ELSEVIER

Sensors and Actuators A 95 (2002) 175–182

SENSORS AND ACTUATORS A PHYSICAL

www.elsevier.com/locate/sna

CMOS-MEMS membrane for audio-frequency acoustic actuation

John J. Neumann Jr.*, Kaigham J. Gabriel

MEMS Laboratory, Electrical and Computer Engineering Department, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213, USA

Abstract

Using CMOS-MEMS micromachining techniques we have constructed a prototype earphone that generates audible acoustic output from 40 Hz to 10 kHz. The fabrication of the acoustic membrane consists of only two additional steps subsequent to the prior post-CMOS micromachining steps developed at Carnegie Mellon University [Sens. Actuat. A 57 (1996) 103–110]. The ability to build a membrane directly within a standard CMOS chip, integrating mechanical structures with signal processing electronics will enable a variety of applications including economical earphones, microphones, hearing aids, high-fidelity earphones, cellular phones and noise cancellation earphones and devices. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Acoustics; CMOS; Membrane; MEMS

1. Introduction

The fundamental challenge of using MEMS for acoustic applications is how to achieve significant sound pressures with a device of relatively small area. This is a greater challenge for audio frequency MEMS than for ultrasonics, because the radiation impedance (which determines the pressure generated for a given membrane displacement) decreases with increasing wavelength. For this reason, we have restricted our study of audio acoustic actuators to those that sit in the ear canal, where the small volume of the canal presents a larger acoustic impedance than is achievable in the free air.

Our goal is to create a high-performance, economical CMOS-MEMS earphone for speech or music playback. In order to create a strong low-frequency response with a small device, we made an airtight seal an integral part of the device concept. Its importance is readily apparent when comparing similar devices, e.g. a cantilever acoustic actuator, with and without this isolation of front and back, which prevents an acoustic short circuit [2,3]. By keeping to a few custom fabrication steps, complexity is decreased and the yield is maximized. Previous MEMS audio devices such as dome-shaped actuators [4], have been made with large output and good frequency response, but require numerous custom fabrication steps and materials. In contrast, the sealed membrane of the present CMOS-MEMS device is formed in a single chemical vapor deposition using standard polymer materials.

CMOS-MEMS acoustic devices have the additional advantage of on-chip electronics. In the case of a microphone, sensing circuitry (capacitive or piezoresistive) can be placed close to the mechanical structure, minimizing electromagnetic pickup and parasitic capacitance. One unique advantage is that the CMOS-MEMS membranes can have as many different conductive layers as the CMOS process (e.g. three for MOSIS HP 0.5 μm process), and these conductors can be arranged in arbitrary geometries (for example to achieve a tailored force distribution, or to use one conductor for sensing and another for actuation). One possible application for this would be a self-correcting earphone that adjusts in real time to changing acoustic environments by monitoring the position of the membrane and/or the sound pressure near the device. Another use of this would be to correct the driving voltage for non-linearities in the membrane mechanical response.

2. Concept and methods

To construct a speaker or microphone in CMOS-MEMS, we pattern a mesh or screen-like structure in one of the metal layers, much like the proof mass in CMOS accelerometers [5]. After removal of the underlying silicon using the CMU CMOS-MEMS micromachining process [1], a polymer is conformally deposited on the mesh to form a continuous membrane and to make it airtight. To operate as a speaker, the CMOS metal layer of the mesh structure embedded

* Corresponding author.
E-mail addresses: jneumann@ece.cmu.edu (J.J. Neumann Jr.), kgabriel@ece.cmu.edu (K.J. Gabriel).

0924-4247/02/$ – see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0924-4247(01)00728-2

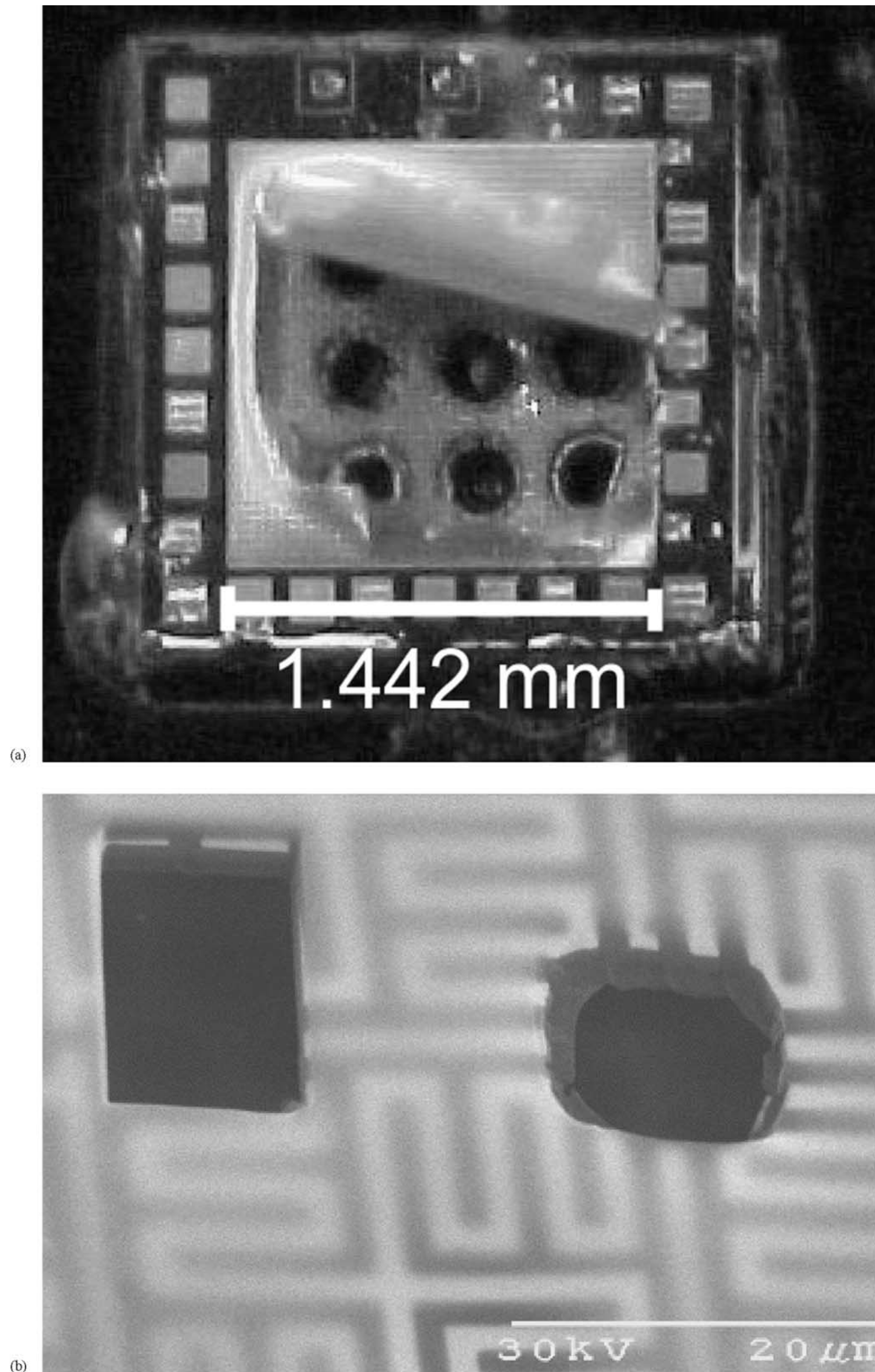

Fig. 1. (a) Photograph showing top of chip. The membrane, formed of CMOS metal-oxide beams and polymer coating, is peeled back revealing the vent holes leading from the back of the chip into the sub-membrane chamber. Located around the periphery of the membrane are the two bond pads and 27 smaller mesh structures used for testing purposes. (b) Scanning Electron Micrograph of section of membrane, showing holes formed by a focused ion beam. The cleaner hole (left) was created by a smaller beam current. The aluminum sections of the beams are visible through the polymer and also along the edges of the holes.

within the polymer is set to varying electrical potentials with respect to the silicon substrate, generating electrostatic force and motion normal to the chip's surface.

The etch of the silicon underneath the membrane structure is controlled so that a cavity of the desired depth is formed. The depth should be great enough so that the membrane has room to vibrate with the desired amplitude, but not so large that an unreasonably high voltage is required to electrostatically deflect the membrane. Our cavities were typically 60–80 μm deep.

A prototype device is shown in Fig. 1a. The membrane on this device is peeled back, revealing the vent holes that were etched from the back surface, through the silicon substrate and into the sub-membrane cavity. Surrounding the large central membrane are the smaller membranes of varying design (intended for testing purposes) and the two bond pads, which connect to the metal layer in the membrane and the silicon substrate. Fig. 1b is a close-up SEM of the polymer-coated membrane, showing two holes made with different focused ion beam currents. Note the embedded metal layer in the left rectangular hole.

One difficulty in creating a large membrane with CMOS-MEMS is the buckling caused by stresses inherent in the oxide and metal, as well as from bimorph temperature-dependent stress differences [6]. For a 50 μm-long cantilever beam made of metal and oxide, the out-of-plane curl can be as much as 1 μm, increasing with the square of the length. To solve this problem, we developed the serpentine mesh design shown in Fig. 2. The mesh is composed of serpentine springs running in both the x and y directions, which provide stress relief. The individual members which make up the springs are kept short (<50 μm) so as to limit curl. The gaps between the beams are of uniform size, allowing uniform coating with polymer. In the chips that were fabricated, both the gaps and beam widths were 1.6 μm. The maximum out-of-plane buckling (in the center of the membrane) is between 10 and 20 μm for the 1.4 mm × 1.4 mm prototype membrane.

The device, in its housing, is placed in the ear with the membrane facing into the ear canal. The backside of the chip should provide a path for air to escape to the outside world, without creating an acoustic "short circuit" within the ear canal. This can be accomplished by mounting the CMOS die on a substrate with a hole in it, and sealing around the edge of the chip with epoxy, as shown in Fig. 3. Holes of various sizes are etched into the back of the chip, leading into the cavity under the membrane. These have important effects on the behavior of the device, such as reducing the acoustic impedance on the backside of the membrane (allowing greater displacements for a given electric force), and adding a resistive component that damps out the unwanted resonances.

Fig. 2. The layout of the serpentine mesh pattern that defines the membrane. Beams are represented by gray, gaps by black. One unit cell is 32 μm across.

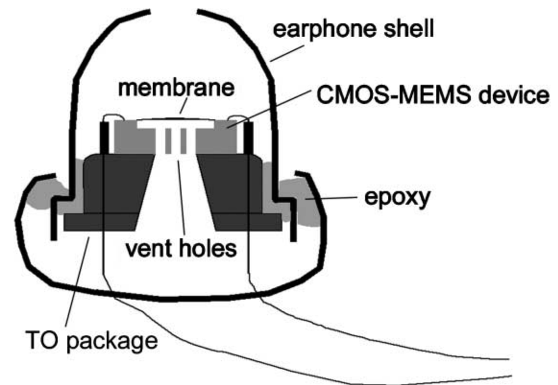

Fig. 3. Drawing of earphone construction. The TO package is drilled to allow air to vent from the backside of the chip. Epoxy is used to seal around the edge of the TO package to prevent an acoustic short circuit between the front and back of the chip.

3. Fabrication process

Before the patterning of the CMOS structures on the front side, the backside vent holes must be etched using the oxide layer as an etch-stop. The hole sizes are about 100 μm across, with the total areas of multiple hole patterns (3 × 3 grid) being around 50,000–100,000 μm^2. After the vent holes are etched, processing of the CMOS (front) side of the chip follows the steps shown in Fig. 4:

1. The chip comes from a CMOS fabrication foundry covered with a layer of protective glass (silicon dioxide). Regions meant for the mechanical structures are patterned in one of the metal layers, usually the topmost layer.
2. The oxide is etched anisotropically down to the silicon substrate, the metal layers acting as a mask to define the mesh structure.
3. The underlying silicon substrate is etched with a 60-min DRIE anisotropic deep etch followed by a 7-min isotropic etch. At this point, the structure is released from the underlying silicon and the desired cavity is formed. In the figure, we see a CMOS-MEMS beam and the metal layers inside which can be used as electrodes for sensing and actuation, or wires for connecting to the on-chip circuitry.

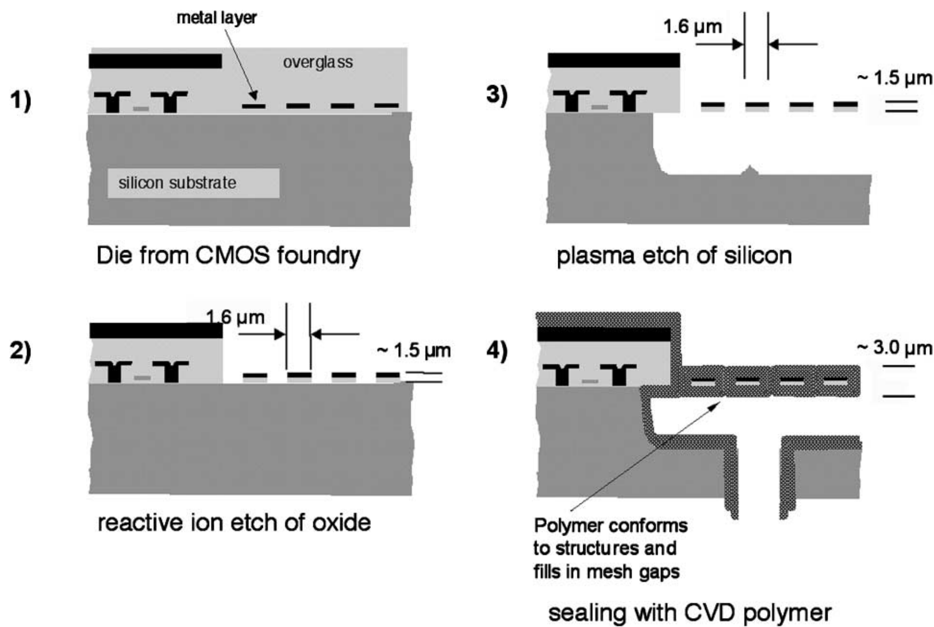

Fig. 4. Process sequence for constructing the CMOS-MEMS membrane.

4. In the final step, the released CMOS-MEMS structure is coated with polymer in a chemical vapor deposition process. The polymer conforms to all sides of the beams, until all the gaps are sealed, creating an airtight membrane suspended over the gap.

4. Simulations and calculations

A simplified acoustic model was developed to investigate the effects of the design parameters on the behavior of the MEMS earphone (Fig. 5a and b). The wavelengths of interest (audio frequencies) are much greater than the size of both the device and the ear canal, so we may ignore the effects of wave propagation and model the system as discrete components (the sound pressure is assumed uniform throughout each of the air volumes). The volumes of air trapped in the sub-membrane cavity and the ear canal comprise compliances C_1 and C_2. The backside vent hole and the leak between the ear canal and the device form the resistances R_1 and R_2. Both the inertance of the air in the vent and the viscosity of the air are small enough to be ignored.

In the equivalent electrical model, we model the sound pressures as voltages and airflow (volume velocity) as current. U is the volume velocity of the air moving into the ear canal, and y is the displacement of the membrane (positive towards the substrate). S is the effective area of the membrane, which takes into account the curved bulge shape. The current source $U = -j\omega Sy$ is dependent on the signal pressure under the membrane (p') and in the ear canal (p), as well as the electrostatic force F applied to the membrane. These forces interact dynamically with the mass m and

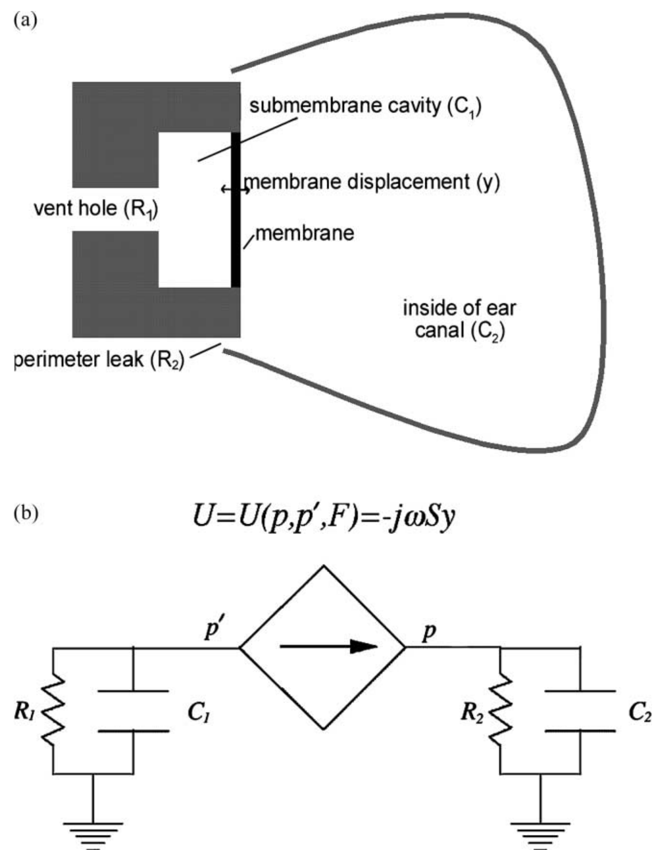

Fig. 5. (a) Drawing of the CMOS-MEMS earphone concept. Parameters for the acoustic calculation are shown on the figure near the relevant parts. (b) Electrical equivalent model of acoustic system. p' is the pressure (deviation from normal atmosphere) under the membrane, and p is the pressure in the ear canal. Dependent current source U depends on pressures p' and p, the mass and stiffness of the membrane structure, and the (real-life) electrical signal applied to the membrane.

stiffness k of the membrane as described by Newton's Second Law (in the frequency domain),

$$-m\omega^2 y = (p' - p)S - ky + F$$

R_1, C_1, R_2, and C_2 are used to calculate p and p' in a way exactly analogous to the voltages in an electrical circuit. The sound pressures p and p' are related to the air volume velocity U through the acoustic impedance $Z = p/U$, and the acoustic resistances and compliances are calculated from the device geometry and the air properties:

$$R_1 = \frac{\rho c}{A}, \qquad C_1 = \frac{V_1}{\rho c^2}$$

where A is the cross-sectional area of the backside vent holes, and V_1 is the volume of air trapped between the membrane and substrate.ρ is the density of air, and c is the sound speed of air at standard temperature and pressure. R_2 and C_2 are calculated in a similar fashion. Typical values are $R = 4E5$ acoustic ohms (g cm^{-4} s^{-1}) for the vent holes, and $C = 1E6$ s/(acoustic ohm) for the compliance of the ear canal.

It was assumed that the device would be placed in a 2 cm^3 ear canal, with some leakage (modeled by R_1) around the device into the outside world. The prototype chips have a membrane that is 1442 μm across. The membrane was simplified and assumed to be a single sheet of material of density 1.4E−3 ng/μm^3 and Young's modulus 3000 MPa. These values are typical of other polymers, and as the material properties of the particular polymer deposited are being measured, they were taken as a reasonable starting point. Simulation of moderate to large sections of the serpentine mesh required more disk space than was available, so no attempt was made to calculate an effective stiffness for the mesh. The serpentine mesh by itself was examined experimentally and found to be extremely compliant, so we further assumed the stiffness of the polymer dominates the behavior of the membrane.

5. Results

Optical measurements were made of static deflection as a function of voltage (Fig. 6) in order to determine an operating voltage, in this case 67 V. The deflection shown in the graph is given relative to the zero-volt position of the membrane, which may be several microns above the chip surface because of the buckling or bulging mentioned above. The chip was mounted on a TO-8 package and wire bonded. The TO-8 package was then epoxied inside the housing from a Radio Shack 33-175B earphone in such a way as to seal off the front of the TO-8 package from the back (Fig. 3).

As the targeted use of this MEMS device is as an earphone, we chose to measure the acoustic behavior in a Brüel and Kjaer (B&K) 4157 Ear Simulator rather than at an arbitrary distance from a free-field reference microphone such as the B&K 4939 (formerly 4135). This presents a more realistic acoustic load to the device, and more sensitivity for small sound pressures. The plastic earphone shell was pressed into the rubber ring adaptor supplied with the B&K 4157. This gave a consistent position and air volume between the measurements. The earphone-measurement microphone pair was put inside a B&K 4232 anechoic test chamber that provides about 40–50 dB isolation.

The responses of both the unmodified Radio Shack earphone and the MEMS device mounted in the Radio Shack shell were measured (Fig. 7a and b). The Radio Shack earphone was driven directly from the function generator with 50 mV-peak signal. The MEMS earphone was driven with a 14.3 V-peak signal on the top of a 67 V dc bias. Two

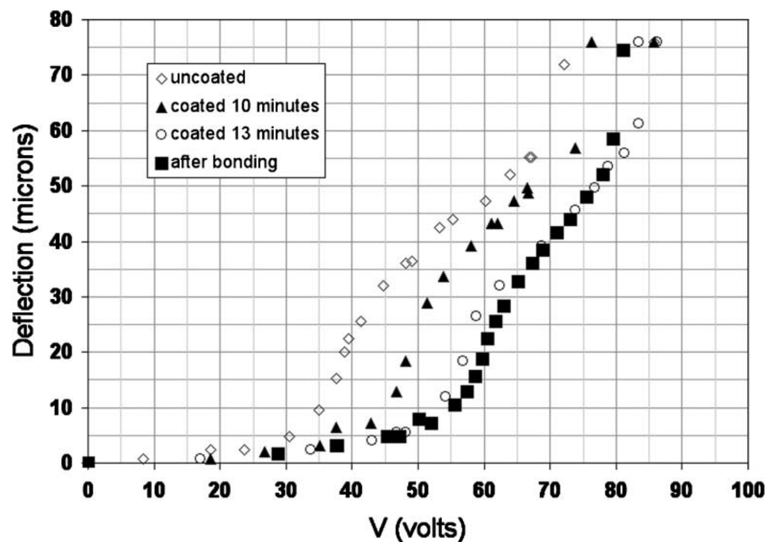

Fig. 6. Graph of static deflection as a function of applied electrostatic potential. Zero deflection corresponds to the rest position of the membrane center, which may be above the chip surface because of the slight (10 μm) buckling.

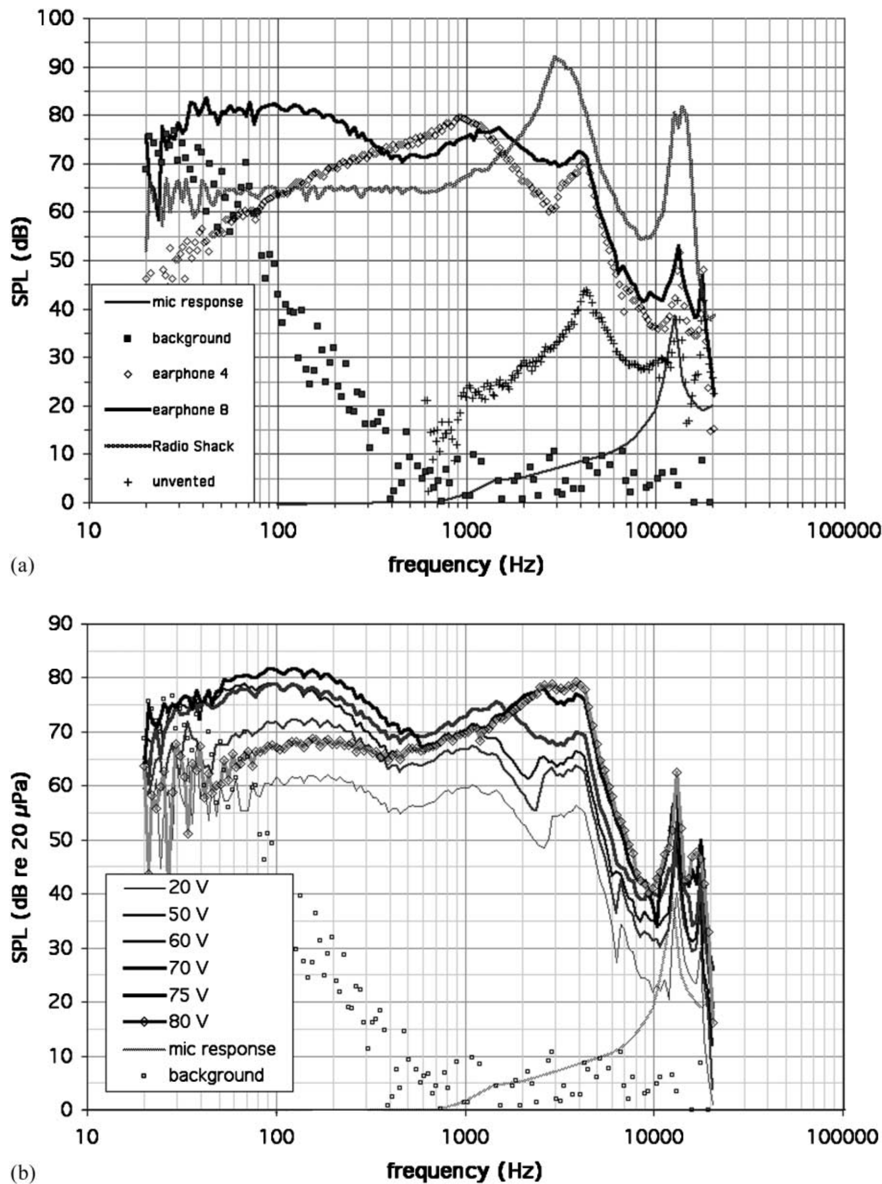

Fig. 7. (a) Graphs of responses of two CMOS-MEMS earphones, compared with a Radio Shack 33-175B earphone. The MEMS earphones are driven with 9.75 V rms on top of a 67 V bias; the Radio Shack earphone is driven with 35 mV rms directly from the function generator. Six earphones were measured and two of these were chosen to show the envelope of variation. Also shown are the response of a CMOS-MEMS earphone without vent holes, and the microphone correction and the background noise. The peak at 4 kHz (3 kHz for the Radio Shack earphone) is due to a resonance within the earphone housing's air volume; (b) Graphs of response as a function of varying bias voltage.

of the prototypes, earphones #4 and #8, were chosen to represent the extremes in the variation between devices. However, four of the six earphone responses fell within a range close to earphone #8. The peaks in the spectra near 4 kHz (3 kHz for the Radio Shack earphone) are due to the air volume trapped in the earphone shell. When the original diaphragm is taken out and the TO-8 package is inserted, this volume is decreased, leading to the upward shift in frequency of the peak relative to the unmodified Radio Shack earphone. This peak is in nearly the same place (within 5%) for the earphone without vent holes, despite the 25 dB difference in output between the unvented earphone and earphones #4 and #8, supporting the idea that this resonance depends only on the plastic housing geometry. The peaks near 12 kHz are due to the frequency response of the microphone (ear simulator) itself, which mimic the resonance of the ear canal. This peak's center frequency is also sensitive to the trapped air volume between the earphone diaphragm and the ear simulator's microphone, so it may not coincide with the peak in the calibration data supplied with the microphone. For this reason, we have chosen to present the data without correction for the microphone response to avoid misinterpretation.

Fig. 7b illustrates the effect of varying bias on the frequency response. The overall magnitude of the response increases, as the bias voltage is increased from 20 V to about

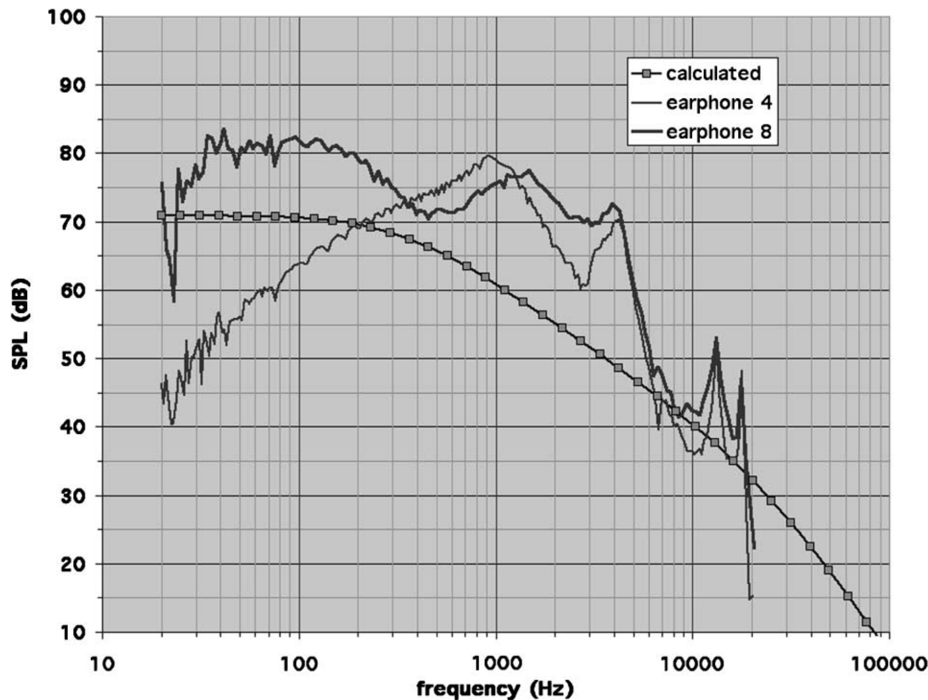

Fig. 8. Graph showing measured earphone responses alongside predictions based on our simplified acoustic model. The measured data are the responses of two of the MEMS earphones with 1.4 mm membranes. Model parameters were chosen which placed the calculated curve close to the two sets of data.

70 V. At around 75 V bias, a qualitative change in behavior occurs, with a broad peak forming between 2 and 4 kHz. Further increase of the bias to 80 V completes the transition, with a sudden drop in the low-frequency (20–200 Hz) response of about 15 dB, while the 3 kHz peak remains. Though not shown on the graph, we also perceived with our ears an increase in distortion at the higher input signal levels in this region of operation. We believe that this change in behavior corresponds to the snapping down of the membrane between its two bistable states (slightly convex and slightly concave), which occurs around 75–80 V in most of the chips we studied.

Simulation results (Fig. 8) were compared with the experimental data to help us to estimate the polymer material properties. The parameters of density and Young's modulus were adjusted in order to bring the simulation result in line with the experimental data. The set of known quantities for the device include the etch depth (70 μm), membrane size (1442 μm^2), applied voltage (67 V bias + 9.75 V rms signal), approximate membrane thickness (2 μm), and total vent hole area (450,000 μm^2). Other knowns include the density of air and the speed of sound. It was estimated that the earphone housing fills about half of the volume of a typical ear canal or our ear simulator, so we use 1 cm^3 for the ear canal volume. Unknown quantities include the material properties of the polymer, and the leak size around the perimeter of the device. As initial trial values, we used a density of 1.4E−3 ng/μm^3 for the polymer density and 3000 MPa for the Young's modulus. We assumed that the observed peaks near 4 and 12 kHz are artifacts of the earphone housing and microphone response (as explained earlier), so we did not try to match these. For the leak area, we started with no leak and increased the leak area.

Fig. 8 illustrates the results of the model calculations. Starting with the parameters described above, we adjusted the unknowns so as to make a compromise between the two sets of data shown. First, the leak area around the device appears to be very small, as increasing it above zero decreases the overall calculated response of the earphone far below the measured response. This is to be expected since the earphone shell fits tightly into the rubber gasket supplied with the ear simulator. Second, the effect of polymer stiffness was investigated. It was possible to raise or lower the low-frequency (20–200 Hz) part of the calculated curve by varying the Young's modulus. This response at 80 Hz varies from about 62 dB SPL at 9000 MPa Young's modulus and increases to about 78 dB SPL at 1000 MPa. Both of these values lie within the experimentally measured range for the set of earphones. In a similar fashion we varied the polymer density from 1E−3 ng/μm^3 to 5E−2 ng/μm^3. As the density was increased, the high-frequency rolloff became steeper, and the corner frequency decreased in frequency and became better defined. At 5E−2 ng/μm^3, the predicted corner frequency is about 300 Hz. The corner frequency of the measured earphones is ambiguous, especially given the fact that the housing itself seems to cause a high-frequency rolloff in the response of both the Radio Shack and MEMS earphones. We found that a significant range of values bracketing the initial estimates of the polymer

material properties lead to calculated results that fall within the range of the experimental data. Thus it is impossible to specify the polymer material properties more accurately than our initial estimates. However, this suggests that the earphone design is robust and relatively insensitive to the specific material property values.

6. Summary

The demonstrated earphone prototype has opened the door to a variety of potential applications in a variety of areas. With some refinement and additional research, high-performance microspeakers and microphones are achievable. Arrays of microspeakers will allow greater acoustic power and range of frequency response, as well as applications such as beam steering. Arrays of microphones will allow lower-noise operation and directionality. Adjusting the geometry of the device (gap, membrane size and thickness) may tailor it for other uses, such as ultrasonic transduction. On-chip electronics allow both sensitivity and new array architectures for a range of consumer and defense acoustic applications. We are currently fabricating and testing a 3 mm membrane. Based on our simulations, we expect that large, compliant membranes will make sensitive microphones and louder microspeakers.

Acknowledgements

The authors would like to thank Lars Erdmann for advice and assistance in processing and Brett Diamond for comments on the manuscript. This project was supported in part by DARPA contract No. 00019-98-K-0817, and in part by the Pennsylvania Infrastructure Technology Alliance, a partnership of Carnegie Mellon, Lehigh University, and the Commonwealth of Pennsylvania's Department of Economic and Community Development. It was also supported in part by a grant from Adtranz.

References

[1] G.K. Fedder, S. Santhanam, M.L. Reed, S.C. Eagle, D.F. Guillou, M.S.-C. Lu, L.R. Carley, Laminated high-aspect-ratio microstructures in a conventional CMOS process, Sens. Actuat. A 57 (1996) 103–110.

[2] S.S. Lee, R.P. Ried, R.M. White, Piezoelectric cantilever microphone and microspeaker, J. MEMS 5 (4) (1996) 238–242.

[3] C.H. Han, E.S. Kim, Fabrication of piezoelectric acoustic transducers built on cantilever-like diaphragm, in: Proceedings of the 14th IEEE International Conference on Micro Electro Mechanical Systems (MEMS 2001), Interlaken, Switzerland, 21–25 January 2001, pp. 110–113.

[4] C.H. Han, E.S. Kim, Parylene-diaphragm piezoelectric acoustic transducers, in: Proceedings of the 13th IEEE International Conference on Micro Electro Mechanical Systems (MEMS 2000), Miyazaki, Japan, 23–27 January 2000, pp. 148–152.

[5] H. Xie, G.K. Fedder, A CMOS z-axis accelerometer with capacitive comb-finger sensing, in: Proceedings of the 13th IEEE International Conference on Micro Electro Mechanical Systems (MEMS 2000), Miyazaki, Japan, 23–27 January 2000, pp. 496–501.

[6] M.S.-C. Lu, X. Zhu, G.K. Fedder, Mechanical property measurement of 0.5-μm CMOS microstructures, Mat. Res. Soc. Symp. Proc. 518 (1998) 27.

Biographies

John J. Neumann Jr. is a postdoctoral researcher in the Carnegie Mellon University MEMS Laboratory. He received a BS in physics (1990) from the State University of New York at Buffalo, and a PhD in theoretical nuclear physics (1996) from Kent State University, Ohio. Before joining the Carnegie Mellon University MEMS Lab, he was Assistant Professor of physics at Hendrix College, Conway, Arkansas. His research interests center on acoustic applications of MEMS, and CMOS MEMS in particular.

Kaigham J. Gabriel (Ken) is Professor of Electrical & Computer Engineering and the Robotics Institute at Carnegie Mellon University. He received his SM and PhD in Electrical Engineering and Computer Science from the Massachusetts Institute of Technology. In 1985, he joined AT&T Bell Labs in the Robotic Systems Research Department, where he pioneered the field of MEMS and started the silicon MEMS effort, and led a group of researchers in exploring and developing IC-based MEMS for applications in photonic and network systems. During a sabbatical year from Bell Labs, Dr. Gabriel was a Visiting Associate Professor at the Institute of Industrial Science, University of Tokyo in Japan, where he led joint projects at IBM Japan Research, Toyota Central Research Laboratories and Ricoh Research Park. After leaving Bell Laboratories in 1991, he spent a year as a visiting scientist at the Naval Research Laboratory transferring micromechanics processing technology to the Nanoelectronics Processing Facility. From 1992 to 1997, Dr. Gabriel was at DARPA, where he started and managed the Microelectromechanical Systems (MEMS) Program (1992–1996) taking it from zero to over $70M/year with over 80 projects in 1997. He was then appointed as Deputy Director of the Electronics Technology Office (1995–1996), and finally as Director of the Electronics Technology Office (1996–1997), where he was responsible for roughly half of the Federal electronics technology investments totaling over $400M/year and spanning programs in advanced lithography, electronics packaging, MEMS, optoelectronics, millimeter and microwave integrated circuits, and high-definition displays.

Dr. Gabriel counts among his honors the Carlton Tucker Prize for Excellence in Teaching from MIT, his appointment to the Senior Executive Service, Co-Chairing the Defense Science Board 1999 Summer Study on 21st Century Technology Trends and Evolution, and two personal technology briefings to then Secretary of Defense William Perry.