances and mechanical coupling, and allow for packaging with minimal package stress effects.

The temperature-sensitive bimorph is cantilevered which allows unconstrained deflections and prevents undesirable package stresses and hysteresis from affecting the performance. The strain ε_{app} induced at the surface of the paddle is uniform along its length [5] and is given approximately by (for $h_2 << h_1$)

$$\varepsilon_{app} = 3\Delta T \Delta\alpha \frac{(h_1 + h_2)h_2}{h_1^2}\left(\frac{E_2}{1-\upsilon_2}\right)\left(\frac{1-\upsilon_1}{E_1}\right) \quad (1)$$

where ΔT is the temperature shift, $\Delta\alpha$ is the TCE mismatch, h_1 and h_2 are the thickness of the silicon paddle and bimorph material, respectively, and E and υ are the elastic moduli and Poisson's ratio. The dependence of the resonant frequency of a microbeam on applied strain is given by

$$f_r = \frac{b_n}{2\pi}\left(\frac{E}{\rho}\right)^{1/2}\left(\frac{h_b}{l_b^2}\right)\left[1+\gamma_n\left(\frac{l_b}{h_b}\right)^2(\varepsilon_o + \varepsilon_{app})\right]^{1/2} \quad (2)$$

where E and ρ are Young's Modulus and density, h_b and l_b are the beam thickness and length, E is the intrinsic strain, and b_n and γ_n are coefficients with values 6.4586 and 0.2949 for the fundamental mode [6]. The temperature sensitivity of the microbeam may then be found by combining equations (1) and (2). If intrinsic strain in the microbeam is neglected, the temperature sensitivity may be expressed as:

$$\frac{\Delta f}{\Delta T} = \left(\frac{3\gamma_n}{2}\right)\Delta\alpha \frac{(h_1 + h_2)h_2}{h_1^2}\left(\frac{l_b}{h_b}\right)^2\left(\frac{E_2}{1-\upsilon_2}\right)\left(\frac{1-\upsilon_1}{E_1}\right) \quad (3)$$

From this equation it can be seen that the sensitivity increases with TCE mismatch and thickness of the bimorph as expected. The microbeam itself is mildly temperature sensitive: due to small density changes and a slight softening of the elastic moduli, a temperature sensitivity of about -40 ppm/°C can be added to the result of equation (2). The TCE-mismatched material can be placed on either side of the silicon paddle, though for expansion coefficients larger than that of silicon, the bimorph material should be placed on the backside. An additional film with a smaller TCE than silicon such as silicon nitride can be placed on the front, which passivates the microbeam and forms an antireflection coating. Consequently, the output frequency will have a negative coefficient.

ANSYS modeling was used to verify and further guide the design. Of particular concern is the avoidance of structural resonances of the paddles or die. Flexural or torsional modes of the cantilevered paddle can couple with microbeam motion, causing a loss in vibration amplitude and gaps in the frequency response. A plot of the structural modes for four microbeams on individual paddles is shown in Figure 2, with nonoverlapping microbeam modes avoiding the paddle modes.

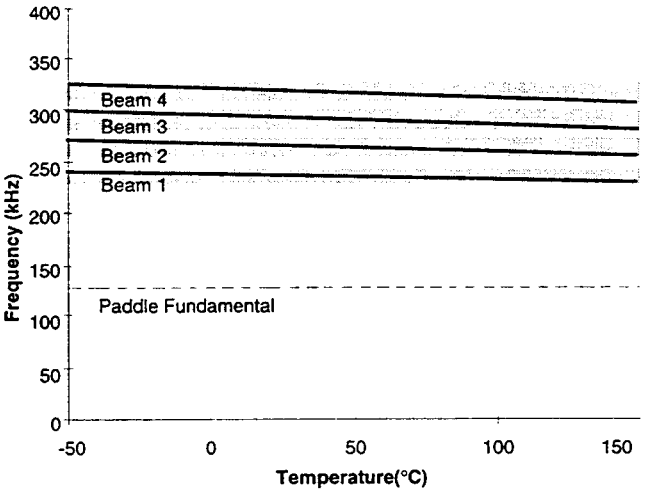


Figure 2. Modal analysis plot of four-beam STORM die. The structural modes of the paddles have been designed to be well above and below the microbeam frequencies to avoid unwanted mechanical coupling.

Mechanical modeling was also done on an alternative support design for the resonant microbeam. This design allows reduction of the thermal cycles during microbeam processing by using a deposited sacrificial layer rather than a thermally grown oxide. This microbeam design is supported by small posts formed by opening windows in the lower sacrificial layer prior to deposition of the beam poly, avoiding non-planar features spanning the entire beam width. The posts support the microbeam and provide ample coupling of applied strain to the ends (see Figure 5). ANSYS modeling shows minimal frequency variations with narrow or wide posts, or varying number of posts. Calculations also show little shift in the gage factor with varying post width. Most importantly, the new microbeam supports result in minimal static deflection of the free-standing microbeam with non-zero internal or applied strain, resulting in planar microbeams and consistent gaps with respect to the substrate and shell.

OPTICAL DESIGN

The design of optically resonant microbeam strain transducers has been described previously for optically driven microbeams [3] and self-resonant microbeams [4]. Although self-resonance has the attractive feature of not requiring laser modulation, it can easily produce optical fringe folding resulting in strong harmonic generation. Therefore for the optical network sensors, an optical design was chosen that has negative internal feedback rather than the positive internal feedback that leads to self-resonance. In either case, the microbeam thickness is chosen to be an odd number of quarter waves so that the rays reflected from the upper and lower surface interfere constructively to produce maximum modulation of the reflected light. The sum of the thicknesses of the two cavities above and below the microbeam is a multiple of a half wave for the same reason. Self-resonance can be avoided by simply interchanging the thicknesses of the upper and lower cavity so that the total reflectivity increases, rather than decreases, as the beam moves away from the substrate. With the proper choice of cavity and microbeam thicknesses, the reflectivity can be high and still readily modulated by the microbeam vibration over a fairly wide wavelength region. Calculated reflectivity as a function of microbeam position is illustrated in Figure 3, which shows a 20% modulation of the reflected light for a microbeam amplitude of 10 nm.

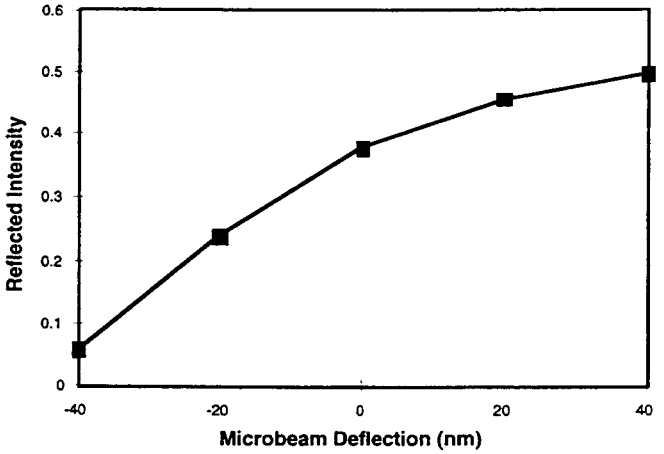


Figure 3. Reflectivity as a function of microbeam position for a case where the cavity thicknesses are 445 and 335 nm, and the incident wavelength is 780 nm. The microbeam and shell thicknesses are 2.26 μm and 2.16 μm, respectively.

DIE DESIGN AND FABRICATION

Four STORM die were designed with multiple microbeam sensors per die, each having a different resonant frequency for use on a different node of the network. Two temperature sensor die have four microbeams on four individual paddles, arranged in a 2x2 array; one with enlarged circular regions and circular photodiodes, the other with rectangular microbeams and elongated photodiodes. A sixteen-beam temperature and a three-beam pressure sensor complete the set. A plot of the four-beam device is shown in Figure 4.

The microbeam fabrication sequence uses six photomasks and begins with a deposition of Low Temperature Oxide (LTO) which is patterned and etched to form the post regions at the ends of the microbeams. An implant is done through the LTO using a photoresist mask to form a shallow photodiode directly under each microbeam. The microbeam is then formed by depositing LPCVD polysilicon, implanting it to provide ohmic contact to the substrate and removing the poly in the field region away from the microbeams. A second LTO layer is deposited and etched in ring-shaped regions around each microbeam to expose the substrate prior to deposition of the shell poly. A thin LTO layer is deposited and patterned in the ring region to form etch channels for subsequent sacrificial etching. The shell poly is deposited, implanted and patterned to cover each microbeam. A hydrofluoric acid solution is used as a sacrificial etchant to etch away the LTO in the field, in the etch channel regions, and in the regions immediately above and below the microbeam. Following carefully selected rinse cycles, a modification of the sublimation sequence previously reported [7] is used to freeze the final rinse fluid, which is then pumped out at reduced temperatures to produce free-standing beams. A final LPCVD film is deposited to seal the vacuum cavities as the thin etch channels grow shut, followed by a thin layer of silicon nitride that further enhances the seal and provides an antireflection coating. No metal depositions are used in the process sequence.

Microstructure formation follows microbeam fabrication and uses an SF_6-based RIE etch to form trenches around the microbeams. The microbeams and trenches are passivated with an additional layer of LTO to provide protection during the subsequent anisotropic etch. A KOH or EDP etchant is used after patterning the oxide on the backside of the wafer to locally thin the wafers, followed by removal of the passivation LTO to free the silicon paddles. A thin film of TCE-mismatched material is

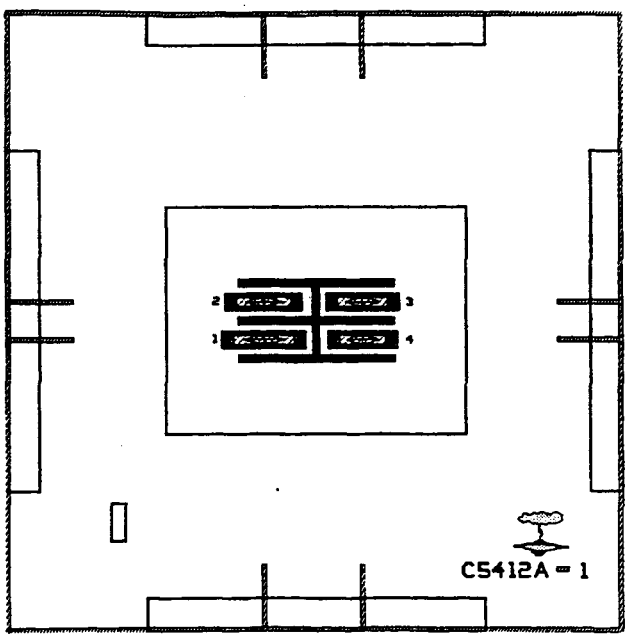


Figure 4. Four-beam STORM temperature sensor die. An array of microbeams ranging from 240 μm to 324 μm long are formed on individual paddles 400 μm long. Deep trenches are formed around the paddles, and a backside anisotropic etch is used to locally thin the silicon. The die footprint is 3.3 mm x 3.3 mm.

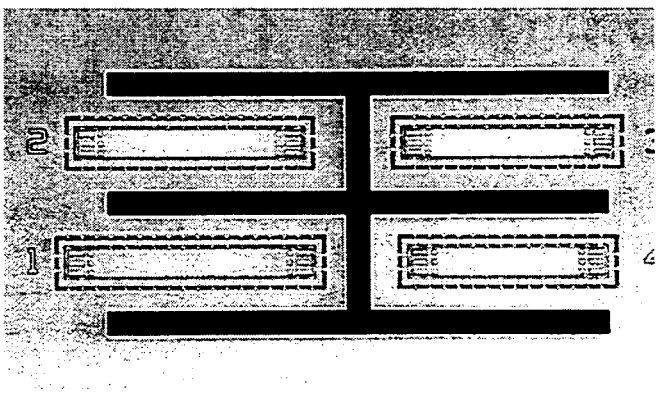


Figure 5. Optical micrograph of STORM temperature sensor die. Four microbeams are located on individual silicon paddles and operate at different frequencies. Close inspection reveals the post-style supports and the elongated embedded photodiodes.

deposited from the backside of the wafer onto the paddles to form the bimorph. A completed four-microbeam die is shown in Figure 5.

PACKAGING AND TESTING

The package currently being utilized for the STORM temperature sensors consists of a STORM die, glass ferrule with funnel-shaped ends to hold an optical fiber, and a metal housing (see Figure 1). The glass tube containing the optical fiber is attached to the STORM die. The tube has a thermal expansion coefficient similar to silicon and provides excellent stabilization of the end of the fiber with respect to the STORM chip. A fixture has been developed to aid in the packaging,

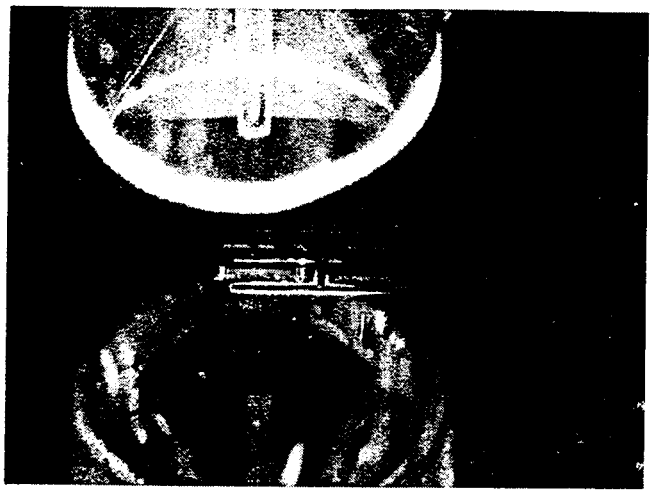


Figure 6. STORM packaging fixture. An optical fiber and glass ferrule are seen directly and reflected in a STORM 4-beam die as viewed through a microscope. Magnetic clamps hold the fiber and ferrule while vacuum holds the STORM die. Coarse alignment is achieved visually followed by fine adjustment using active alignment.

consisting of a video camera and monitor to allow the operator to visually align the optical fiber to the appropriate microbeam. Active alignment by monitoring the microbeam resonance is used for fine position adjustment of the optical fiber with independent x, y and z-axis control. The STORM chip is held with a vacuum chuck that can be electrically heated to 250°C so that solders or thermal set adhesives can be used to bond the glass ferrule and silicon die. The assembly is placed in a threaded metal package for external attachment and a strain relief for the pigtailed fiber. Current packages are 5 mm x 5 mm x 15 mm, but smaller packages are planned.

Testing of the STORM devices has been done at the wafer level and in assembled packages. During wafer level testing, the incident light strikes the microbeam at an angle and a second fiber is used to collect the reflected light. Resonant peaks are detected using a spectrum analyzer or a gain-phase analyzer. A plot of resonant frequencies from a sixteen-beam STORM die is shown in Figure 7. Resonant frequencies on test chips varied from 197 kHz for a 400 μm long beam to 1.17 MHz for a 100 μm long beam. Q values as high as 100,000 were measured on some beams. A hot chuck has been used to measure the temperature shift of the resonant microbeams. The first device tested had a base frequency of 207.5 kHz with a frequency sensitivity of -115 Hz/°C. A four-beam sensor has also been demonstrated, with baseline frequencies of 207.7 kHz, 250.4 kHz, 297.4 kHz and 340.6 kHz. The corresponding temperature sensitivities are between -112 Hz/°C and -118 Hz/°C for each beam. Higher sensitivities were obtained by depositing additional material on the back of the paddle, with sensitivities increasing linearly with material thickness, approaching -265 Hz/°C. A sensitivity of -30 Hz/°C was measured with only a silicon nitride passivation layer. Unpackaged die have been used to establish operating temperature ranges above 200°C (see Figure 8). Although noise measurements are yet to be made on networked sensors, measurements have been taken on optically-coupled self-resonant microbeams. An RAV of 0.014 Hz was measured at a sample time of 0.3 seconds, indicating that temperature resolution less than 100 μK is attainable with the optically interfaced sensors. A laser vibrometer has also been used to measure the structural resonances of the paddles [6], which are also in good agreement with the ANSYS modeling.

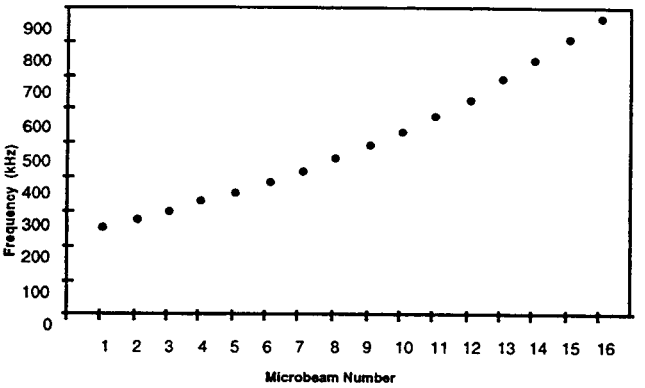


Figure 7. Measured resonances on a sixteen-beam sensor die. The sixteen frequencies vary from 251 kHz to 865 kHz (room temperature) for the 1.102 μm thick microbeams with signal and guard bands of approximately 10% between microbeams.

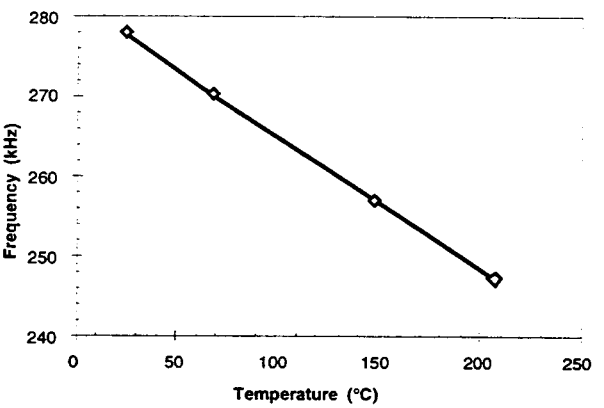


Figure 8. Frequency output of a single resonant microbeam with temperature. A 30 kHz shift was obtained over a 175°C temperature range for a STORM die mounted on a hot chuck. Operation above 200°C has been demonstrated.

OPTICAL NETWORKS

Two different network topologies have been investigated, one based upon a lightly tapped linear bus, and the second based upon a star bus. An advantage of a linear tapped bus is that a single fiber can be used to connect the entire network. However, this can also be a disadvantage if the fiber is damaged or broken since all sensors downstream from the break would no longer function. The star bus has the advantage that the signal levels are higher and there is no single point failure mechanism except if the star coupler is damaged.

Figure 9 illustrates the star bus design. It is possible to use several different wavelength lasers together with wavelength filters at the STORM chip in order to use the same resonant frequency several times on the bus. Thus the combination of Wavelength Division Multiplexing (WDM) with Frequency Division Multiplexing (FDM) techniques allows for the connection of much larger optical sensor networks.

The signal obtained from an actual STORM device is shown in Figure 10. The noise floor is at -90 dBm, and the signal is about +45 dB above the noise floor. The signal to noise ratio (SNR) of this signal is very large making signal processing much easier. In Figure 10, the spectrum analyzer's filter bandwidth is convolved with this resonant peak, so the actual peak is much sharper than shown in this figure. Figure 11 shows data obtained from a small 2x2 star network similar to

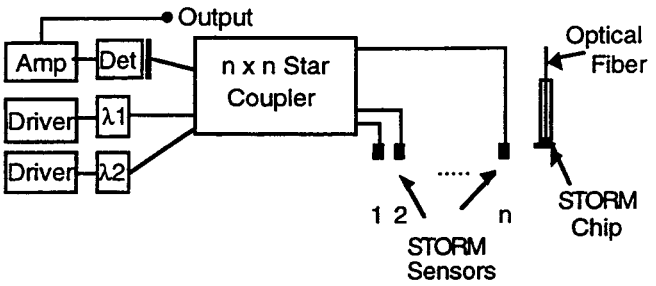


Figure 9. Optical fiber network block diagram. One to n STORM die are attached to system optics using an n x n star coupler.

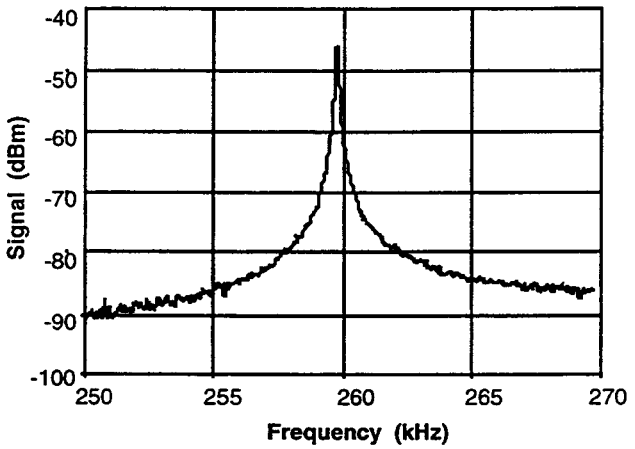


Figure 10. The resonant frequency spectrum of a STORM device .

that detailed in Figure 9. Two temperature sensors each with different resonant frequencies were connected to this network. Note that each frequency is free from inter-modulation harmonics and each peak has an excellent SNR.

CONCLUSIONS AND FUTURE DIRECTIONS

STORM sensors represent a new and novel optoelectro-mechanical sensor technology which combines the strength of silicon resonant microbeam technology - including high sensitivity, stability and robustness - with that of optical fiber technology - including EMI immunity and intrinsic safety. The STORM technology provides the generic basis for a wide range of sensors that offer significant advantages over conventional silicon-only or fiber-only sensors. Fiber optic networks for silicon-based resonant sensors are being demonstrated, with first applications to temperature sensing. The resonant microbeams are highly sensitive to axial strain and can be configured with suitable silicon microstructures to form precision pressure sensors, accelerometers, force sensors, vibration sensors, rotation sensors and load cells as well as temperature sensors, with possibilities of multiple sensors per node if desired. Optical fiber interfacing to resonant microbeam sensors can be achieved without deterioration of the output signals due to EMI, no loss in ambiguity as in many interference-based sensors and without any on-chip electronics. The limits of networkability have not been firmly established, though a variety of bus architectures appear feasible. Further development of the optically self-resonant and non-self-resonant concepts are needed to address questions relevant to producibility, fiber-to-chip coupling and tolerances on optical components.

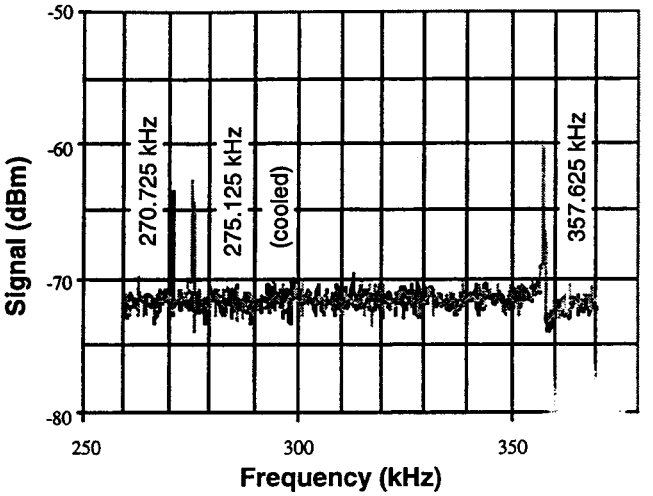


Figure 11. Output from a sensor network with two devices. Two overlapping scans are shown, one taken later in time with (only) one node reduced in temperature and showing a frequency shift of 4400 Hz (ΔT = -30°C). The sensors were separated by several meters of optical fiber, joined by a four-port coupler.

ACKNOWLEDGMENTS

The authors would like to gratefully acknowledge Dane Larson and Jerry Huber for sensor processing, Cindy Bassett, Ernest Satren and Jerry Bolser for sensor testing and packaging, Hirak Chanda for FEM calculations, Peg Wilson for layout, Barry Cole for optical modeling, Professor Guckel at the University of Wisconsin for his support and technical contributions, and ARPA (contract #DAAL01-94-C-3427) for partial funding of this effort.

REFERENCES

1. J. D. Zook, D. W. Burns, H. Guckel, J. J. Sniegowski, R. L. Engelstad and Z. Feng, "Characterization of Polysilicon Resonant Microbeams," Sensors and Actuators A, Vol. 35 (1992), pp. 51-59.

2. H. Guckel, M. Nesnidal, J. D. Zook and D. W. Burns, "Optical Drive/Sense for High Q Resonant Microbeams," Technical Digest of the 7th International Conference on Solid-State Sensors and Actuators, pp. 686-689.

3. J. D. Zook, D. W. Burns, J. N. Schoess and H. Guckel, "Optically Resonant Microbeams," SPIE, Photonics West '95, Feb. 10, 1995.

4. J. D. Zook, D. W. Burns, W. R. Herb, H. Guckel, J.-W. Kang and Y. Ahn, "Optically Excited Self-Resonant Strain Transducers," Transducers '95, pp. 600-603.

5. S. Timoshenko, "Analysis of Bi-Metal Thermostats," J. of the Optical Society of America, Vol. 11, 1925, pp. 233-255.

6. D. W. Burns, J. D. Zook, R. D. Horning, W. R. Herb and H. Guckel, "Sealed-Cavity Resonant Microbeam Pressure Sensors," Sensors and Actuators A, Vol. 48 (1995), pp. 179-186.

7. H. Guckel, J. J Sniegowski, T. R. Christenson and F. Rassi, "The Application of Fine-Grained, Tensile Polysilicon to Mechanically Resonant Transducers," Sensors and Actuators, A21-A23 (1990), pp. 346-351.

MICROMACHINED DUAL INPUT AXIS ANGULAR RATE SENSOR

Thor Juneau, A. P. Pisano
Berkeley Sensor & Actuator Center, 497 Cory Hall
University of California at Berkeley
Berkeley, CA 94720
e-mail: thorj@nobelium.me.berkeley.edu

ABSTRACT

The need for inexpensive yet reliable angular rate sensors in fields ranging from automotive to consumer electronics has sparked prolific micromachined angular rate sensor research. The vast majority of research has focused on single input axis angular rate sensors based upon either translational resonance, such as tuning forks, or structural mode resonance, such as vibrating rings. However, this work presents an entirely novel approach based on *angular* resonance of a rotating rigid rotor suspended by torsional springs. The inherent symmetry of the circular design allows angular rate measurement about two axes simultaneously. Using a commercial micro-accelerometer production process, two versions of a fully integrated, 660 μm diameter dual input axis angular rate sensor have been designed, fabricated, and tested. Minimum detectable signals as small as 1.2°/sec over a 25 Hz bandwidth were achieved with untuned rate sensors having mismatched mode frequencies. However, with mode frequency tuning the performance was improved to 10°/min over an equivalent bandwidth.

INTRODUCTION

The advent of micromachining has raised the possibility of fabricating high performance sensors without the associated high costs. In particular, there has been great interest in manufacturing micromachined angular rate senors [1]. Until now, the high cost of even medium performance rate sensors has precluded their use in most consumer products and high volume military applications, such as automatic vehicle braking systems, automobile navigation, camera stabilization, and artillery munitions guidance.

A number of angular rate sensors have been recently reported, but the majority of these designs are single axis sensors based on translational vibration [2] or structural mode vibration [3]. In addition, many reported rate sensors require fabrication processing not compatible with standard VLSI processes. In contrast, this research has focused on surface micromachining a dual input axis angular rate sensor in the fully integrated Analog Devices Inc. (ADI) microaccelerometer production process [4]. This proven fabrication process allowed inexpensive production with integrated electronics to improve noise performance. Unlike most previous designs, the dual axis rate sensor presented here is based on rotational resonance of a polysilicon disk. Since the disk is symmetric in two orthogonal axes, the sensor can sense rotation equally well about both these two orthogonal axes. When combined with a z-axis rate sensor [5] or [3], the angular rate about all three axes can be measured using one die.

BASIC OPERATING PRINCIPLE

The micromachined dual input axis angular rate sensor is composed of an inertial rotor, an actuation means to drive this inertial rotor into resonance, and a sensing system to measure angular deflections of the inertial rotor due to Coriolis accelerations. The inertial rotor is a 2 μm thick polysilicon disk with a 150 μm radius. As depicted in Fig. 1, this inertia rotor is suspended on four symmetrically placed beams anchored to the substrate. These beams provide a torsional suspension for the rotor allowing compliance about three mutually perpendicular axes. To achieve maximum mechanical sensitivity, the suspension was designed to match the frequencies of all three rotational modes sufficiently well so that electrostatic tuning may compensate for fabrication process variation.

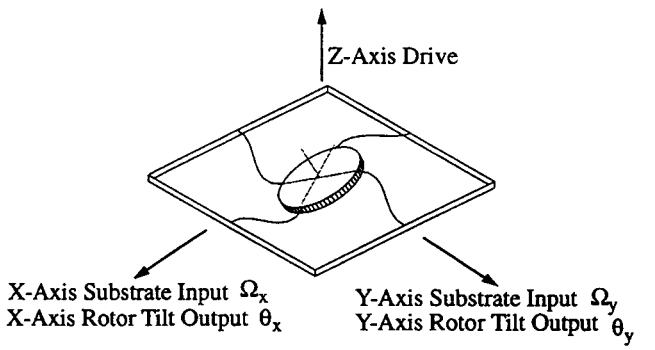


Figure 1: Conceptual illustration of inertial rotor and four quad symmetric suspension beams.

The inertial rotor is driven at angular resonance about the z-axis (perpendicular to the substrate) using the electrostatic comb drive described in a later section. The resonating inertial rotor is responsive to substrate rotation rates about both the x- and y-axes parallel to the substrate as defined in Fig. 1. A rotation rate about the x-axis induces a Coriolis angular oscillation about the y-axis, and likewise a rotation rate about the y-axis induces a Coriolis angular oscillation about the x-axis. Sensing these Coriolis oscillations and relating them to the original substrate rate inputs forms the foundation of angular rate sensor operation.

The Coriolis angular oscillation or tilting about each sense axis is embodied in the dynamical equation of motion below where θ_x is the sense axis tilt angle, I_{xx} is the moment of inertia, C_x is the damping coefficient, and K_x is the spring constant all reckoned about the x-axis. Only the x-axis is given since the equation for the y-axis is essentially identical except x and y subscripts are reversed and the right hand side is positive.

$$I_{xx}\ddot{\theta}_x + C_x\dot{\theta}_x + K_x\theta_x = -I_{zz}\Omega_y \times \theta_{z0}\omega_z Cos(\omega_z t) \qquad (1)$$

The term on the right hand side is the Coriolis acceleration due to conservation of angular momentum. The angular rate input about the y-axis to be measured is Ω_y. This is multiplied by the z-axis moment of inertia I_{zz} and the resonating z-axis angular rate which is an ω_z frequency cosine with amplitude $\theta_{z0}\omega_z$. Hence the Coriolis acceleration and the ensuing Coriolis angular oscillation of the rotor both can be thought of as amplitude modulated signals with a carrier frequency ω_z. The amplitude of this modulated Coriolis oscillation is directly proportional to the y-axis input rate Ω_y. This Coriolis oscillation is sensed using the change in capacitance between the rotor and four sense electrodes beneath the inertial rotor. Since the Coriolis oscillation is in phase with the

0-9640024-1-8 $3.00 © 1996 TRF

resonating z-axis angular rate, simply demodulating the sense signal with the resonant drive signal yields the desired rate measurement. The dynamics of the second sense axis are similar.

The fact that the Coriolis oscillation is in phase with the drive forcing signal is of paramount importance. Unfortunately due to fabrication imperfections, the disk will wobble slightly in phase with resonance position. The wobble can masquerade as a false rate measurement called quadrature error [5] and can be larger than 200°/sec. Ideally, the final demodulation with the drive signal will eliminate the quadrature error from the measurement because the quadrature error is 90° out of phase. However, in practice small phase lags (~0.5°) in the electronics allow a small percentage of quadrature error to corrupt the measurement. Such factors as the performance required, the lag of the electronics, and the stability of resonant amplitude dictate whether the DC offset due to quadrature error may be tolerated or require active cancellation. With the basic operation described, attention is now focused on specifics.

MECHANICAL STRUCTURE

The heart of the dual input axis angular rate sensor is the micromachined structure itself. It is fabricated in true surface micromachining tradition by using a structural polysilicon layer deposited over a sacrificial oxide and patterned [4]. Etch holes in the sacrificial oxide provide anchor points to the underlying substrate after the sacrificial oxide is removed. The combined inertial rotor and spring suspension were carefully designed to satisfy many design constraints placed on the natural frequencies of the structure. For example, natural frequencies in the upper audio range are desirable for shock and random vibration robustness. The rate sensor should be insensitive to all translational accelerations, so the natural frequency of all three translation directions must be high. In this design, the x-axis and y-axis directions parallel to the substrate are very stiff. Although the z-axis perpendicular to the substrate is not nearly as stiff (approx. 15 kHz), the differential sensing scheme discussed later nulls z-axis translation sensitivity to first order. Ideally, the drive natural frequency and the pair of sense natural frequencies should be designed to be nearly equal. However, process variation will result in mismatch unless electrical tuning is used to artificially match frequencies. Electrostatic forces due to voltage differentials are always attractive, hence spring constants and natural frequencies can only be decreased. For this reason, the sense axis natural frequencies were designed to be 10% above the drive natural frequency so that they may be down tuned to match the drive frequency. This is implemented by using spring suspension beams with a cross-section of 2 μm thick by 1.8 μm wide with the thicker dimension stiffening the sense axis spring constants.

With the above constraints in mind, two distinct rate sensor versions were fabricated. Each version has identical rotor sizes, but different suspension designs. A compliant serpentine suspension shown in Fig. 2 gave a low natural frequency of 21 kHz. The main spring suspension beam is 180 μm in length, while the shorter serpentine beams are 120 μm in length. The second version has a less compliant suspension composed of a single 180 μm beam with some stress relief (as shown later in Fig. 6) which gave a higher natural frequency of 28.6 kHz. In each version, an external suspension extends from the outside of the inertial rotor. This suspension design can accommodate warping of the inertial rotor due to large residual stress gradients in the polysilicon [6].

ELECTROSTATIC RESONANT DRIVE

There can be no rate sensing unless the inertia rotor is in motion. To this end, the inertial rotor is driven into rotational resonance about the z-axis perpendicular to the substrate. Electrostatic comb drive was used as it has the distinct advantage of remaining linear despite large displacements. However, this technique requires that the angular rate sensor be operated with ambient pressure below 150 mTorr.

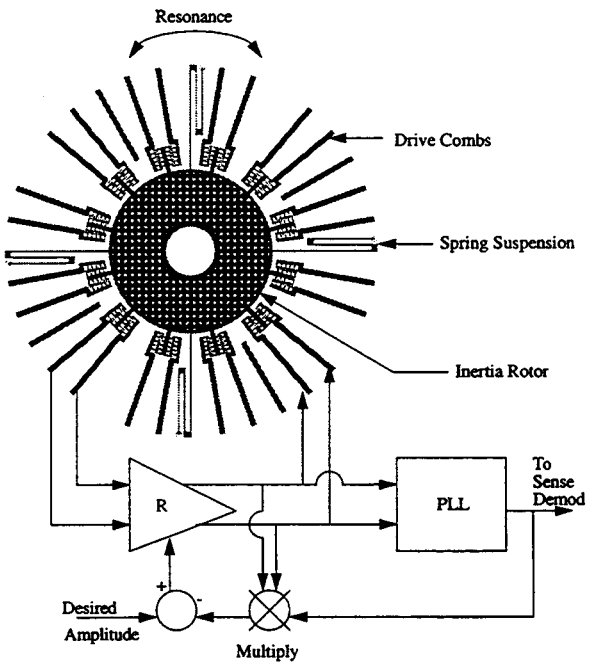


Figure 2: Diagram of polysilicon dual input axis rate sensor looking down on die. Schematic shows trans-resistance resonant drive with automatic gain control to hold oscillation amplitude constant plus phase-lock-loop for Coriolis demodulation.

As discussed in Nguyen [7] a trans-resistance amplifier provides positive feedback which induces resonance. The amplifier and combs (see Fig. 2) were designed to ensure adequate forcing to overcome air damping. This design used a differential drive comprised of 60 comb teeth and requiring approximately 3 MΩ drive with a 10 volt difference between the inertial rotor and drive combs. In addition, the driving circuitry has low phase lag and feedthrough to reduce the effect of quadrature error. The resonant amplitude is carefully controlled using automatic gain control because scale factor is directly proportional to resonant amplitude. A phase-lock-loop provides a clean signal for Coriolis signal demodulation discussed next.

DUAL AXIS CORIOLIS OSCILLATION SENSING

The inertial rotor tilts due to Coriolis acceleration and this Coriolis oscillation is measured by sensing the change in capacitance between the inertial rotor and the four n+ diffusion sense electrodes beneath the inertial rotor. As shown in Fig. 3, when a differential modulation voltage is applied to a pair of diametrically opposed sense electrodes (each shaped as a quarter-pie section), a rotation of the rotor about the sense axis of these electrodes produces a voltage on an integrator electrically connected to the rotor. Since the pair of sense electrodes provides differential

sensing, any translation of the rotor up or down along the z-axis direction produces no change in sense output. Dual axis operation is achieved by using a different voltage modulation frequency (200 kHz and 300 kHz) for each pair of sense electrodes. Separate demodulation circuits for each sense axis provide two voltage outputs proportional to the two angular rate inputs. The output voltages must be demodulated twice: first a demodulation removes the high sense modulation frequency, and second a demodulation removes the drive resonant frequency leaving the base band rate input measurements.

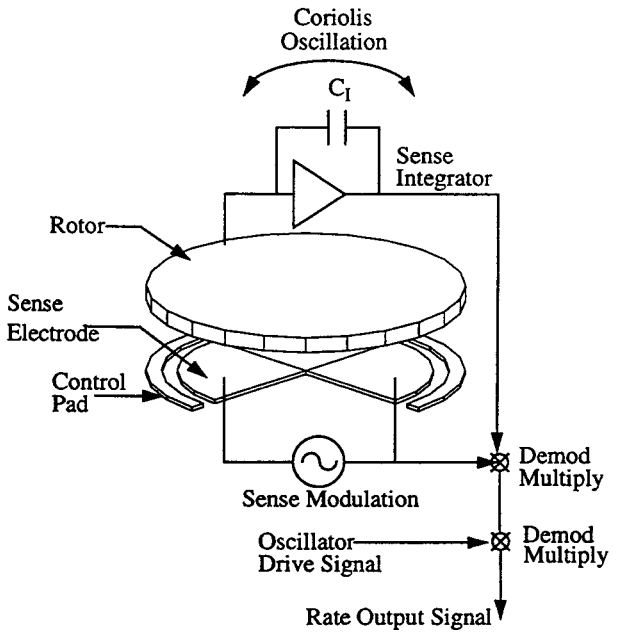


Figure 3: Conceptual side view showing rotor with underlying diffusion sense electrodes for one axis. Please note only two of four sense electrodes are shown for simplicity.

The afore mentioned single ended integrator scheme (using an integrating capacitor C_I of 50 fF) was used for two principle reasons. First, the structure itself has far less noise producing parasitic capacitance than the n+ diffusion electrodes. Second, the integrator improves linearity [5]. Absence of distortion in the sense circuitry is critical, for any odd harmonic distortion can lead to quadrature error signals mixing with true Coriolis measurement signals. One drawback of sensing on the driven structure is that despite fully differential comb drive, a small inherent voltage at twice the drive frequency appears on the integrator. In general, this signal is small and is far from the crucial sense modulation frequencies.

SENSITIVITY

The mechanical sensitivity relating Coriolis rotor oscillation amplitude to substrate angular rate input for each sense axis can be derived from Eq. 1 with the final form shown below for the x-axis.

$$\left|\frac{\theta_x}{\Omega_y}\right| = \left|\frac{2\theta_{z0}\omega_z}{\omega_x^2 + \frac{j\omega_x\omega_z}{Q_x} - \omega_z^2}\right| \qquad (2)$$

This is the standard second order frequency response with θ_{z0} the resonant drive amplitude, ω_z the resonant drive frequency, ω_x the sense axis frequency, and Q_x the quality or inverse damping coefficient. Since the rate sensor is operated in partial vacuum, the quality factor Q_x is generally above 800. Thus, a large resonant peak exists which provides excellent gain when drive and sense frequencies match.

The combined mechanical and electrical sensitivity must also be considered since mechanical sensitivity alone can be misleading. The combined mechanical system, sense electrodes, and integrator can be summarized as below with sense voltage V_s, integrating capacitor C_I, and electrode capacitive sensitivity $\partial C/\partial\theta_x$.

$$\left|\frac{V_{out}}{\Omega_y}\right| = \left|\frac{\theta_x}{\Omega_y}\right|\left|\frac{V_{out}}{\theta_x}\right| = \left|\frac{2\theta_{z0}\omega_z}{\omega_x^2 + \frac{j\omega_x\omega_z}{Q_x} - \omega_z^2}\right|\left|2\frac{\partial C}{\partial\theta_x}\frac{V_s}{C_I}\right| \qquad (3)$$

Clearly increasing the sense voltage differential V_s between the rotor and the quarter circle sense electrodes will increase overall sensitivity. Unfortunately, this sense voltage also results in an electrostatic force which tends to pull the inertial rotor down to the substrate. The critical pull down voltage limits the maximum applicable sense voltage and is proportional to the natural frequency for z-axis translation perpendicular to the substrate. For this reason, the sense voltage for weaker suspensions must be reduced, nearly eliminating all mechanical sensitivity gains. However, the sense voltage also has another side effect which complicates simple sensitivity generalizations. The sense voltage down tunes the sense axes natural frequencies as discussed in the following section.

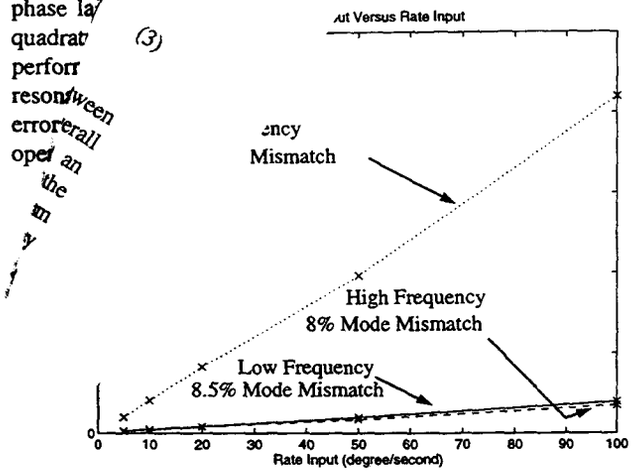
FREQUENCY TUNING

The voltage applied between the inertial rotor and the four quarter circle sense electrodes results in an attractive electrostatic force which down tunes the natural frequencies of both the rotational sense axes. This frequency tuning can be used to nearly match frequencies and vastly improve sensitivity. The tuning range is limited because any tuning voltage must be below the critical pull down voltage. Hence, the design target values for the drive mode and the sense mode frequencies must be far enough apart to accommodate process variation, but not so far apart that they are outside the tuning range.

Under open-loop operation, matching natural frequencies *exactly* has several drawbacks. Because the gain and phase change radically at the resonant peak, gain is not constant over the bandwidth, frequency drift cause large scale factor changes, and the phase swings from 0 to 180 degrees. Cross-sensitivity is also compromised as large resonant peak gained signals many cross-couple. Therefore during open-loop operation it is advantageous to match modes only within 2-5%. On the other hand, the use of closed-loop feedback should alleviate most the these difficulties.

EXPERIMENTAL RESULTS *WITHOUT* TUNING

Two versions of the dual input axis angular rate sensor were fabricated using the ADI surface micromachining processes. In addition to performance, those parameters with greatest impact on performance, such as natural frequency, quality factor Q, and maximum sense voltage, have been experimentally identified. The two versions have identical inertial rotor sizes but different suspensions giving drive natural frequencies of 28.6kHz and 21kHz. Both sensors were tested in 60 mTorr vacuum which gave the high frequency version a quality factor Q of 960 and the low frequency



igure 4: Graph showing rate sensor voltage amplitude output versus rate amplitude input for a 1Hz sinewave.

EXPERIMENTAL RESULTS *WITH* TUNING

Due to the high quality factor Q, great performance improvement may be reaped using electrostatic tuning. As displayed in the top graph of Fig. 5, increasing voltage between the quarter circle sense electrodes and the rotor down tunes the sense frequency more than 2500 Hz, while down tuning the drive frequency only slightly.

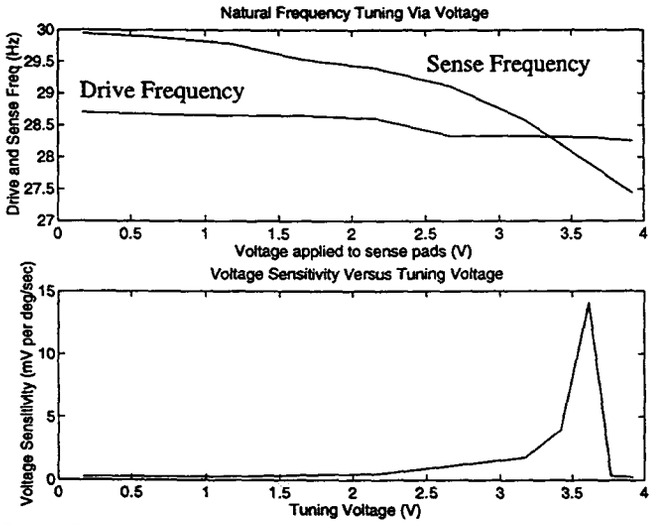


Figure 5: The top graph shows drive and sense frequency down tuning due to applied voltage. The lower graph shows the dramatic impact frequency matching has on sensitivity.

The bottom graph in Fig. 5 shows sensitivity is maximized roughly at the point where the frequencies match with some offset due to experimental error. Tuning resulted in performance of 10 °/ min over a 25 Hz bandwidth.

CONCLUSION

The underlying operating principles and resulting design methodology for a dual input axis angular rate sensor based on angular resonance were presented. Fabricated rate sensors (as shown in Fig. 6) without tuning provided good noise performance. However, electrostatic frequency tuning improved performance significantly.

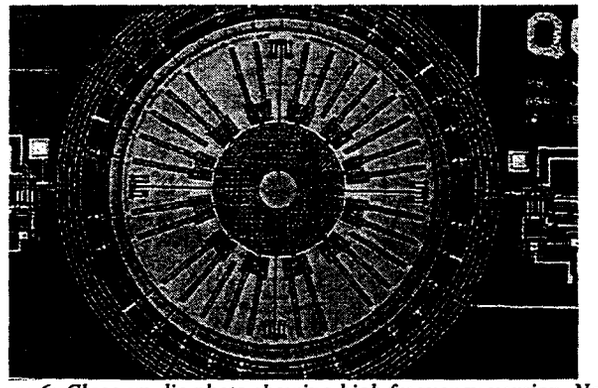


Figure 6: Close-up die photo showing high frequency version. Note that the beam suspension is not serpentine.

ACKNOWLEDGEMENTS

First, the authors must thank ADI for all device fabrication using the excellent ADI surface micromachining process. Also, without the financial support of both ARPA (grant TP-26890) and the California PATH project, this research could not continue. The insights and guidance of Professor R. T. Howe have been invaluable. The first author also must thank Per Ljung for his previous rate sensor work. However, the most heart felt thanks on behalf of the first author goes to Bill Clark for his help with key circuitry.

REFERENCES

1. J. Soderkvist, "Micromachined Gyroscopes", Sensors and Actuators A, pp. 65-71 (1994 v43).

2. J. Bernstein and M. Weinberg, "A Micromachined Comb-drive Tuning Fork Rate Gyroscope", *21st Joint Serv. Data Exchng for G. N. and C.*, Palm Springs, CA, Oct. 1992.

3. M. Putty and K. Najafi, "A micromachined Vibrating Ring Gyroscope", *1994 Solid-State Sensor and Actuator Workshop*, Hilton Head, pp. 213.

4. R. S. Payne, S. Sherman, S. Lewis, R. T. Howe, "Surface Micromachining: From Vision to Reality to Vision", 1995 *IEEE International Solid State Circuits Conference*, pp. 164.

5. W. A. Clark, R. T. Howe, R. Horowitz, "Surface Micromachined Z-Axis Vibratory Rate Gyroscope", *1996 Solid-State Sensor and Actuator Workshop*, Hilton Head.

6. P. B. Ljung, T. Juneau, A. P. Pisano, "Micromachined two input axis angular rate sensor", ASME International Mechanical Engineering Congress and Exposition, session DSC-16, (1995).

7. C. H. Nguyen, "Micromachanical Signal Processors", Ph.D. Thesis, University of California at Berkeley, Berkeley, CA 1994.

SUBJECT INDEX

G

H, I

J, K, L

M

N, O

P, Q

R, S

T

U, V, W

X, Y, Z

AUTHOR INDEX

J, K

L

M

N, O

P, Q

R, S

T, U, V

W

X, Y, Z