A silicon MEMS structure for characterization of femto-farad-level capacitive sensors with lock-in architecture

J Wei\textsuperscript{1}, C Yue\textsuperscript{1,2}, Z L Chen\textsuperscript{2}, Z W Liu\textsuperscript{2} and P M Sarro\textsuperscript{1}

\textsuperscript{1} Dimes, Delft University of Technology, Delft, The Netherlands
\textsuperscript{2} IMETU, Tsinghua University, Beijing, People’s Republic of China

E-mail: j.wei@tudelft.nl

Received 15 December 2009, in final form 12 April 2010
Published 1 June 2010
Online at stacks.iop.org/JMM/20/064019

Abstract
This paper presents a silicon MEMS capacitive structure to investigate a test methodology for fF-level capacitive sensors’ measurement. The device mimics a capacitive sensor with a changing intermediate layer between the electrodes. A single mask bulk micromachining process is used to fabricate the device, which has a nominal capacitance of 1.2 fF. A high performance measurement setup based on lock-in principle is developed to detect the capacitance variation. The maximum capacitance variation of the fabricated device is 0.31 fF, and the capacitance detection limit is 0.095 aF Hz\(^{-1/2}\).

(Some figures in this article are in colour only in the electronic version)

1. Introduction

MEMS capacitive sensors offering high resolution, low intrinsic temperature coefficient, large dynamic range and low power consumption have been widely used in many applications, such as strain gauges \cite{1}, inertial sensors \cite{2} and micro-fluidic sensors \cite{3}.

There are three basic configurations for converting the mechanical motions into capacitance variation, as indicated in figure 1. The capacitance variation can be generated by changing the distance between two parallel plates (figure 1(a)). This configuration has large sensitivities but nonlinear characteristics. The second configuration (figure 1(b)) changes the overlapping area of the plates, which gives linear measurement with reduced sensitivity. The last configuration (figure 1(c)) changes the intermediate layer between the electrodes. This configuration is especially interesting for microfluidic devices \cite{3} and biomedical detections \cite{4}: the electrodes are placed on the sidewall or at the bottom of the fluidic channel, so that the impedance or capacitance changes caused by the change of liquid level and the appearance of particles in the channel are detected. However, as the intermediate layer can have different electrical properties, signal detection in this configuration is more complex.

The miniaturization of sensing systems requires the sensor dimensions to be shrunk down. Consequently, the nominal capacitance of the sensor to be measured could be as small as fF-level \cite{5}. This results in a high-impedance signal path and limits the signal to nA-levels. Furthermore, the tiny signal is often degraded by parasitic components and disturbed by electrical noise. These factors limit the resolution of the sensor, posing great challenges for both MEMS design and detection strategy.

To address the above-mentioned issues, this paper presents a specifically designed MEMS test structure which allows the investigation of measurement methods for detecting fF-level capacitive sensors with aF-level variation. It particularly focuses on detecting the capacitance variation generated by the change of a conductive intermediate layer. The concept of the device is given in section 2, and device fabrication and the measurement system are described in sections 3 and 4, respectively. Finally, the results and discussion are presented in section 5 of the paper.
2. Concept

To study the behavior and the measurement principle of fF-level capacitive sensors, the proposed device is designed to mimic a real capacitive sensor with a changing conductive intermediate layer. There are several design requirements. Both the nominal capacitance and the variation of the device capacitance need to be at fF-level. The variation of the capacitance needs to be externally controllable. The parasitic components in the device should be well limited to allow the investigation of proper measurement methods. Since it is a test structure, the fabrication of the device should be simple.

To meet the above-mentioned requirements, the test structure is specifically designed as schematically shown in figure 2. The structure contains two parts: the ‘electrodes’ and the ‘blade’. The ‘electrodes’ part has two pairs of electrodes placed parallel to each other as the sensing electrodes and reference electrodes, respectively. The nominal capacitance between the electrodes can be designed by tuning the geometric dimensions. The ‘blade’ part has a suspended blade supported by flexible beams. It has two basic states, ‘stand-by’ and ‘inserted’, as indicated in figure 2. The blade can be inserted between the two fixed sensing electrodes by external forces, and the position of the blade can be monitored with a microscope. The blade is designed to be conductive to mimic a conductive intermediate layer in a real capacitive sensor, and the motion of the blade generates the desired capacitance variation.

Since the signal of interest is as small as nA-level, the influence of the parasitic components should be reduced as much as possible. To do this, the coupling of the electrodes with the substrate needs to be suppressed. This can be achieved by removing the material surrounding the electrodes to separate them from the substrate. Guard rings for ground shielding are placed around the electrodes to suppress unwanted disturbances from the environment. The sensing and reference electrode pairs have the same geometric shape and size, so that a pair of anti-phase input signals can be applied to cancel unwanted influence during measurement.

The main geometric parameters of the device are indicated in figure 3.

3. Device fabrication

To fabricate the proposed device, a SCREAM (single-crystal reactive etching and metallization) process [6] is chosen. The process contains three general steps: (1) formation of a suspended silicon structure; (2) deposition of a dielectric layer on the silicon structure and (3) sputtering of a metal layer for electrical connections.

To reduce the influence of the parasitic components, the electrodes are suspended from the substrate during the first step of the process. The dielectric layer on the silicon structure is created with thermal oxidation. This provides uniform coverage on the 3D surface of the electrode and a thick layer up to 2–3 μm can be easily achieved. At the same time, the mechanical support, such as the silicon anchors under the electrodes, is partly turned into oxide during oxidation. This further reduces the coupling between the electrodes and the substrate.

Figure 4 shows the complete process flow, with a view across the cross-section A–A’ in figure 3. The process uses only one lithography step at the beginning of the fabrication to define the horizontal dimensions of the electrodes and beams. Oxide from plasma-enhanced chemical vapour deposition (PECVD) is used as a hard mask for the deep reactive-ion etching (DRIE) process (figure 4(a)). A combination of DRIE anisotropic etching and isotropic silicon etching with SF6 is used to define the vertical size of the structure and release movable structures from the substrate (figure 4(b)). The suspended connection (in figure 4(b)) to each electrode is released by isotropic etching from both sides of the connection. To isolate the structures, wet oxidation of silicon is performed to create a 2.5 μm uniform insulation layer (figure 4(c)). Finally, the device is coated with a 3 μm thick aluminum layer for reliable electrical connections (figure 4(d)).
The fabricated test structure is shown in figures 5 and 6. The electrodes and blade are 35 µm wide, 50 µm high and suspended about 40 µm above the substrate. The distance between the two electrodes is 40 µm. When the suspended blade is inserted, there is a 3 µm gap between the blade and the electrodes on each side. Other geometric parameters can be found in figure 3. The suspended beams supporting the blade are flexible enough to allow sufficient stroke for complete blade insertion. The symmetrical design of the supporting beams restricts the motion of the blade in the direction parallel to the electrodes’ surface, preventing contact between the blade and the electrodes during insertion. A ruler is integrated on the beam of the blade, allowing inspection of the blade position.

4. Measurement strategy

4.1. Configuration of the device

The fabricated device is first measured with an impedance analyzer. The nominal capacitance between the sensing electrodes is 1.2 fF at 100 kHz in the ‘stand-by’ state. For measuring such a small capacitance during the insertion of the suspended blade, the behavior of the blade needs to be understood. Since the suspended blade is conductive, the input signal will partly leak into the ground through the blade. This effect is illustrated in figure 7.

The coupling between the two electrodes without the insertion of the blade is represented as two capacitors as shown in figure 7(a). $C_1$ represents the coupling of the parallel-plate capacitor between the two electrodes, which is sensitive to blade insertion. $C_P$ represents the remaining coupling of the
the magnitude of the coupling impedance and thus can be treated as a constant value. Depending on in the figure, does not depend on the insertion of the blade, and therefore, can be treated as a constant offset in this simple model. When the blade is partly inserted between the two electrodes up to a given position $x$ (figure 7(b)), the capacitance of $C_1$ decreases due to the reduction of the effective overlapping area. The appearing capacitor $C_2$ represents the coupling between the electrodes and the side surface of the blade, and its capacitance increases with $x$. The two $C_2$ capacitors, connected in series, form a second signal path between the two electrodes. This path has lower impedance than the $C_1$ path, since the distance between the electrodes is much reduced. When the blade is fully inserted, as depicted in figure 7(c), the coupling through the $C_1$ path is replaced completely by the $C_2$ path.

The behavior of the $C_2$ path is significantly determined by the coupling between the conductive blade and the ground potential. This coupling, represented by the impedance $Z_{\text{GND}}$ in the figure, does not depend on the insertion of the blade, and thus can be treated as a constant value. Depending on the magnitude of the coupling impedance $Z_{\text{GND}}$, the sensor behaves in three different ways:

1. If $Z_{\text{GND}}$ is very small, most of the input signal through the $C_2$ path leaks into the ground, and there is little signal arriving at the output electrodes through the $C_2$ path.
   In this case, the equivalent capacitance of the device is dominated by $C_1$. Since $C_1$ decreases with blade insertion, the equivalent capacitance between the input and output electrodes decreases as well.

2. On the other hand, if $Z_{\text{GND}}$ is very large, there is limited coupling between the blade and the ground. Most of the signal energy entering the $C_2$ path arrives at the output electrodes. Since the $C_2$ path has lower impedance, the total signal arriving at the output electrodes through the two paths increases with the insertion of the blade. Consequently, the equivalent capacitance increases.

3. For a $Z_{\text{GND}}$ in between the above-mentioned two conditions, the situation becomes complex. At the beginning of the insertion, due to the leakage of the $C_2$ path, the equivalent capacitance drops. With further insertion, the energy entering the $C_2$ path increases dramatically. Despite the leakage, there are signals arriving at the output electrodes. Thus, an increase in the equivalent capacitance can be observed during further insertion. In this case, the device is not suitable for sensor applications, since the relationship between the equivalent capacitance and the blade position is non-monotonic.

Considering a constant and pure capacitive coupling between the blade and the ground ($C_{\text{GND}}$), the equivalent capacitance ($C_{\text{eq}}$) of the sensing electrode pair, without the $C_P$ offset, is derived as follows:

$$C_{\text{eq}}(x) = C_1(x) + C_2(x) \frac{1}{2 + \frac{C_{\text{GND}}}{C_2(x)}}.$$  \hspace{1cm} (1)

$C_1$ and $C_2$ are considered as ideal parallel-plate capacitors:

$$C_1(x) = \varepsilon_0 h (w - x)/d_1$$  \hspace{1cm} (2)

$$C_2(x) = \varepsilon_0 h x / d_2,$$  \hspace{1cm} (3)

where $\varepsilon_0$ is the permittivity of free space, $h$ and $w$ are the height and the width of the electrodes, respectively, $d_1$ is the distance between the two electrodes, $d_2$ is the gap between the electrodes and inserted blade and $x$ is the length of the blade inserted between the electrodes (figure 7(b)).

In principle, a $C_{\text{GND}}$ value satisfying either condition 1 or 2 in the above analysis can be chosen to achieve successful detection. However, in the real case, it is difficult to meet the requirement for condition 2, and only condition 1 can be chosen. As can be seen in equation (1), the behavior of the $C_2$ path is determined by the $C_{\text{GND}}/C_2$ ratio. In order to achieve low signal leakage as mentioned in condition 2, $C_{\text{GND}}$ needs to be smaller than $C_2$. Since $C_2$ is in the range of several
Additionally, the resistive part of the silicon substrate, and therefore it is much larger than fF-level. To further illustrate this, the equivalent capacitance of the device as a function of the blade position \( x \) is calculated with the geometric parameters shown in figure 3 and is reported in figure 8.

In the case of large coupling (condition 1), \( C_{\text{GND}} = 0.5 \) pF, the equivalent capacitance decreases linearly with the blade position \( x \). In the case of \( C_{\text{GND}} = 0.05 \) pF (condition 3), the device is in the transition condition, the equivalent capacitance has a non-monotonic relation with the blade position and, thus, the device is not suitable for sensor applications. When \( C_{\text{GND}} \) is 0.005 pF (condition 2), or an even smaller value, the equivalent capacitance has a nonlinear increase with the blade position. For a fabricated device, \( C_{\text{GND}} \) is mainly caused by the coupling between the metal layer on the blade (together with the flexible support) and the silicon substrate, and therefore it is much larger than fF-level. Additionally, the resistive part of \( Z_{\text{GND}} \) will further increase this coupling. Therefore, for an fF-level capacitive sensor, it is very difficult to achieve small enough coupling between the intermediate material and the ground potential to satisfy the requirement of condition 2. Condition 1 is thus chosen in the experiment.

Based on the above analysis, to achieve a proper measurement of the device, a small \( Z_{\text{GND}} \) needs to be chosen. A simple way to do this is to connect the blade directly to the ground potential. In this case, the \( C_{\text{GND}}/C_1 \) ratio is maximized, and the equivalent capacitance is consequently dominated by \( C_1 \), which has a linear dependence on the blade position \( x \). At the same time, since the accuracy of \( C_2 \) has less influence on the measurement, the requirement for precisely controlling the gap between the fixed electrodes and the side surface of the suspended blade becomes less important.

The above analysis only applies to capacitive sensors whose coupling between the intermediate layer and the ground potential is a constant value. In the case of variable coupling, other methods can be used, such as the capacitive sensor presented in [7].

### 4.2. Principle of the measurement circuit

To measure an fF-level capacitive sensor, a dedicated readout strategy needs to be developed. This strategy can be generally applied to all kinds of capacitive sensors. A fabricated test structure is used in the experiment to investigate the feasibility and performance of the strategy. The experiment targets the detection of the fabricated device, with a resolution of 1 aF with a 100 Hz bandwidth. The two main challenges in this case are noise and parasitic capacitance.

- **Noise.** The fF-level capacitive sensor has high impedance. For example, if the excitation signal is at 100 kHz, the impedance of 1 fF capacitance is about 1.6 GΩ, and consequently, the current signal is at nA-level for a 10 V peak-to-peak supply voltage. Moreover, the capacitance variation is even smaller. To successfully measure such a weak signal, noise control is a key challenge [8]. A measurement principle to suppress noise, especially the 1/f noise (flicker noise) dominating the noise level at a low frequency band, needs to be applied.

- **Parasitic capacitance.** Sensitivity to parasitic capacitance is a severe drawback of capacitive sensors [2]. For fF-level capacitive sensors, this issue is even more serious. The parasitic capacitances between the electrodes and the ground potential, usually at pF-level, are 1000 times larger than the sensor itself. These parasitic capacitances decrease the sensitivity of the sensor system and make the signal too small, and therefore their influence must be reduced [9].

In order to address these challenges, lock-in architecture [10] is chosen to measure the fabricated device. The schematic diagram is shown in figure 9.

The test structure is put in the middle of the schematic diagram with its suspended blade connected to the ground potential. Two anti-phase sine waves are fed into the sensing electrode pair and the reference electrode pair as drive signals. Thus, the variation of the equivalent capacitance between the sensing electrodes is amplitude-modulated into current variation at the frequency of the driving signal \( (f_{\text{drive}}) \) where thermal noise dominates. This current signal is then converted into corresponding voltage by a charge amplifier, and fed into a lock-in amplifier for demodulation. This shifts the sensor signal back to its low frequency baseband, while the 1/f noise of the charge amplifier is shifted to \( f_{\text{drive}} \). In this way, the 1/f noise, which originally dominates the noise level at a low frequency band, is separated from the sensor signal and can be filtered out by a low pass filter (LPF). Figure 10 shows the spectrum shift of the signal and noise during measurement. At the output of the lock-in amplifier, noise is determined by the thermal noise floor and the bandwidth of the LPF. The cutoff frequency of the LPF is selected according to the frequency band of the interested signal, which is 100 Hz in this experiment.

The chosen readout system is also insensitive to the parasitic capacitance. The parasitic capacitance at the input node \( (C_{\text{PI}}) \) is connected directly to the signal generator, and thus can be ignored. The sensor output node is connected to the inverting input node of the charge amplifier. When the
open loop gain of the charge amplifier is sufficiently large, the input node is a virtual ground, and the parasitic capacitor ($C_{P2}$) can be ignored as well. The reference capacitor ($C_R$) is driven by $V_{in-}$ to generate a compensating current, which reduces the offset signal induced by $C_P$. Therefore, the output signal of the lock-in amplifier is sensitive only to the capacitance change between the sensing electrodes caused by the insertion of the suspended blade.

For correct measurement, the charge amplifier and its feedback components need to be selected carefully. $C_F$ allows capacitive feedback, and feedback resistance ($R_F$) is used to provide a dc path for the feedback signal. To select a proper $C_F$, the following considerations need to be taken.

- **Sensitivity.** The output signal of the charge amplifier is determined by [11]

\[
V_{out} = V_{drive} \times \frac{\Delta C_S}{C_F}.
\]

The sensitivity of the measurement is determined by the ratio $V_{drive}/C_F$. In order to measure an fF-level variation of the sensing capacitor $C_S$ with a reasonable output signal level, the capacitance of $C_F$ needs to be small. Otherwise, a high voltage is required for the input signal $V_{drive}$.

- **Noise.** The noise density at the output of the charge amplifier can be written as [1]

\[
\frac{V_{noise}^2}{V_{nA}^2} = \left(1 + \frac{C_P}{C_F}\right)^2 + \frac{I_{nR}^2}{R_F C_F} \left|\frac{1}{\beta \omega C_F}\right|^2
\]

where $V_{nA}$ is thermal noise density of the first-stage field-effect transistor (FET) and $I_{nR}$ is thermal noise density caused by the feedback resistance $R_F$. Since a small $C_F$ brings more noise into the output signal, the capacitance of $C_F$ needs to be compromised to balance the sensitivity of the charge amplifier and noise level at the output for an optimized signal-to-noise ratio (SNR).

- **Frequency response.** The frequency response of the charge amplifier has a band-pass transfer function. In the high frequency region, the amplification is limited by the gain bandwidth product (GBW) of the amplifier; while in the lower frequency band, the amplification is influenced by the $RC$ time constant of the feedback components ($R_F$ and $C_F$). In the middle region, the amplifier has a stable closed-loop gain, which is suitable for amplification. The lower limit of this suitable frequency region is determined by the cut-off frequency of the high-pass filter formed by $R_F$ and $C_F$ in parallel. In order to keep this frequency limit low enough to allow the use of the function generator and lock-in amplifier, $C_F$ should not be too small.

Although a smaller capacitance of $C_F$ can provide a higher sensitivity, it, at the same time, will increase the noise level and limits and frequency response of the system. Considering the above-mentioned reasons, a $C_F$ of 1 pF is chosen in the experiment. To ensure a stable closed-loop gain of the charge amplifier, the GBW is the most important criterion of the amplifier. The required GBW can be written as

\[
GBW = \frac{1}{\beta} \times f_{close}
\]

where $f_{close}$ is the closed-loop bandwidth of the charge amplifier, which is typically required to be one order larger than the frequency of the driving signal, and $\beta$ is the feedback coefficient of the charge amplifier. It can be written as

\[
\beta = \frac{V_{input}}{V_{output}} \approx \frac{C_F}{C_F + C_{P2} + C_S + C_R} \approx \frac{1}{1 + C_{P2}/C_F}
\]

where $V_{input}$ and $V_{output}$ are the voltage signals at the input and output of the charge amplifier, respectively. As $C_S$ and...
$C_R$ are much smaller than $C_F$ and $C_{P2}$, their influence can be ignored. As can be seen from equations (6) and (7), a smaller $C_F$, related to $C_{P2}$, requires a larger GBW. In the real case, the parasitic component $C_{P2}$ is of the order of several pF-levels. Since a compromise $C_F$ is selected to be as small as 1 pF, the required GBW of the amplifier is, consequently, calculated to be higher than 10 MHz.

5. Results and discussion

A fabricated sensor is directly bonded to a printed circuit board (PCB) (figure 11), together with a charge amplifier. The charge amplifier is put close to the sensor to avoid large parasitic components generated by long connections. Two anti-phase sine-wave-driving signals and a synchronous signal are provided externally by a low-distortion signal generator. The amplitude of the driving signals is 20 V peak-to-peak to guarantee a mV-level output signal. The demodulation and low-pass filters are implemented by the lock-in amplifier (SR830). The time constant of the lock-in amplifier is chosen to be 10 ms. An operational amplifier MAX4489 with a GBW of 42 MHz, a $C_F$ of 1 pF and an $R_F$ of 10 MΩ is used as the charge amplifier. The sensitivity is calculated to be 7.07 mV fF$^{-1}$ by using equation (4). The suspended blade is connected to the ground potential as previously discussed. A needle of the probe-station is used to control the capacitance variation by pushing the blade while observing its position under a microscope.

As the blade is connected to the ground potential, the equivalent capacitance of the $C_S$ path keeps decreasing during the insertion of the blade, while the anti-phase signal through the $C_R$ path maintains a constant value. The output of the lock-in amplifier, which detects the absolute difference between the two paths, has a monotonic increase. When the blade is at the ‘stand-by’ position, the output signal reaches the minimum value. When the blade is completely inserted, the output signal is maximized. Figure 12 shows the total change of the output signal (between the minimum and maximum conditions) as a function of the driving signal frequency. The measured outputs show a low frequency part of a typical band-pass frequency response of the feedback amplifier. The $–3$ dB frequency as the low-pass corner is 11 kHz, which is determined by the $RC$ time constant of the feedback loop. The frequency of the driving signal is chosen to be 100 kHz, where the charge amplifier has a stable closed-loop gain.

The insertion of the suspended blade between the two electrodes is divided into six steps. The corresponding output voltage of the lock-in amplifier for each step is shown in figure 13 with the blade position $x$ as the horizontal axis. At step 0 (‘stand-by’ position), the blade is far from the electrodes, and the signal leakage of the input signal through the blade to the ground is limited. The difference of the equivalent capacitance between the sensing and reference paths is small. This difference, together with the slight variation of the dimensions between the two electrode pairs caused by process non-uniformity, generates the initial offset of the output signal. During the insertion of the blade, the leakage of the signal becomes more significant, causing a continuous increase in the output signal. When the blade reaches step 6, the output signal is maximized, and at the same time, the equivalent
capacitance between the sensing electrodes decreases to its lowest level.

The relation between the output signal and the blade insertion is nonlinear. During steps 1–5, the front edge of the blade moves between the two electrodes, where the electric field is at maximum density. Consequently, this causes a rapid drop in the equivalent capacitance. This coincides with the measured curve, where the signal increases faster in the corresponding region.

The total increase of the output signal is 2.18 mV. Considering a sensitivity of 7.07 mV fF$^{-1}$, the corresponding capacitance change is 0.31 fF. The noise level at the lock-in output is about 0.67 μV Hz$^{-1/2}$, which is equivalent to a detection limit of 0.095 aF Hz$^{-1/2}$. Thus a resolution of 1 aF can be reached with 100 Hz signal bandwidth.

6. Conclusion

A silicon fF-level MEMS device is designed to investigate the measurement strategy for an fF-level capacitive sensor. The device is fabricated using a single mask bulk micromachining process, and provides a nominal capacitance of 1.2 fF. The behavior of the device is analyzed, and a high performance measurement setup based on the lock-in principle is developed. The measurement strategy presented can be applied to other capacitive sensors where capacitance variation as small as aF-level needs to be detected. The maximum capacitance variation of the fabricated device is measured to be 0.31 fF, with a capacitance detection limit at 0.095 aF Hz$^{-1/2}$.

Acknowledgment

The authors would like to thank Dr X Li and Professor K A A Makinwa from Electronic Instrumentation of TUDelft/DIMES for many useful discussions, and the DIMES IC-process group for technical support. This work is supported by the IOP program (IPT04003) of the Dutch government.

References