

# CMOS-MEMS MEMBRANE FOR AUDIO-FREQUENCY ACOUSTIC ACTUATION

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

Electrical and Computer Engineering Department, Carnegie Mellon University

Pittsburgh PA 15213 USA

email: jneumann@ece.cmu.edu, kgabriel@ece.cmu.edu

## ABSTRACT

Using CMOS-MEMS micromachining techniques we have constructed a prototype earphone that is audible from 1 to 15 kHz. The fabrication of the acoustic membrane consists of only two steps in addition to the prior post-CMOS micromachining steps developed at CMU [1]. The ability to build a membrane directly on 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. The large compliance of the CMOS-MEMS membrane also promises application as a sensitive microphone and pressure sensor.

## INTRODUCTION

The fundamental challenge of using MEMS for acoustic applications is how to achieve significant sound pressures with a device of 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, since the small volume of the canal presents a larger acoustic impedance than is achievable in free air.

Our goal is to create a high-performance, economical CMOS-MEMS earphone for speech or music playback. A strong low-frequency response is especially desirable for music. By keeping the number of custom fabrication steps low, complexity is decreased and yield maximized. Previous MEMS audio devices, such as dome-shaped actuators [2], have been made with large output and good frequency response, but require numerous custom fabrication steps. In contrast, the sealed membrane of the present CMOS-MEMS device is formed in a single chemical vapor deposition.

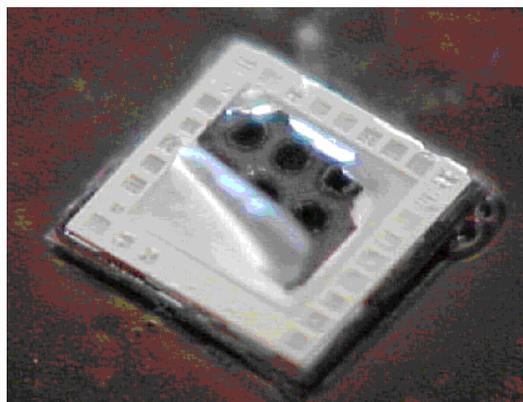
In addition to the simple fabrication sequence, the CMOS-MEMS acoustic devices have the advantage of on-chip electronics. In the case of a microphone, sensing circuitry (capacitive or piezoresistive) can be placed close to the mechanical structure, minimizing electromagnetic pickup and parasitic capacitance. Another advantage is that the CMOS-MEMS membranes can have as many different conductive layers as the CMOS process (typically 3), 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 which 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.

## CONCEPT AND METHODS

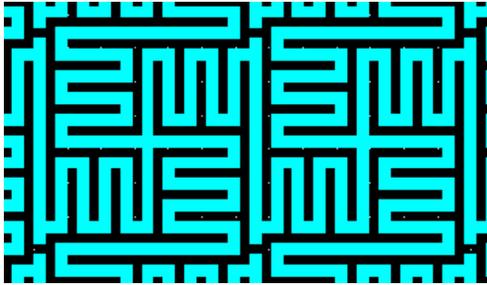
To construct a speaker or microphone in CMOS-MEMS, we make an airtight membrane on the surface of the chip by defining a mesh or screen-like structure, much like the proof mass in CMOS accelerometers [3]. This mesh is covered with polymer to make it air-tight, and the CMOS metal layer in the mesh structure is set to a varying electrical potential with respect to the silicon substrate, causing motion out of plane.

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 that the membrane has room to vibrate with the desired amplitude, but not so large that an unreasonably high voltage is required to operate. Our cavities were typically 60 to 100  $\mu\text{m}$  deep.



*Figure 1. Photograph of speaker chip, showing peeled-back membrane, and underlying cavity and vent holes.*

The device (Figure 1) is placed in the ear with the membrane facing into the ear canal. The back side 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. 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 back side of the membrane (allowing greater displacements for a given electric force), and adding a resistive component which damps out unwanted resonances.

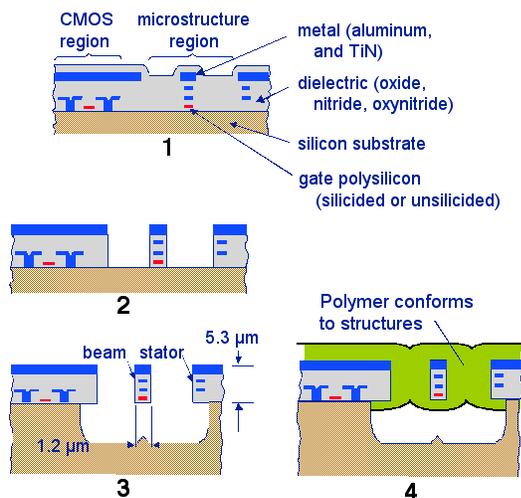


**Figure 2.** Layout of serpentine mesh design, showing two unit cells.

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 [4]. For a 50  $\mu\text{m}$  cantilever beam made of metal and oxide, the out of plane curl can be as much as 1  $\mu\text{m}$ , and increasing with the square of the length. To solve this problem, a serpentine mesh design was developed (Figure 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 ( $<50 \mu\text{m}$ ) so as to limit curl. The gaps between the beams are of uniform size, allowing efficient and uniform coating with polymer. In the chips that were fabricated, both the gaps and beam widths were 1.6  $\mu\text{m}$ . The maximum out-of-plane buckling is between 10 and 20  $\mu\text{m}$  for the 1.4 mm by 1.4 mm prototype membrane.

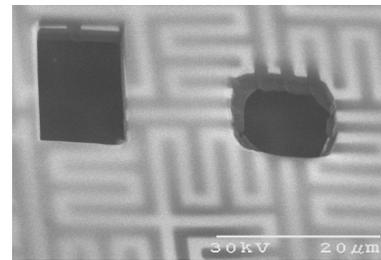
### FABRICATION PROCESS

Before the patterning of the CMOS structures on the front side begins, the backside vent holes must be etched, using the oxide layer as an etch-stop. The hole sizes are about 100 microns across, with total areas of multiple hole patterns (3x3 grid) being around 50,000 to 100,000  $\mu\text{m}^2$ . After the vent holes are etched, processing of the CMOS (front) side of the chip follows the steps shown in Figure 3:



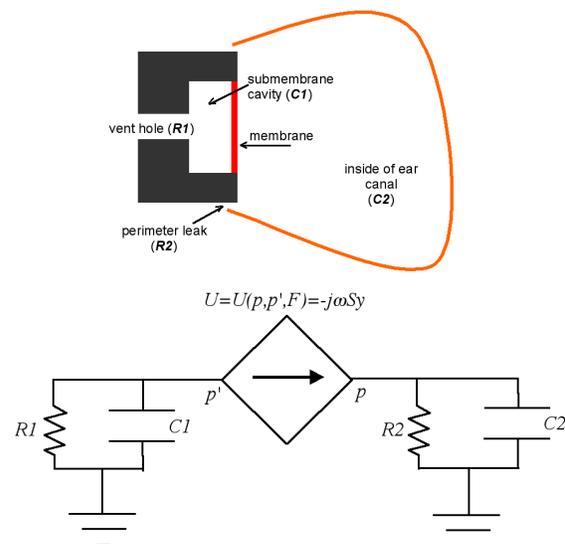
**Figure 3.** Processing steps for CMOS-MEMS earphone.

- (1) The chip comes from a CMOS fabrication foundry (e.g. MOSIS) covered with a layer of protective glass (silicon dioxide). Regions meant for mechanical structures are patterned in one of the metal layers, usually the topmost layer.
- (2) The oxide is etched anisotropically (directionally) 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-minute DRIE (Bosch) anisotropic deep etch followed by a 7-minute isotropic etch. We found the membrane was less likely to stick to the bottom after actuation than after only an isotropic etch, perhaps because of a rougher surface (about 4 micron roughness) on the bottom of the cavity. At this point, the structure is released from the underlying silicon and the desired cavity is formed. In Figure 3 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.



**Figure 4.** SEM of serpentine mesh sealed with polymer, with two holes from focussed ion beam.

- (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 (see Figure 4). The metal layers inside the beams allow the membrane to be treated as an electrode for electrostatic actuation.



**Figure 5.** Simplified drawing of device and equivalent electrical model.

## SIMULATIONS AND CALCULATIONS

A simplified acoustic model was developed to investigate the effect of the design parameters on the behavior of the MEMS earphone (Figure 5). 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, i.e. 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  $C1$  and  $C2$ . The backside vent hole and the leak between the ear canal and device form resistances  $R1$  and  $R2$ . Both the inertance of the air in the vent and the viscosity of the air are small enough to be ignored.

In the electrical equivalent model, we model sound pressures as voltages and air flow (volume velocity) as current.  $U$  is the volume velocity of the air moved into the ear canal, and  $y$  is the displacement of the membrane (positive toward the substrate).  $S$  is the effective area of the membrane, which takes into account the curved bulge shape. The current source  $U = -j\omega S y$  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 stiffness  $k$  of the membrane as described by Newton's 2<sup>nd</sup> law (in the frequency domain),

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

$R1$ ,  $C1$ ,  $R2$ , and  $C2$  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 air properties:

$$R1 = \frac{\rho c}{A}, \quad C1 = \frac{VI}{\rho c^2}$$

where  $A$  is the cross-sectional area of the backside vent holes, and  $VI$  is the volume of air trapped between the membrane and substrate.  $\rho$  is the density of air, and  $c$  is the sound speed of air at standard temperature and pressure.  $R2$  and  $C2$  are calculated in a similar fashion. Typical values are  $R = 4e5$  acoustic ohms ( $\text{g cm}^{-4} \text{sec}^{-1}$ ) for the vent holes, and  $C = 1e6 \text{ sec}/(\text{acoustic ohm})$  for the compliance of the ear canal.

It was assumed that the device would be placed in a  $2 \text{ cm}^3$  ear canal, with some leakage around the device into the outside world. The width of the leak is assumed to be  $500 \mu\text{m}$ , around the perimeter of the device's housing, which is assumed to be  $4 \text{ cm}$ . The prototype chips have a membrane that is  $1442 \mu\text{m}$  across. The membrane, which in reality is composed of the serpentine mesh and polymer, was simplified to a single sheet of material of density  $1.4e-3 \text{ ng}/\mu\text{m}^3$ , Young's modulus  $3000 \text{ MPa}$ . These values are typical of other polymers, and as the material properties of the particular polymer deposited are not yet known, 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 assume the stiffness of the polymer dominates the behavior of the membrane.

## 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 volts. 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 to 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.

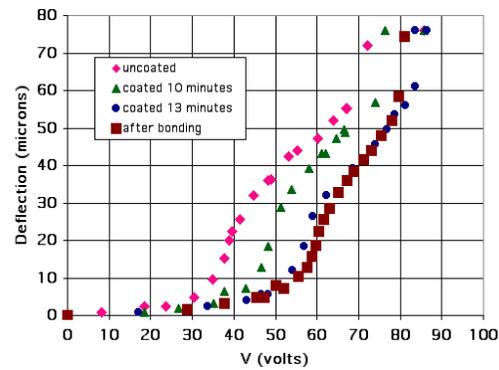
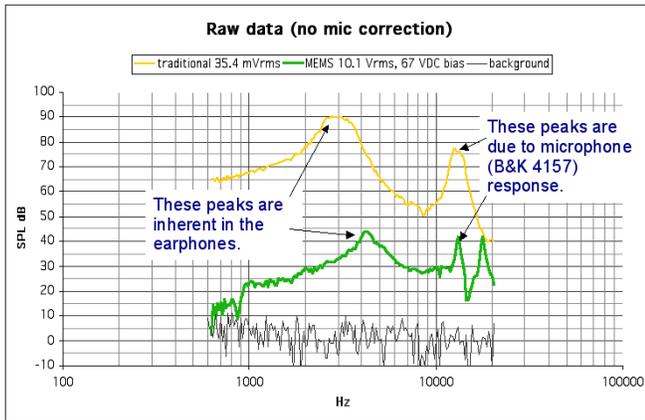


Figure 6. Static deflection of membrane vs. voltage, for various thicknesses of polymer.

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 4157. This gave a consistent position and air volume between measurements. The earphone-microphone pair was put inside a B&K 4232 anechoic test chamber which provides about 40-50 dB isolation.

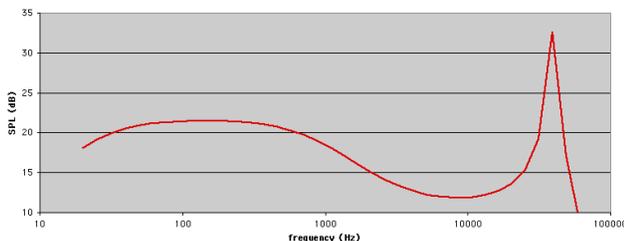
The response of both the unmodified Radio Shack earphone and the MEMS device mounted in the Radio Shack shell were measured (Figure 7). The Radio Shack earphone was driven directly from the function generator with 50 mV-peak signal. The MEMS earphone was driven with a 14.3 volt-peak signal on top of a 67 volt DC bias. The peaks in the spectra near 3-4 kHz 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. 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



**Figure 7.** The MEMS earphone in its housing is compared with a traditional earphone (Radio Shack 33-175B). Resonant peaks are due to the earphone housing air volume (3-4 kHz) and microphone response (12 kHz).

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 the correction (0 dB around 600 Hz up to 19 dB at 10 kHz.) for the microphone response to avoid misinterpretation.

Simulation results were compared with the experimental data to help us 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 20 dB SPL at 1000 Hz and 10 dB SPL at 10 kHz. Matching far below or far above this range was ignored because it is beyond the range of measurement. We assumed that the observed peaks near 4 kHz and 12 kHz are artifacts of the earphone housing and microphone response (as explained earlier), so we did not try to match these. We found that the simulation results were insensitive to variations in the Young's modulus, with values from 1000 to 9000 MPa all yielding similar results. This suggests that the air trapped under the membrane (compliance  $CI$ ) dominates the stiffness of the membranes. The density of the polymer also has a weak effect on the simulated frequency response, where a 5 fold change in density only produces about 1.5 to 2 fold change in resonance frequency. The predicted resonant frequency for a typical polymer density of  $1.4 \times 10^{-3} \text{ ng}/\mu\text{m}^3$  is 40 kHz, well above the range of the measurement. Setting the density of the polymer in the simulation to  $5 \times 10^{-3} \text{ ng}/\mu\text{m}^3$  causes a resonance near 18 kHz, about twice the expected frequency based on the membrane in a vacuum. It



**Figure 8.** Sound pressure level predicted by simulation. Design parameters are listed in text.

appears then that the resonances of the device are dominated at least as much by the acoustic environment (packaging) as by the properties of the membrane.

In contrast to the polymer material properties, the effect of the vent holes in the simulation was dramatic, with a 500  $\mu\text{m}$  square vent hole increasing the SPL from 20 dB to 40 dB at 1000 Hz. The simulation results which match our measurements assumed a single 100  $\mu\text{m}$  square hole, which is smaller than what was really used. This suggests that our vent holes were not as effective as they should be, possibly because of acoustic short circuiting around the chip. This explanation is also supported by the large (1 mm) perimeter leak which gives this result.

Figure 8 shows the result of the simulation for a membrane size of 1442  $\mu\text{m}$  square, etch depth of 88  $\mu\text{m}$ , 4  $\mu\text{m}$  membrane thickness, polymer density of  $1.4 \times 10^{-3} \text{ ng}/\mu\text{m}^3$ , polymer Young's modulus of 3000 MPa, ear canal volume of 2  $\text{cm}^3$ , total perimeter leak area of 40  $\text{mm}^2$ , vent hole size of 100  $\mu\text{m}$  square, DC bias of 67 volts, and signal of 10.1 volts RMS.

## SUMMARY

The working earphone prototype is exciting because of the variety of applications proven possible. With some refinement, high-performance microspeakers and microphones are achievable. 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 "intelligence" for a range of consumer and military uses.

It remains to see the effect of a larger membrane size on behavior. Large, compliant membranes will make sensitive microphones and louder microspeakers. Calculations, using parameters extracted from the above comparison of simulation and experimental results, predict higher (30-40 dB for a 3 mm square membrane) sound pressures and improved low-frequency response.

## ACKNOWLEDGEMENTS

This project was supported by DARPA.

## REFERENCES

1. G.K. Fedder, S. Santhanam, M.L. Reed, S.C. Eagle, D.F. Guillou, M.S.-C. Lu and L.R. Carley, "Laminated high-aspect-ratio microstructures in a conventional CMOS process," *Sensors and Actuators A* 57 (1996) 103-110.
2. C.H. Han and E.S. Kim, "Parylene-Diaphragm Piezoelectric Acoustic Transducers," *Proceedings of the 13<sup>th</sup> IEEE International Conference on Micro Electro Mechanical Systems (MEMS 2000)*, pp. 148-152, Miyazaki, Japan, Jan. 23-27, 2000.
3. H. Xie, G.K. Fedder, "A CMOS z-axis accelerometer with capacitive comb-finger sensing", *Proceedings of the 13<sup>th</sup> IEEE International Conference on Micro Electro Mechanical Systems (MEMS 2000)*, pp. 496-501, Miyazaki, Japan, Jan. 23-27, 2000.
4. M. S.-C. Lu, X. Zhu, G.K. Fedder, "Mechanical Property Measurement of 0.5- $\mu\text{m}$  CMOS Microstructures", *Mat. Res. Soc. Symp. Proc.* Vol. 518, p. 27 (1998).