A single-sided process for differentially cooled electrothermal micro-actuators

H Veladi\(^1\), R R A Sym\(^2\) and H Zou\(^3\)

1 Electrical Engineering and Computer Department, Tabriz University, Bolvar 29 Bahman, Tabriz 51666-14776, Iran
2 Electrical and Electronic Engineering Department, Imperial College, Exhibition Road, London SW7 2AZ, UK
3 School of Mechanical Engineering, Dalian University of Technology, 2 Linggong Road, Dalian 116024, People’s Republic of China

E-mail: r.syms@imperial.ac.uk

Received 10 February 2008, in final form 28 March 2008
Published 21 April 2008
Online at stacks.iop.org/JMM/18/055033

Abstract

A simple method for increasing the thermal efficiency of shape bimorph electrothermal micro-actuators is proposed, based on a reduction of gas conduction cooling beneath the hot arms by a deep, localized undercut. A single-sided, single-mask SCREAM-type process for fabricating differentially cooled actuators in a bonded silicon-on-insulator material is demonstrated. The process uses deep reactive ion etching and undercut to form suspended parts and isotropic reactive ion etching and lift-off of sacrificial shields to form localized mesas. The advantage of the method is confirmed using folded electrothermal actuators, and an approximate halving of the drive power is demonstrated by variations in the substrate profile in the vicinity of a series of actuators with the same mechanical design.

1. Introduction

Electrothermal actuators are widely used to power motion in microsystems because of their simplicity, small size, long travel and high force. Originally demonstrated as material bimorphs [1, 2], electrothermal actuators are now generally fabricated from a single material using standard processes such as surface micromachining of single crystal silicon [3], polysilicon [4] and bonded silicon-on-insulator (BSOI) [5].

Their main disadvantage is continuous and relatively high power consumption, especially in silicon, which has a poor value of the ratio \(\alpha/k\), where \(\alpha\) is the thermal expansion coefficient and \(k\) is the thermal conductivity.

The main variants are structures in which a suspended elastic element is thermally strained against the substrate and shape bimorphs in which one suspended element is strained against another. The former type (figure 1(a)) consists of a built-in-beam, which is heated by passing a current between its two anchors and deflects by buckling. Smooth deflection in a preferential direction can be forced by the use of a chevron layout [6–9]. The latter is more compact, but has many variants. Figure 1(b) shows the original design introduced by Guckel, which has been investigated thoroughly [10–16]. It consists of a U-shaped suspended beam, which is again heated by passing a current. However, differential heating is achieved by a variation between the cross-sections of the two arms, and causes deflection by a combination of bending and buckling. Differential heating may also be achieved by using different arm lengths [17, 18] (figure 1(c)) or by passing a current only through a split hot arm [19] (figure 1(d)).

Electrothermal actuators have been used in a variety of linear and rotary stages and engines [20–23], and in optical fibre aligners [24, 25], switches [26] and variable attenuators [27–29]. Other applications include electrical circuit breakers [30].

Clearly, any improvement in the efficiency of electrothermal actuators is beneficial. To date, most authors have concentrated either on localization of heating or on optimization of the elastic deflection by the variation of geometric parameters in the flexible elements, rather than on control of the distribution of cooling. In this paper, we show how the thermally induced deflection may be increased further, by localizing gas conduction cooling to the substrate to the cold arm. We introduce a single-sided, single-mask fabrication process for forming terraced structures based on the modification of the single crystal reactive etching and metallization (SCREAM) process [3]. The new process uses a BSOI material rather than bare silicon and allows
differentially cooled structures to be realized by combining simple modifications to the layout with an additional undercut step. We demonstrate an approximate halving of the drive power by the stepwise application of the method to the same folded actuator layout. This method is extremely flexible and may easily be applied to other materials, processes and actuator designs. The design approach is introduced in section 2. Device fabrication is described in section 3 and experimental results in section 4. Conclusions are presented in section 5.

2. Theory

The thermal performance of electrothermal micro-actuators may generally be investigated independent of their elastic design using a one-dimensional model corresponding to an unfolded version of the layout [14, 16]. For example, figure 2(a) shows a shape bimorph similar to figure 1(d), which consists of three suspended cantilevers linked by a rigid beam at their free ends. This device was successfully demonstrated in [31], but with the recessed areas entirely removed by an expensive double-side process involving deep etching from the rear of the wafer. The outer arms are electrically heated, while the unheated central arm is a tether. The hot arms are inclined at a shallow angle $\theta_H$ and the cold arm at a different angle $\theta_C$. The width $w_H$ of the hot arms is smaller than the widths $w_C$ of the cold arm to promote heating. However, the cold arm is often subdivided into sections of different widths, to increase flexibility near its ends. With beams of a suitable aspect ratio, preferential expansion of the hot arms forces an in-plane deflection in the direction shown.

Ignoring the link beam, the arrangement may be unfolded as in figure 2(b). The two hot arms in parallel form a first section running from $x = 0$ to $x = L$. Here they connect to the cold arm, which forms a second section from $x = L$ to $x = 2L$. The main thermal effects are electrical heating, solid conduction cooling via the anchors, gas conduction cooling via the substrate and convection cooling. Radiation can typically be neglected, since this involves elevated temperatures that result in extremely short lifetimes. In the $i$th section, the steady-state temperature variation may be found from the heat conduction equation [14, 16]:

$$k_S w_i t_i d^2 \Delta T_i / dx^2 - (k_{\text{Air}}(w_i/g_i) + h S_i) \Delta T_i + P_i \rho_S / w_i t_i = 0. \quad (1)$$

Here $\Delta T_i$ is the temperature above ambient, $x$ is the position along the beam, $I_i$ is the heating current, $w_i$ is the width, $t_i$ is the thickness of the beam and $k_S$ and $\rho_S$ are the thermal conductivity and resistivity of the beam material, respectively. $S_i$ is the surface area of the beam per unit length and $h$ is the convection heat transfer coefficient. Finally, $g_i$ is the thickness of the air layer beneath and $k_{\text{Air}}$ is its thermal conductivity. The term $I_i^2 \rho_S / w_i t_i$ can also be written as $P_i / L_i$, where $P_i$ and $L_i$ are the electrical power and length of the section, respectively. Convection can be seen to have a similar effect to gas conduction cooling. However, since the value of $h$ is difficult to estimate, its significance will be assumed to be small for simplicity.

The values $w_i$, $t_i$ and $g_i$ compared with the length $L$ determine the relative significance of the various terms. At constant power, decreasing $t_i$ without altering other values decreases solid conduction cooling. Similarly, increasing $g_i$ decreases gas conduction cooling. Here, our interest is to see the effect of varying $g_i$ to mimic the effect of a localized recess that may reduce gas conduction cooling beneath the hot arms. Clearly, the prediction of equation (1) will be inaccurate if $g_i \gg w_i$ when a two- or even three-dimensional model is
The temperature difference between the hot and cold arms will generate thermal strains, which in turn convert to a lateral deflection in a well-designed actuator. It is not our purpose here to provide a detailed thermoelastic analysis or a prediction of the deflection. We simply point out that the main effects will be determined in any given actuator by the difference in average temperature,

$$\Delta T_{\text{avg}} = \left\{ \int_0^L \Delta T_1 \, dx - \int_{L}^{2L} \Delta T_2 \, dx \right\} / L,$$

between the hot and cold arms [16]. Figure 3(b) shows this variation as a function of the hot arm undercut depth $g_1$, for different values of the cold arm width $w_2$. Clearly, $\Delta T_{\text{avg}}$ rises as $g_1$ increases, but saturates when $g_1$ is large, so that the majority of the increase is already obtained when $g_1 \approx 50 \mu m$. A slightly decreased effect is obtained as the cold arm width $w_2$ increases, when gas conduction cooling in the hot arms becomes relatively less significant.

These results are only qualitative because of the simple thermal model, especially the neglect of convection and the three-dimensional nature of gas conduction cooling when the undercut depth is large. Despite these caveats, they suggest that an advantage should be obtained if the undercut beneath the hot arms can be increased, and that this advantage is likely to be an increase in the deflection per unit power of around 2–3 times.

3. Device fabrication

All that is required to increase the undercut beneath the hot arms is a bi-level etching scheme. We have developed a suitable process based on a BSOI material. The process starts by using anisotropic reactive ion etching to define the suspended mechanical parts and a set of closely spaced sacrificial shields. Isotropic reactive ion etching is then used to define mesas outlined by the shields, and finally wet etching is used to undercut the mechanical parts and remove the etching. Differential structuring is therefore achieved using single-sided processing and single-step lithography. The process has few topographic limitations, and can be used with a wide range of structures. The only real requirement is for the sufficient surface area to allow inclusion of sacrificial shields that are large enough to prevent their undercut by the isotropic dry etch.

Figure 4(a) shows the process, which has many features in common with SCREAM [3], such as the use of re-oxidation to protect deeply undercut features. The process was applied to 100 mm diameter (1 0 0) oriented BSOI with a 20 $\mu m$ bonded silicon thickness and 2 $\mu m$ oxide thickness (Icecms Technology, Belfast), and yielded approximately 600 devices per wafer.

To form a hard mask, a 1.2 $\mu m$ thick layer of thermal oxide was first grown by dry thermal oxidation for 48 h at 1100 °C, and a 1000 Å thick layer of Cr metal was deposited by RF sputtering. The wafer was then spin-coated with a 4.5 $\mu m$ thick layer of Shipley S1828 photosist, which was soft-baked, patterned using a UV contact mask aligner and then hard-baked. The mask layout simultaneously defined anchor
points, suspended mechanical parts and sacrificial shields. The shields were small areas of silicon that outlined the regions to be protected from isotropic etching, before their own later removal by undercut etching. The shields were perforated to assist in this step.

The pattern was transferred through the Cr layer using wet etching in ceric ammonium nitrate, and then through the oxide layer using reactive ion etching in a parallel plate RIE (Plasmalab 80+, Oxford Instruments, Yatton, UK) using CHF₃, O₂ and Ar gases. The pattern was then transferred through the bonded silicon layer by deep reactive ion etching using an inductively coupled plasma etcher (Single Chamber Multiplex, Surface Technology Systems, UK) operating a cyclic etch/passivation process based on SF₆ and C₂F₆ gases. Finally, the pattern was then transferred through the buried oxide layer by a second parallel plate RIE step. The photoresist and Cr layers were then removed by wet processing, and the wafer was cleaned in an oxygen asher (step 1).

A 4000 Å thick oxide layer was then grown by dry thermal oxidation for 4 h at 1100 °C (step 2). This layer was thick enough to protect the mechanical parts from subsequent etch steps, but thin enough to minimize buckling of narrow beams after release. Oxide was then removed from exposed areas of the substrate by a third parallel plate RIE step. The overall effect was to leave the mechanical parts entirely surrounded by thick oxide on their horizontal surfaces and thin oxide on their vertical surfaces (step 3).

Exposed areas of the substrate were then isotropically etched to a depth of 50 μm, by using SF₆ alone in the STS DRIE. The extent of the lateral etching was sufficient to undercut narrow mechanical parts fully, but wide parts only partially (step 4). Although the results of figure 3 suggest that most of the thermal advantage will be obtained with a shallower etch, it was found that this did not lead to complete release of link beams. Clearly, the mesa dimensions were eroded significantly in the process. Exposed surface oxide was then removed by wet etching in buffered HF. This process not only freed suspended mechanical parts, but also allowed the shield areas to be removed by floatation (step 5). The released structure was then freeze-dried in a water/methanol mixture, and coated with 1000 Å Cr metal and 400 Å Au metal by RF sputtering to allow electrical contact. Isolation between different parts of the structure was ensured using a long HF etch, which removed sufficient interlayer oxide that sputtered metal could not track down to the substrate.

The use of a BSOI material rather than bare silicon in a SCREAM-type process clearly allows the hot arms to be deeply undercut, while suspending the cold arms at a much smaller distance above the substrate, which can be accurately defined by the oxide interlayer thickness. The use of sacrificial shields allows the region protected from a deep undercut to be selected by layout and to be extended on one or both sides of the cold arms to allow for lateral motion. This flexibility allowed the effect of additional undercut etching to be investigated using a set of actuators with exactly the same mechanical layout, but with variations in cooling determined only by their precise shield arrangement.

Figure 4(b) shows one set of variations investigated. Here the hot and cold arms were each provided with a sacrificial shield. For reasons of space, the shields were only located on one side of each beam, towards the direction of motion. The cold arms were provided with full-length shields, while the hot arms were provided with partial shields of length Lₕ and a corresponding undercut of length Lₜ. Decreasing Lₜ therefore increased the fraction of the hot arm Lₕ/L that was deeply undercut and consequently reduced the gas conduction cooling. Devices with Lₜ/L = 1 correspond to the optimum achievable with the new process, while devices with Lₜ/L = 0 should have a performance similar to that achievable using standard BSOI processing involving deep reactive ion etching and wet undercut alone, especially at large deflections.

Actuators were constructed with a length L = 2000 μm, hot arm parameters \( \theta_H = 2° \) and \( w_H = 8 \mu m \) and cold arm parameters \( \theta_C = 1° \) and \( w_C = 40 \mu m \) in a central region of length 1200 μm and 6 μm in end sections each of length 400 μm. The link beam length was 300 μm, and variants were prototyped with different end links formed as trusses and perforated beams respectively. No difficulties were found in fabrication, apart from occasional failures to remove shield elements due to an incomplete HF undercut.
Figure 5. SEM photographs of completed actuators with (a) $L_u/L = 1$ and a truss-type end link and (b) $L_u/L = 0.5625$ and a perforated beam end link.

Figure 5(a) shows a scanning electron microscope photograph of a completed device, which has fully undercut hot arms ($L_u/L = 1$) and a truss-type end link. Here a central mesa may clearly be seen running alongside and slightly beneath the uppermost edge of the cold beam. Figure 5(b) shows a similar device, with a hot-arm mesa along part of its length and partial undercut ($L_u/L = 0.5625$) and a perforated beam end link.

Figure 6(a) shows a magnified view near the free end of a truss-ended device, which shows the hot arms and the link beam to be completely undercut and the cold arm to be fully undercut in its narrow, flexible sections and partially undercut in its wide central section. Occasional stalagmitic feature characteristics of an isotropic undercut may be seen; however these are unimportant in what is essentially an average cooling effect, especially since (as has previously been noted) an over-etch was used for undercut. Variants with perforated beam end links yielded larger spurious features. Figure 6(b) shows an alternative view of the free end, which shows the local cooling mesa for the cold arm in greater detail.

4. Experimental results

Device performance was evaluated by measuring in-plane deflection as a function of power, with the device mounted on a heat sink. The electrical power was calculated from the separate measurement of the voltage across and the current through the hot beam anchors. Device resistance was found from the ratio of the voltage and current as $\approx 100 \Omega$, although a slow increase with heating power was observed as discussed below. In-plane displacement was found using an optical microscope equipped with a calibrated on-screen cursor measurement system, and deflections of $\approx 50 \mu m$ were obtained using $\approx 100 \text{ mW}$ power. Damage to the Au coating occurred at $\approx 120 \text{ mW}$ in devices with fully undercut hot arms and at steadily increasing power up to $\approx 200 \text{ mW}$ as $L_u/L$ decreased. AC displacement was measured in the same way. A first-order thermal time constant with a 3 dB point at a drive frequency of $\approx 40 \text{ Hz}$ was observed, together with a mechanical resonance at $\approx 2.8 \text{ kHz}$.

Figure 7(a) shows the detailed variation of deflection with the drive power, for a set of devices with different fractional
lengths \( L_u/L \) of a hot arm undercut. In each case, deflection is a quasi-linear function of power. The slopes of the data are all similar for low deflections, but greater differences are obtained at larger deflections when the hot arms are partially suspended over their cooling mesas. As \( L_u/L \) increases, and gas conduction cooling of the hot arms is reduced, there is a clear reduction in the power needed to achieve a given deflection. The effect was verified using the measurement of multiple devices of similar type, and almost identical variations were obtained from the devices with perforated beam end links. Similar behaviour was obtained for devices with different cold arm lengths. Figure 7(b) shows the variation with \( L_u/L \) of the power \( P_{50} \) needed to obtain 50 \( \mu \text{m} \) deflection, showing that the drive power is approximately halved by increasing \( L_u/L \) from zero to unity, in qualitative agreement with the arguments of section 2.

The thermal nature of the effect was confirmed by measuring variations in device resistance. Figure 8 shows the variation of the resistance with the drive power, for devices with approximately zero length and full-length hot-arm cooling mesas. Here the resistance has been normalized against its low power asymptote, to eliminate variations caused by small differences in metal thickness. There is some scatter in the data but the effects are clear. In each case, the normalized resistance rises with the drive power, suggesting that this quantity can provide a form of temperature measurement. However, the rise is considerably more rapid in the device with \( L_u/L = 1 \), i.e. with lower gas conduction cooling in the hot arms.

Even greater differences in performance would clearly have been obtained between devices with cooled and uncooled hot arms if shields had been extended on either side of the cold arm. These features would allow isotropic etch damage to the cold arm mesa to be eliminated almost entirely, especially near the hottest (and therefore most important) flexible section close to the link beam. Further advantage could also have been obtained by retaining a mesa below the link beam itself, since this is wide enough to ensure a rapid reduction in temperature before the heat flow can reach the cold arm. However, even this simple demonstration shows the basic approach to be effective, and the use of identical elastic test structures confirms the thermal basis of the performance improvement.

In any electrothermal device, there is a trade-off between the drive power and speed of response. Experimentally, it was found that devices with reduced gas conduction cooling were slightly slower. However, it was difficult to quantify the effect because the cooling was deflection dependent. Slow actuators find uses in preventing transients from exciting resonances in under-damped mechanical oscillators. In electrothermal circuit breakers [30], for example, it is advantageous to reduce the sensitivity to transient overload. In systems with complex mechanical dynamics, such as the multi-blade iris variable optical attenuator in [31], it is advantageous to prevent asynchronous oscillations that might cause mechanical collision.

5. Conclusions

A new method has been developed for enhancing the performance of electrothermal microactuators based on a reduction of gas conduction cooling beneath the hot arms by a deep, local undercut. A single-sided, single-mask SCREAM-type process for fabricating differentially cooled actuators in a BSOI material has been demonstrated. The process uses deep reactive ion etching and undercut to form suspended parts and isotropic reactive ion etching and lift-off of sacrificial shields to form local mesas. The advantage of the method has been confirmed using folded electrothermal actuators, and an approximate halving of the drive power has been demonstrated.
The method is flexible, since it allows variations in gas conduction cooling to be achieved simply by the layout of additional sacrificial features. Consequently, it could be applied to a wide range of actuator designs. Provided damage to relatively thin mechanical parts due to oxidation-induced buckling can be avoided, it should be directly applicable to polysilicon micromachining. Analogous processes should be applicable to other material systems that also allow isotropic removal of a substrate material by differential etching. One obvious example is electrothermal actuators based on nickel mechanical parts on oxide insulators on silicon substrates, obvious example is electrothermal actuators based on nickel mechanical parts on oxide insulators on silicon substrates, which would take advantage of the increased value of $\alpha/k$ available in Ni. Although additional process steps (protective oxidation and isotropic undercut) are required for the BSOI devices demonstrated, oxidation may not be required in all material systems, and an isotropic dry undercut may already be used for release (see e.g. [30]). Other dry undercut methods such as XeF$_2$ etching could also be used, together with protective materials other than SiO$_2$. Finally, it is worth noting that a similar approach could be adopted for local control of damping in vibrating microstructures.

References


