J. Micromech. Microeng. **10** (2000) 380–386. Printed in the UK PII: S0960-1317(00)12324-1

A fabrication process for integrating polysilicon microstructures with post-processed CMOS circuits

Y B Gianchandani†‖, H Kim‡, M Shinn‡, B Ha‡, B Lee‡, K Najafi§ and C Song‡

† Department of Electrical and Computer Engineering, University of Wisconsin, 1415 Engineering Drive, WI 53706-1691, USA
‡Samsung Advanced Institute of Technology, Suwon, Korea
§ Center for Integrated Microsystems, EECS Department, University of Michigan, Ann Arbor, MI 48109-2122, USA

E-mail: yogesh@engr.wisc.edu (Y B Gianchandani)

Received 3 March 2000

Abstract. A MEMS-first fabrication process for integrating CMOS circuits with polysilicon micromechanical structures is described in detail. The overall process uses 18 masks (22 lithography steps) to merge a p-well LOCOS CMOS process that has one metal and two polysilicon layers with a surface micromachining process that has three layers of polysilicon. The microstructures are formed within recesses on the surface of silicon wafers such that their uppermost surfaces are coplanar with the remainder of the substrate. No special planarization technique, such as chemical–mechanical polishing, is used in the work described here. Special aspects of the process include provisions to improve lithography within the recesses, to protect the microstructures during the circuit fabrication, and to implement an effective lead transfer between the microstructures and the on-chip circuitry. The process is validated using a test vehicle that includes accelerometers and gyroscopes interfaced with sensing circuits. Measured transistor parameters match those obtained in standard CMOS, with NMOS and PMOS thresholds at 0.76 V and −0.96 V, respectively.

1. Introduction

The integration of circuitry with surface micromachined polysilicon microstructures has been traditionally performed by first fabricating the circuitry, and then depositing and patterning the structural polysilicon [1–3]. This approach offers the advantage that large topographical variations associated with the micromechanical structures do not affect the lithography for fabricating the circuitry. However, it limits the thermal budget of the deposition and annealing of the micromechanical structures to whatever small amount the circuitry can accommodate. This constraint has motivated efforts to reverse the traditional approach and fabricate the micromechanical structures first [4–6]. The rationale for these approaches is that the micromechanical structures are less affected by the thermal budget of the circuit fabrication than the circuit is affected by that of the microstructure fabrication [7]. This issue is discussed further in section 3.

This paper describes a MEMS-first fabrication process for integrating CMOS circuits with surface micromachined polysilicon structures, in the same vein as [4–6]. Micromechanical structures are formed within recesses on the surface of silicon wafers such that their uppermost surfaces are coplanar with the remainder of the substrate. The motivation for this effort is to develop a process that (i) can accommodate a conservative CMOS process with a relatively high thermal budget and (ii) can be implemented with a relatively small tool set. For example, all the lithography in this process is performed with contact aligners. Additionally, no special planarization technique such as chemical–mechanical polishing (CMP) is used. The fabrication process is detailed in section 2. Section 3 addresses the impact of the thermal budget of the circuit fabrication upon the structural polysilicon. Experimental results are presented in section 4, and the conclusions that can be drawn from this effort are summarized in section 5. The preliminary results of this work are discussed in [8].

2. Fabrication

The overall process uses 18 masks (22 lithography steps) to merge a p-well LOCOS CMOS process that has one metal and two polysilicon layers with a surface micromachining process that has three layers of polysilicon. The number of masks is less than the number of lithography steps because one

‖ Author to whom correspondence should be addressed.

0960-1317/00/030380+07$30.00

particular mask with very large feature sizes and non-critical alignment is used repeatedly. Special aspects of the process sequence include the use of this mask, the inclusion of steps to ensure the protection of the microstructures during circuit fabrication and the lead transfer between the microstructures and the circuits.

The layer thickness and etch depths in the following description are provided as guidelines only. Many different values were used in our experiments; the numbers presented are typical and representative, but scalable.

The fabrication process is summarized in figure 1. It uses n-type (100) wafers with 1.0–1.5 Ω cm resistivity, as required for the CMOS transistors. Portions of the wafers are recessed by an KOH etch to house the microstructures. This depth is adjusted, as necessary, to equal the total thickness of all the structural layers and sacrificial layers between them. For example, wafers for accelerometers (which have two levels of microstructural polysilicon) are recessed by 6 μm, whereas those for gyroscopes (which have three levels of structural polysilicon) are recessed by 9 μm. Unrecessed regions of the wafers are protected from the KOH by thermal oxide. Trichloroethane (TCA) is used in the thermal oxidation to minimize the generation of surface defects in the circuit regions of the wafer. Following the recess etch the masking oxide is stripped and a double layer of dielectrics is deposited. This includes, in order, 1 μm thermal TCA oxide for electrical isolation of the microstructures and 3000 Å of LPCVD nitride to serve as an etch-stop while wet etching the sacrificial oxide in subsequent steps. The thickness of the nitride is intended to compensate for wear in subsequent processing, such as its gradual removal during the wet-etch in buffered hydrofluoric acid (BHF) when patterning the sacrificial oxides above it. It is important to retain some of the nitride through the entire process in order to protect the 1 μm thermal oxide beneath it during the release etch. It is also needed to ensure that this oxide is not etched away prematurely from the regions where the circuitry will be located, exposing them to plasma damage, etc.

Ideally, the first layer of polysilicon would be placed directly above the nitride. However, the polysilicon is dry etched using a recipe that aggressively attacks nitride. Thus, an optional 200 Å thick layer of LPCVD oxide is placed as an etch-stop between the nitride and the polysilicon. The thickness of this layer is constrained to 200 Å to minimize the undercutting of the first polysilicon during the final release etch.

The first polysilicon is deposited, doped and patterned for use in interconnect and for electrodes within the recess. It is only 3000 Å thick, and does not significantly affect the topography. A relatively high value of sheet resistance can be accommodated in this layer since microstructures in the test vehicle use capacitive transduction, so it does not sustain any dc current.

The line widths on the first polysilicon layer are relatively large and non-critical, so the lithography at the trench bottom does not pose a problem. However, it is worth noting that resist thickness can vary near the recess edges if the resist is spun on and the recesses are deep. In order to minimize the potential impact of this variation upon lithography, circuits and microstructures with critical dimensions are located more than 100 μm from recess edges. Spun-on resist with

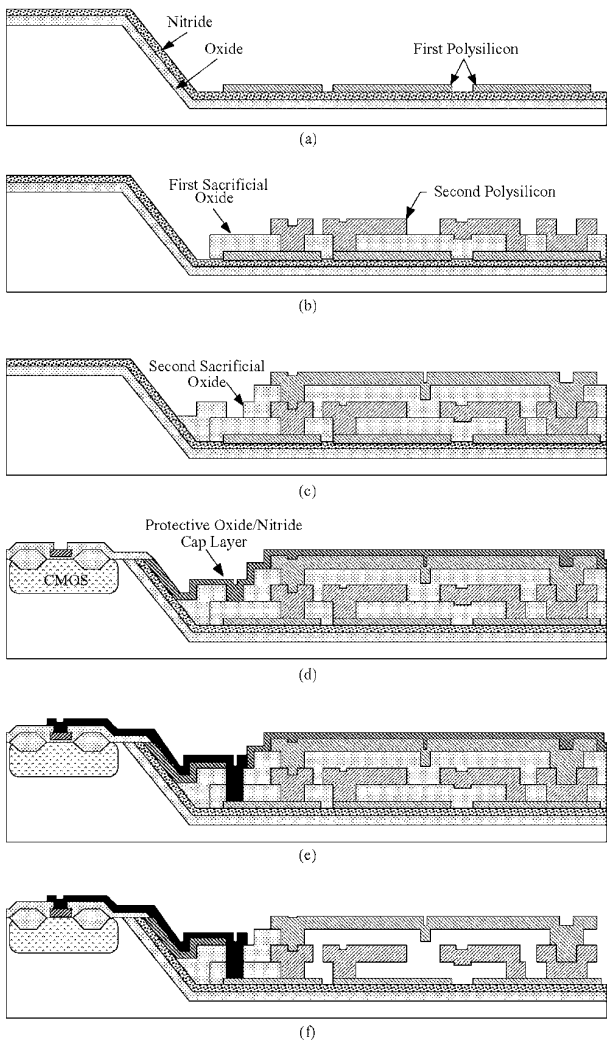


Figure 1. A schematic diagram summarizing the process flow. (a) Create trench using KOH and a thermal oxide mask, deposit bottom oxide and nitride insulation layers, deposit and pattern first polysilicon layer. (b) Deposit and pattern first sacrificial oxide layer, deposit and pattern second (microstructural) polysilicon layer. In both of these steps the films are stripped from the circuit region before patterning the layers at the bottom of the trench. (c) Deposit and pattern second sacrificial oxide, deposit and pattern third polysilicon layer. (d) Strip the nitride from unrecessed regions using hot phosphoric acid, deposit and pattern the top protective cap layer consisting of the oxide and nitride to protect the MEMS parts, proceed with the complete circuit process until the metal interconnect layer. (e) Open contact holes in the MEMS regions, deposit and pattern aluminum for circuit interconnect and lead transfers between circuit and MEMS regions. (f) Cover aluminum metallization with a protective layer, remove protective cap layer from the MEMS regions, etch all sacrificial layers in BHF, remove protective layer from circuit regions.

nominal thickness from 1.3 μm to more than 3 μm has been successfully used throughout this fabrication sequence.

After patterning the first polysilicon layer the first sacrificial oxide is deposited (2 μm thick). It is patterned using a two-step process: a clear-field version of the recess mask is first used to selectively wet etch the oxide outside the recess. The minimum feature size on this mask is over 100 μm, and the alignment is non-critical, so it does not

significantly add to the complexity of the overall sequence. This serves to reduce the step height between the upper surface of the remaining oxide and the top surface of the wafer. This oxide is then patterned by dry and/or wet etching in the second step. The two-step process improves the spatial resolution that can be achieved with contact aligners, which have a very limited depth of focus. It is used to pattern the first and second sacrificial oxides, as well as the second and third structural polysilicon layers. Consequently, up to four non-critical alignment steps and one mask are added to the process to circumvent the depth of focus limitations in the lithography equipment.

The second polysilicon layer is the primary structural layer, and forms the proof mass and suspension for both accelerometers and gyroscopes. It is 3 μm thick and can easily be scaled up or down, within the constraints of the device design and the deposition technique. It is reached electrically by a buried contact from the first polysilicon layer. In the gyroscope process the second polysilicon is followed by the second sacrificial oxide and the third polysilicon layer, each 2 μm thick. The polysilicon forms the suspended bridge electrode above the proof mass. Its anchors are coincident with those of the second polysilicon layer to reduce their vertical step height.

The dry etch for the second polysilicon should be highly anisotropic so that lateral capacitive transduction can be used effectively. There is no problem with stringers in patterning the second polysilicon because the first polysilicon layer that lies below it is very thin. However, lithography and patterning of the third polysilicon layer require special considerations. The third polysilicon layer is conformal to the substantial topographical variations created by patterning the second polysilicon. Portions of the third polysilicon that must be etched away lie within narrow trenches, between the tines of comb actuators defined in the second polysilicon. If the etch mask is defined using a clear-field pattern for this step, the duration of the exposure required for the trench bottoms causes the resist on the upper surfaces to be substantially overexposed. This problem is alleviated by using the lift-off technique and a dark-field pattern for a thin-film metal etch mask. Note that the dry etch used to pattern the third polysilicon must have an isotropic component to remove stringers. The lateral dimensions between features on this layer are less critical in our applications, since it is used as an electrode plate for the proof mass below it in the second polysilicon layer. The electrode gap is not affected by the isotropic etch, whereas the area of the electrode is compensated in the layout. Additionally, it is important to note that stringers may arise not only from the third structural polysilicon, but also from the two levels of polysilicon used by the CMOS. These must also be etched isotropically in order to ensure complete removal from undesired locations. Stringers that block BHF access holes in the structural polysilicon layers can easily double the duration of the release etch, and if not removed during the release, they can also prevent the device from operating.

After depositing the last structural polysilicon layer the wafers are prepared for the CMOS process sequence. By this point the 3000 Å thick nitride layer which was deposited early in the sequence is uneven and worn outside the recesses,

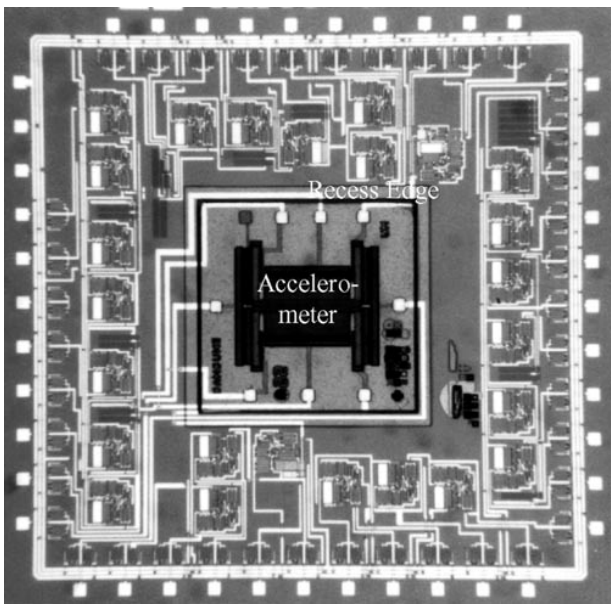


Figure 2. An optical micrograph of a capacitive polysilicon microaccelerometer in a recess, surrounded by CMOS sense circuitry.

since it is attacked by BHF when the sacrificial oxides are patterned. The wear is greater for the gyroscope wafers than for the accelerometer wafers, which have been through fewer processing steps at this point. This nitride is stripped by an unmasked wet etch in hot phosphoric acid, which does not attack the polysilicon microstructures or the oxide in the unrecessed regions. A capping oxide (1 μm) and nitride (1200 Å) are now deposited by LPCVD to protect the microstructures during CMOS processing. Note that if the nitride is too thick it will crack at high stress points (e.g. the corners of the microstructures) during p-well drive-in, which has a high thermal budget. In order to avoid this, it is necessary to stress compensate the dielectric cap or to use a low-stress nitride. The nitride is patterned by dry etching to remain only above the recessed regions. The oxide on the upper surface of the wafer, about 2 μm thick at this point, is thinned by a timed wet etch down to 1 μm before patterning it to expose the regions in which the p-well will be implanted. This step is performed to reduce the amount of lateral undercutting that occurs when it is patterned by wet etching. A wet etch is used to prevent plasma damage to the silicon surface in the p-well regions.

From this point on, the wafers are subjected to a standard LOCOS CMOS process until the contact etch prior to metallization [9]. The p-well implant is followed by a drive-in diffusion, which has the highest thermal budget of all the steps in the CMOS process, being performed at 1200 °C for 16 h in an ambient of nitrogen and oxygen. The microstructures are not exposed to the ambient gases because they are under the capping layer. Following this, the wafers are patterned for the field implants and field oxidation. Field oxidation has the next highest thermal budget in the fabrication sequence, at 1000 °C for a total of about 4.5 h. The wafers are then patterned for threshold implants and gate oxidation. The circuit process has two layers of polysilicon in addition to the three layers used for microstructures. The first

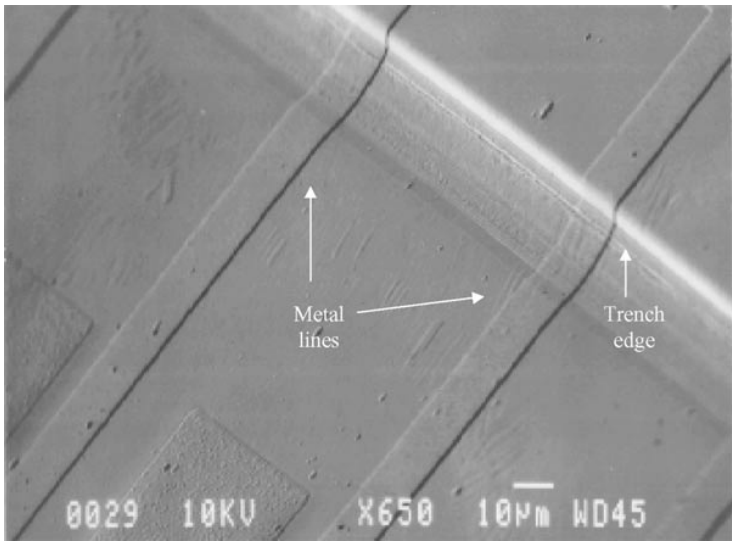


Figure 3. A scanning electron micrograph of the aluminum lead transfer over sidewalls of the recess.

of these is the gate polysilicon, which is 6000 Å thick and is phosphorus-doped to a sheet resistance of 10–15 Ω sq^{-1}. A thermal oxide is grown on it before depositing another polysilicon layer to form capacitors for use in analog circuits. This polysilicon is also used for resistors, and is doped to a level of 25–30 Ω sq^{-1}. After patterning it a layer of LPCVD oxide is deposited.

At this point the wafers are ready for contact cut and metallization. The two-step patterning of the structural polysilicon basically eliminates its use in lead transfer to the circuitry because it is stripped from the sidewalls of the recess in the first of these steps. The connections to the circuit, therefore, travel by metal to the thin polysilicon at the bottom of the recess, from which they pass to the structural members through buried contacts. If CMP planarization is used, as in [4], the metal may run out above the recess region and then drop down at some location to contact the microstructure. The step height at this contact could be reduced by having the metal directly contact the second or third structural polysilicon. In our process, however, the passivation above the recess is conformal to the sidewall of the trench and to the microstructures beneath it. The metal interconnect must, therefore, follow the sidewall of the trench to its bottom, and the lead transfer is best done to the lowest polysilicon. Even so, the contact within the recess must cut through 4 μm of sacrificial oxide, which is substantially more than the contact cuts in the circuit region. In order to prevent excessive undercutting of the contacts in the circuit region they are patterned separately, using a different mask. The interconnect metal (Al with 0.5% Si) is then sputtered to ensure proper coverage of the recess sidewall, which has a convex angle of about 125° to the trench bottom. The wafers are sintered in forming gas, and then passivated.

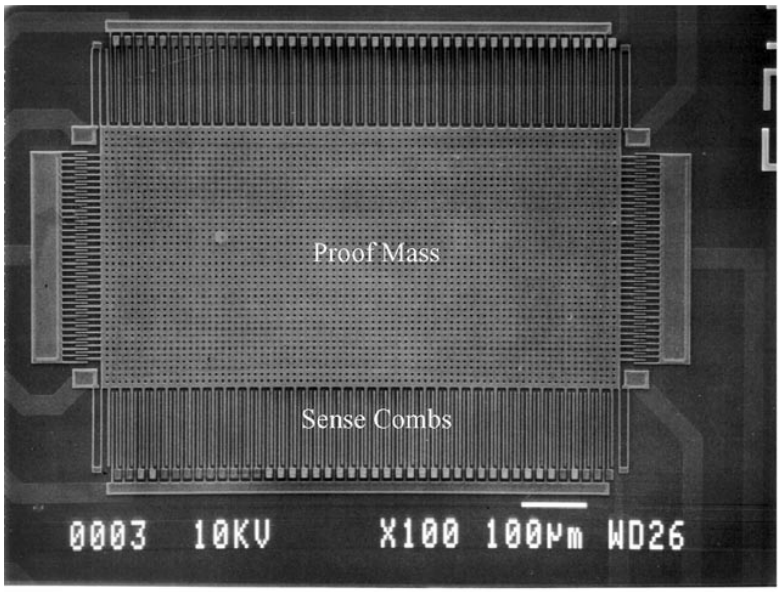
The final step in the fabrication process is to release the microstructures by wet-etching the sacrificial oxide in HF or BHF. The duration of the etch required may be prolonged because of potential densification of the sacrificial oxide during CMOS process steps such as p-well drive-in and field oxidation. It is necessary to carefully protect the circuitry and all the interconnect during the sacrificial etch since Al is rapidly attacked by the etchant. Depending on the duration of the etch, photoresist alone might not provide adequate protection since it peels off after about 30 min in BHF. In such cases a layer of PECVD nitride can provide additional protection for the metal.

3. Material characterization

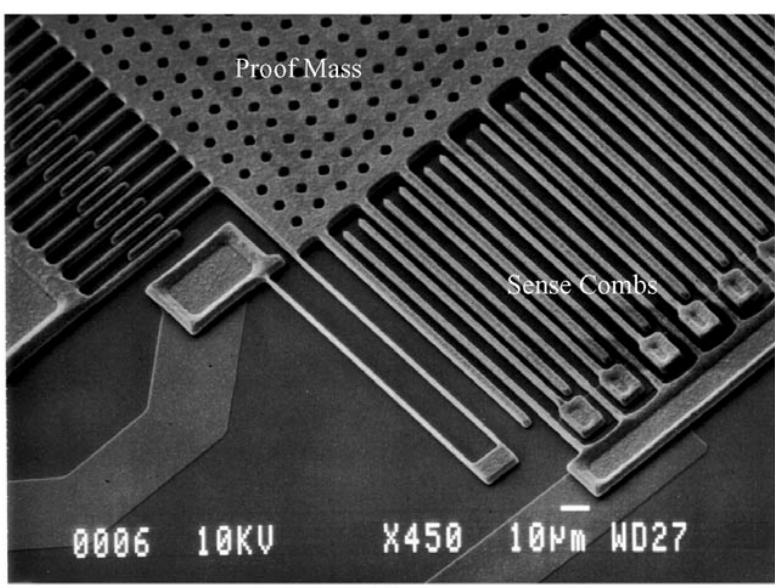
The strategic choice to fabricate the microstructures before the circuitry has the consequence of subjecting the structures to a much larger thermal budget than is typically used in annealing the structural polysilicon. Whereas typical anneals are less than 3 h and less than 1100 °C, the p-well drive-in can be as long as 16 h at 1200 °C. Various samples of polysilicon were subjected to the latter anneal to determine its consequences. Based on these experiments, the following observations can be made.

(1) After the long anneal, undoped polysilicon is in mild tension (less than +20 MPa) regardless of the original deposition temperature and the as-deposited residual stress. Furthermore, the long anneal transforms the dominant grain orientation to (111), and the grain size grows to 2000–4000 Å. Since LPCVD polysilicon deposited at 625 °C has a deposition rate that is three times higher than polysilicon deposited at 570 °C, these results make a case for using the former as the microstructural material in the interest of increased throughput. The only caveat is that the surface roughness of the former is substantially larger, and it can be difficult to pattern narrow linewidths in it.

(2) Doping the polysilicon with phosphorus before the long anneal introduces compressive stress which is linearly related to the doping temperature. Specifically, a 1 h deposition and 1 h soak at 950, 1000 and 1100 °C, causes the final stress to be −5, −35 and −98 MPa, respectively. This also indicates that the compressive stress is inversely related to sheet resistance.



(a)



(b)

Figure 4. (a) A scanning electron micrograph of a capacitive microaccelerometer; (b) a close-up of one corner, showing the sense electrodes.

Overall, the results confirm that a long anneal at 1200 °C for 16 h (in N_2), can be accommodated by the polysilicon microstructures. Additional details on this topic are described in [7].

4. Experimental results

Figure 2 shows a microaccelerometer co-fabricated with its CMOS interface circuitry. A closer view of step coverage over the side of the recess is provided in figure 3.

Figure 4 shows a capacitive microaccelerometer without interface circuitry. Its proof mass, suspension beams, and sense electrode combs are defined in the second structural polysilicon, while electrical connection to this second polysilicon is achieved using the first polysilicon as shown in figure 4(b). The vibratory gyroscope included on the test chip similarly uses the second structural polysilicon for the proof mass, suspension and drive combs. Sense electrodes are formed below and above the proof mass by the first and third polysilicons, respectively [10].

Typical performance parameters of fabricated NMOS and PMOS transistors were measured, and are presented in table 1. The measured output curves of NMOS and PMOS transistors are shown in figure 5. Since the CMOS process followed the MEMS process, no performance changes were expected, or observed, due to the inclusion of MEMS. Preliminary circuit tests have been performed on individual transistors as well as larger circuit blocks such as operational amplifiers. Other on-chip circuits included CMOS switched capacitor charge amplifiers for interfacing with accelerometers and PMOS source followers for interfacing with gyroscopes. Device and circuit details

Table 1. Typical parameters measured on transistors integrated with polysilicon microstructures.

Parameter	VT-NMOS	KP-NMOS	LAMBDA-N	VT-PMOS	KP- PMOS	LAMBDA-P
Value	0.76	28	0.013	−0.96	11	0.007
Unit	V	μA V^{-2}	V^{-1}	V	μA V^{-2}	V^{-1}

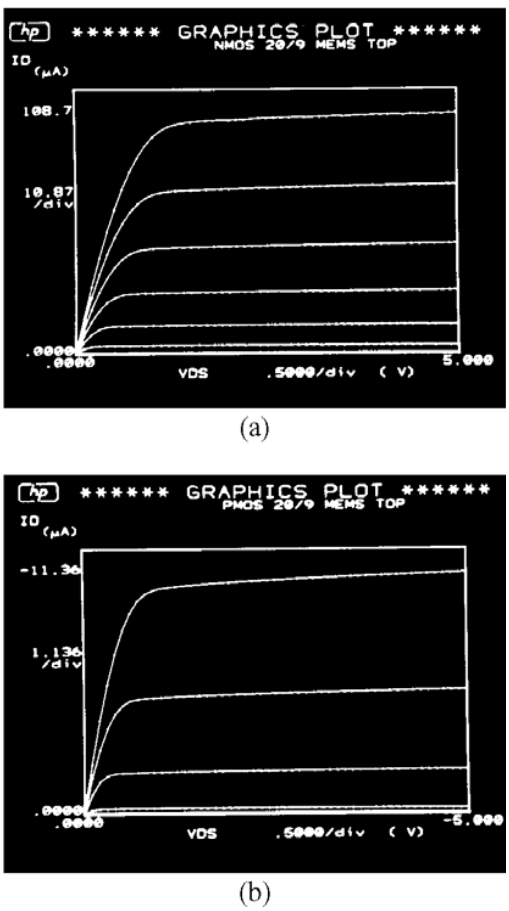


Figure 5. Output curves of (a) NMOS and (b) PMOS transistors co-fabricated with polysilicon microstructures in the process described.

are beyond the scope of this paper. The essential result is that functional CMOS devices with expected performance characteristics, and circuits based on these devices, have been successfully fabricated.

5. Conclusions

A fabrication process has been developed to integrate surface micromachined structures with CMOS circuitry. It eliminates topographical and thermal budget constraints placed upon the fabrication of the microstructures by fabricating them within recesses on the substrate wafer before initiating the circuit process. The process sequence that was developed integrates three level polysilicon microstructures with a one-metal, two-polysilicon, p-well based LOCOS CMOS process. It requires a total of 18 masks and 22 lithography steps. Four of these are non-critical alignments performed with the same mask at different points in the sequence. They may be eliminated if the depth of focus of the alignment equipment is adequate for the lithography necessary at the bottom of the recesses. Other special features of the process address the protection of microstructures during circuit fabrication, and lead transfer between the microstructures and the circuitry. This process does not require the deposition of a thick layer of silicon oxide to fill the trench, and the subsequent CMP step to planarize the wafer. Obviously, the wafer surface around the trench periphery is not planar. However, this does not cause major problems in either lead transfer, or in spinning and patterning photoresist. A metal interconnect covers the step at the periphery of the trench without any difficulty because the trench sidewall has a convex angle to the trench bottom.

The test vehicle for the development effort included micromachined accelerometers and gyroscopes. The primary structural members, i.e. the proof masses and suspension beams were 3 μm thick in these devices. Three structural layers and two sacrificial layers were used, with a total thickness of 9 μm. The number of layers and their thickness can be scaled up or down easily.

Acknowledgments

This project was funded by the Samsung Advanced Institute of Technology (SAIT), Suwon, Korea. The process development was performed at the University of Michigan's Solid State Electronics Laboratory. The authors gratefully acknowledge Mr K Handique, Dr C C Huang, Mr A Salain, Ms C Tom, Dr N Yazdi and Dr B Ziaie of the University of Michigan, and Dr Y Oh of SAIT for contributions to this effort.

References

[1] Howe R T and Muller R S 1986 Resonant-microbridge vapor sensor *IEEE Trans. Electron Devices* **33** 499–506

[2] Yun W, Howe R T and Gray P R 1992 Surface micromachined, digitally force-balanced accelerometer with integrated CMOS detection circuitry *Tech. Digest IEEE Solid-State Sensor and Actuator Workshop (Hilton Head, June 1992)* pp 126–41

[3] Kuehnel W and Sherman S 1994 A surface micromachined silicon accelerometer with on-chip detection circuitry *Sensors Actuators* A **45** 7–16

[4] Smith J H, Montague S, Sniegowski J J, Murray J R and McWhorter P J 1995 Embedded micromechanical devices for the monolithic integration of MEMS with CMOS *Proc. Int. Electron Devices Meeting, IEDM '95 (December 1995)* pp 609–12

[5] Kung J T 1996 Methods for planarization and encapsulation of micromechanical devices in semiconductor processes *US Patent Specification* 5 504 026

[6] Offenberg M, Larmer F, Elsner B, Munzel H and Riethmuller W 1995 Novel process for a monolithical integrated accelerometer *Proc. Int. Conf. on Solid-State Sensors and Actuators, Transducers '95 (June 1995)* pp 589–92

[7] Gianchandani Y B, Shinn M and Najafi K 1998 Impact of high thermal budget anneals on polysilicon as a

micromechanical material *IEEE J. Microelectromech. Syst.* **7** 102–5

[8] Gianchandani Y B, Kim H, Shinn M, Ha B, Lee B, Najafi K and Song C 1998 A MEMS-first fabrication process for integrating CMOS circuits with polysilicon microstructures *Proc. IEEE Workshop on Micro Electro Mechanical Systems, MEMS '98 (Heidelberg, Germany, January 1998)* pp 257–62

[9] Ji J and Wise K D 1992 An implantable CMOS circuit interface for multiplexed microelectrode recording arrays *IEEE J. Solid-State Circuits* **27** 433–43

[10] Oh Y, Lee B, Baek B, Kim H, Kim J, Kang S and Song C 1997 A surface-micromachined tunable vibratory gyroscope *Proc. IEEE Workshop on Micro Electro Mechanical Systems, MEMS '97 (Nagoya, Japan, January 1997)* pp 272–7