Self-Sensing of Oscillation in Parametrically-Resonant MEMS Mirrors with Uncertain Nonlinear Dynamics

Miki Lee  
Dept. of Internal Medicine  
University of Michigan  
Ann Arbor, MI USA  
leemiki@umich.edu

Cameron Harris  
Dept. of Mechanical Engineering  
University of Michigan  
Ann Arbor, MI USA  
harricam@umich.edu

Haijun Li  
Dept. of Internal Medicine  
University of Michigan  
Ann Arbor, MI USA  
haijunl@umich.edu

Xiyu Duan  
Optical Sensing  
Apple, Inc.  
Cupertino, CA USA  
dxy@umich.edu

Yi Chen  
Computer Vision  
Midea Emerging Technology Center  
San Jose, CA USA  
davidsky@umich.edu

Thomas D. Wang  
Depts. of Internal Medicine, Biomedical Engineering, and Mechanical Engineering  
University of Michigan  
Ann Arbor, MI USA  
thomaswa@umich.edu

Kenn R. Oldham  
Dept. of Mechanical Engineering  
University of Michigan  
Ann Arbor, MI USA  
oldham@umich.edu

Abstract—During operation in endoscopic imaging instruments, on-chip motion sensing by microelectromechanical system (MEMS) scanning mirrors can be difficult to perform due to high scanning frequencies and poorly known nonlinear behavior in system dynamics. In this work, we optimize a nonlinear model for on-chip capacitance simultaneously with a simple neural net model for unknown transmission line and circuit dynamics. Use of nonlinear model elements is observed to improve accuracy of predicted sensor output during ringdown, parasitic feedthrough, and multi-axis operation experiments. Calibration can be performed using a single iteration of ringdown and feedthrough training data. Motion estimation accuracy of approximately 0.1 rad of scan angle and 0.15 rad of phase delay is achieved, sufficient for diagnostic use in endoscopic imaging systems.

Keywords—Microelectromechanical Systems, Sensors and Sensing Systems, Identification and Estimation in Mechatronics

I. INTRODUCTION

Near-infrared fluorescence endoscopy has potential to improve cancer detection rates compared to traditional endoscopy. Many near-infrared fluorescence endoscopy systems utilize a microelectromechanical system (MEMS) scanning mirror to control the scanning pattern of a near-infrared laser, which excites fluorescent biomarkers in tissue. Electrostatic MEMS scanners, with motion driven by electrostatic attraction between arrays of capacitive comb fingers, have demonstrated especially high scanning frequencies and scan angles [1] [2]. Imaging is performed by scanning with multiple mirrors or mirror axes in a raster or Lissajous pattern.

As a form of single-pixel imaging, accurate knowledge of mirror position as a function of time during laser scanning has a substantial impact on image registration accuracy. Application of MEMS scanning mirrors to endoscopic imaging poses several challenges for position tracking or estimation. Many MEMS devices incorporate dedicated displacement sensing elements, such as separate capacitive electrodes or piezoresistors [3] [4]. However, extreme size constraints on endoscopy instruments (typically 1.2-5.5 mm in diameter) limit the ability to include added features on MEMS chips or to provide additional sensing connections. Alternately, if a scanner is known to be in motion, with approximate knowledge of amplitude and/or frequency, image processing techniques can be applied to infer attributes of steady-state motion, such as mechanical phase delay, without on-chip sensing [5].

Between extremes of dedicated on-chip sensing elements and no on-chip sensing, self-sensing of motion from the capacitive comb finger arrays used to actuate electrostatic mirrors is attractive. The most common obstacle to self-sensing in electrostatic MEMS devices is that driving voltages and currents are often much larger than those induced by changes in capacitance as a function of displacement. Nonetheless, self-sensing has been demonstrated in a variety of electrostatic MEMS devices, generally using modulation/demodulation strategies to separate displacement information from driving signals [6] [7] [8].

During self-sensing, an endoscopic setting introduces two challenges that are generally not present in prior works. First, the high scanning frequencies needed to achieve video-rate imaging reduce the feasible frequency separation between modulation signals and driving signals. Second, endoscopy systems feature comparatively long transmission lines between scanning mirrors and sensing circuitry. Imperfect and unknown electrical dynamics, and related non-idealities in demodulation...
and filtering circuitry, can combine to substantially reduce accuracy of capacitive measurements. Use of phase lock loops to track changes in resonant frequency has been demonstrated, but this does not provide additional information on amplitude and phase at the chip that can be used in image registration.

Here, we evaluate use of system identification and a simple neural net model to identify and compensate for unknown nonlinear dynamics in self-sensing circuitry for MEMS scanning mirrors used in endoscopy. We have previously described strategies for compensating for system uncertainties through detection of discrete, high-accuracy capacitive features [9] [10]. A limitation of those approaches is dependence on comparatively large amounts of calibration data. Use of a nonlinear model to capture both unknown capacitive information and transmission line/circuit dynamics can substantially reduce the amount of measurement data required to perform motion tracking. In this paper, we will introduce key mirror attributes, a proposed structure for modeling circuitry and nonlinear dynamics, and results of mirror phase and amplitude measurements from a parametrically-resonant electrostatic micro-mirror. Within the current work, self-sensing is performed as a signal extraction process, as opposed to using a dynamic observer or estimator, given a large number of unknowns in system dynamics, but potential use in combination with other methods is briefly discussed.

II. SYSTEM MODEL

A. Electrostatic Micro-Mirror

The electrostatic MEMS mirror used in this work is a multi-axis scanner excited in parametric resonance. Such scanners can be fabricated as purely planar structures with just three electrical connections: driving voltage for rotation about the x-axis, driving voltage for rotation about the y-axis, and ground. A gimbal structure separates inner x-axis mirror rotation (the fast axis) from outer y-axis mirror rotation (the slow axis) driven by corresponding capacitive electrode arrays. When resonance is excited in both axes, a Lissajous scan pattern is obtained for 2D laser scanning during fluorescence imaging. A representative example of a multi-axis, parametrically-resonant scanning mirror developed by the authors is shown in Fig. 1, of a class reported in [11].

For a given axis, capacitance is a nonlinear function rotation angle in that axis. For the purposes of this work, we will focus on fast-axis motion, assigned as the rotation angle about the mirror’s x-axis, as this is the more difficult axis to track effectively. Various functions can be used to approximate nonlinear capacitance, but for this work we use a simple Taylor expansion with respect to angle. For fast axis angle \( \theta_x \), capacitance \( C_x \) is given by

\[
C_x(\theta_x) = C_{0,x} + C_{2,x} \theta_x^2 + C_{4,x} \theta_x^4 + C_{6,x} \theta_x^6
\]  

where parameters \( C_{0,x}, C_{2,x}, C_{4,x}, \text{ and } C_{6,x} \) are constants that are unknown prior to system identification, due to complexity of the mirror’s comb finger geometry and uncertainty in exact finger dimensions following microfabrication. The capacitance function is assumed to be even, due to the planar symmetric structure of the mirror.

Mirror dynamics are those of a mass-spring-damper system driven by electrostatic attraction,

\[
J_x \ddot{\theta}_x + b_x \dot{\theta}_x + k_x \theta_x = \frac{1}{2} \frac{dC_x}{dx} V_x^2
\]  

where \( J_x \) is axis rotational inertia, \( b_x \) is a damping constant, \( k_x \) is a torsional spring stiffness, and \( V_x \) is driving voltage.

Parametrically resonant dynamics of this class of mirrors has been studied previously and are not revisited in detail here [12] [13]. However, for sensing circuit design, it is important to note a few points:

1) in steady-state operation there is an integer-multiple relationship between the driving voltage frequency, \( V_x \), and frequency of motion, with largest amplitudes typical when input frequency is approximately twice the natural frequency of the device. In other words, when a nominal driving signal, \( V_{x,d} \), is applied as a function of time, \( t \), with amplitude \( V_0 \) and frequency \( f_{in} \), or

\[
V_{x,d} = V_0 \sin(2\pi f_{in} t)
\]  

![Fig. 1. Representative multi-axis parametrically-resonant micro-mirror, showing large angle rotation of a central mirror structure (dual axes confocal endoscope mirror configuration; motion about y-axis captured by high-speed camera under optical microscope).](image1)

![Fig. 2. Parametrically-resonant micro-mirrors driven by electrostatic forcing typically exhibit softening behavior that results in bifurcated behavior in frequency response, and rapidly changing phase angle near the highest amplitude motions, illustrated schematically above.](image2)
output motion can be approximated as
\[ \theta_x \approx \theta_{x,0} \sin(2\pi \frac{f}{f_{\text{lo}}^2} t + \phi_x) \] (4)
where \( \theta_{x,0} \) is amplitude and \( \phi_x \) is phase delay of the axis's dynamics.

2) there is commonly a bifurcation in frequency response due to electrostatic softening effects. The highest amplitudes achievable by the mirror may be reached only by sweeping downward in frequency (“down sweep”), while the mirror may be driven at the same frequencies with negligible mirror motion if approached from lower frequencies (“up sweep”). This phenomena is illustrated schematically in Fig. 2.

3) phase angle, \( \phi_x \), is typically changing rapidly near peak resonance, with significant impact on proper image registration.

In practical operation, at a given excitation frequency, \( \theta_{x,0} \) and \( \phi_x \) can vary significantly over time due environmental perturbations on system damping and stiffness. In contrast, \( C_x(\theta_x) \) is effectively time invariant, as it is dictated by geometry of the mirror’s silicon structure. However, \( C_x(\theta_x) \) can only be identified through observation of mirror dynamic and/or self-sensing behavior.

**B. Sensing Circuit**

To perform capacitive self-sensing, a modulation-demodulation scheme is applied, with a periodic, high-frequency carrier wave, \( V_{x,c} \), added to the driving voltage \( V_x \), i.e.
\[ V_x = V_{x,d} + V_{x,c} \] (5)
A transimpedance amplifier with resistor, \( R \), and a high pass filter is then added to the mirror’s ground channel, to amplify high-frequency current oscillations at the carrier frequency. An envelope detector and low-pass filter convert those high-frequency voltage oscillations into an output signal collected by an A/D converter as the sensing circuit output. Under ideal conditions, with \( V_{x,d} = 0 \), wide separation of carrier and driving frequencies, and perfect envelope detection, the sensing circuit output, \( V_{\text{out,ideal}} \), reduces to a functional form
\[ V_{\text{out,ideal}} = AC_x(\theta_x) + B \] (6)
where \( A \) and \( B \) are constants [9].

To expand the model to better capture non-idealities of the sensor output, we first note that the current fluctuations at the carrier frequency arise from two sources: the time derivative of carrier voltage multiplied by capacitance (the primary signal) and the time derivative of capacitance multiplied by carrier voltage. In addition, due to large voltage amplitudes and modest frequency separation between driving voltage (20-30 kHz) and carrier frequency (100 kHz-1 MHz), a substantial component of the driving signal remains present following the envelope detector. The expanded voltage signal exiting the envelope detector, \( V_{\text{mid}} \), becomes
\[ V_{\text{mid}}(t) = a_1 C_x(\theta_x) + a_2 \frac{d}{dt} C_x(\theta_x) + a_3 V_{x,d} \] (7)
this signal then passes through the low-pass filter for an approximate linear system output, \( V_{\text{out,lin}} \)

\[ V_{\text{out,lin}}(s) = G_{\text{LPF}}(s)V_{\text{mid}}(s) \] (8)
The model for self-sensing output described by (7)-(8) represents a compromise between ideal filtered output in (6) and full expansion of frequency components in the proposed sensing circuit. Exact replication of all frequency components in the sensing circuit is difficult due to frequency-dependent transmission line dynamics. However, the proposed model is observed to capture major attributes of sensor output under special circumstances in which either \( \theta_x = 0 \) or \( V_{x,d} = 0 \), including comparative phase shift of terms in (7), and major frequency components present in the output signal. These comparisons will be elaborated on in Section IV.

We hypothesize that the additional nonlinear distortions and higher-order dynamics can occur due to transmission line delays interacting with non-ideal envelope detector and amplifier dynamics. Thus, we propose an additional model for signal output in which \( V_{\text{out,lin}} \) is affected by an unknown nonlinear process \( g() \) to produce a modified voltage output, \( V_{\text{out,nl}} \), represented in discrete time sampled data as
\[ V_{\text{out,nl}}(k) = g(V_{\text{out,nl}}(k - 1), V_{\text{out,nl}}(k - 2), ...) \]
\[ V_{\text{out,lin}}(k - 1), V_{\text{out,lin}}(k - 2), ... \] (9)
where \( k \) is sample index, i.e. \( t = kT \), for discrete sample time \( T \).

**III. Motion Tracking**

A. Objective

The goal of mirror motion tracking for endoscopic imaging is to reconstruct scan angles as a function of time, i.e. to find \( \theta(t) \). In this work, we seek estimated scanning amplitude, \( \theta_{x,0} \) and phase angle \( \phi_x \), to best approximate underlying sinusoidal mirror motion.

B. System Identification

To identify mirror model parameters, the mirror is operated under two scenarios in which mirror behavior is known:

- Passive settling with external displacement data, referred to as “ringdown”. During ringdown, \( V_{x,d} = 0 \) and rotation angle, \( \theta_{\text{rad}}(t) \) is an exponentially decaying sinusoid with amplitude available from an external reference. The corresponding sensor output signal is denoted \( V_{\text{out,rad}}(t) \).

- Driving signal measurement without motion, referred to as “feedthrough”. This measurement collects sensor output voltage, denoted \( V_{\text{out,ft}}(t) \), when the driving signal in (3) is applied at a frequency representative of motion operation, but the mirror motion can be avoided through careful use of bifurcations in parametric resonance.
Model parameters are then selected through minimization of an objective function,

\[ \mathbf{v} \mathbf{w} = \arg \min_{\mathbf{v} \mathbf{w}} \left\{ \lambda_1 \sum_{k=0}^{K} (V_{\text{out,rd}}(t) - V_{\text{out,rd}}(t))^2 + \lambda_2 \sum_{k=0}^{K} (V_{\text{out,ft}}(t) - V_{\text{out,ft}}(t))^2 \right\}. \quad (10) \]

Here, vector \( \mathbf{v} \) contains constants to be identified from (1) and (7), i.e. \( \mathbf{v} = [c_0 \ c_2 \ c_4 \ c_6 \ a_1 \ a_2 \ a_3] \), and \( \mathbf{w} \) contains constants from a shallow neural net intended to capture unknown dynamics in (9), as described below. \( V_{\text{out,rd}} \) and \( V_{\text{out,ft}} \) represent the voltage output predicted by the model, for a given test set of \( \mathbf{v} \) and \( \mathbf{w} \), \( \lambda_1 \) and \( \lambda_2 \) are weighting factors used to evaluate effects of placing greater emphasis on ringdown versus feedthrough measurements, \( K \) is the number of samples of each data set used in calibration, and \( \mathbf{v} \) and \( \mathbf{w} \) are the optimal parameter vectors identified by numerical optimization.

A single layer recursive neural net was tested for improved identification of unknown nonlinear dynamics. In the neural net, each of \( J \) nodes contains a nonlinear basis function, \( h_j \), computed over the sum of weighted values of the nodes and the modeled value for \( V_{\text{out,lin}} \), denoted \( V_{\text{out,lin}}(t) \), at the previous time step.

\[ y_j(k) = h_j \left( \sum_i w_{ji} V_{\text{out,lin}}(k - 1) + w_{j0} V_{\text{out,lin}}(k - 1) \right) \quad (11) \]

The modeled output with resulting nonlinearity model, \( V_{\text{out,nl}} \), is a final weighted sum of nodal contributions.

\[ V_{\text{out,nl}}(k) = \sum_i w_{ji} V_{\text{out,lin}}(k) \quad (12) \]

All weighting coefficients \( w_{ji} \) are collected in vector \( \mathbf{w} \).

During experimental micro-mirror characterization, equation (10) was optimized over \( \mathbf{v} \) and \( \mathbf{w} \) using a genetic algorithm. In this study, the basis functions \( h \) consisted of

\[ h_{1-5} = \{ \cdot \}^1, \{ \cdot \}^2, \{ \cdot \}^3, \{ \cdot \}^{1/2}, \{ 1 \} \quad (13) \]

selected manually from experience with common functional behaviors associated with modulation-demodulation techniques.

C. Motion Estimation

Following optimization of terms in \( \mathbf{v} \) and \( \mathbf{w} \) for a given pair of weights \( \lambda_1 \) and \( \lambda_2 \), resonant motion is estimated by minimizing error in system output without access to any external motion reference,

\[ [\theta_{x,0} \ \phi_x] = \arg \min \sum_{k=0}^{K} \left( V_{\text{out,nl}}(k) - V_{\text{out,nl}}(k) \right) \quad (14) \]

where \( V_{\text{out,nl}}(k) \) is generated through (7) and (11) with parameters fixed by \( \mathbf{v} \) and \( \mathbf{w} \) given hypothesized mirror motion of

\[ \theta_x(t) = \theta_{x,0} \sin(2\pi f \ t + \phi_x). \quad (15) \]

IV. RESULTS

Effectiveness of the proposed model and system identification procedures is evaluated in tracking of changes in amplitude and phase angle of a sample MEMS mirror under varying frequency. Results are compared across models with varying levels of hypothesized nonlinearity.

A. Experimental Setup

A test mirror was operated with the described sensing circuit while measuring angular detection with a benchtop-mounted photosensor array (On-Trak Photonics). Mirror driving voltage was delivered at a fixed 60 V amplitude with frequency and waveform controlled by custom operating software (LabVIEW). Three types of time series data were collected:

1) Ringdown measurements: Mirror voltage was cut to zero once steady-state amplitude was achieved, and motion and sensing circuit output collected during passive settling to rest.

2) Feedthrough measurements: Mirror voltage was applied with frequency selections chosen to prevent excitation of parametric resonance.

3) Resonance measurements: Mirror voltage was applied at varying operating frequencies with parametrically resonant motion.

System identification as described in Section III.B. uses off-chip displacement measurements from a single ringdown event, and knowledge of applied voltage \( V_d(t) \) in both ringdown and feedthrough testing. Motion tracking is performed using knowledge only of applied voltage \( V_d(t) \) during resonance measurements, with off-chip displacement measurements used for validation.

B. Results

Expanding sensor output models with the capacitive model described in (7)-(8) and unknown nonlinear dynamics to be captured by (11)-(13) is found to substantially improve agreement between modeled and observed sensor outputs. Figs. 4-6 show increasing model and experimental agreement during a sample ringdown experiment. Ideal capacitance modeling, as by (6), is insufficient to capture sensor outputs when mirror angle is small, as it neglects dynamic influence from prior behavior when sensor output voltage approaches zero (Fig. 4). Adding circuit dynamics and nonlinear capacitance improves model behavior (Fig. 5), but high amplitude behavior remains difficult to replicate. Finally, incorporating additional nonlinear dynamics (Fig. 6) provides a high fidelity fit between experimentally-measured and modeled sensor output, when calibrated with the off-chip angle data from ringdown.

The resulting model was then used to estimate amplitude and phase angles across the operating range of the micro-mirrors. As described above, estimated angle and phase were chosen to minimize error between measured and modeled sensor output voltage, given model parameters identified from a single ringdown experiment and a feedthrough voltage measurement at a single frequency (22.1 kHz).

We find that nonlinear capacitance modeling, as during ringdown testing, exhibits largest amplitude errors when motion
Fig. 4. Experimental and predicted capacitive sensing output during micro-mirror ringdown under ideal linear capacitance model in (6).

Fig. 5. Experimental and predicted capacitive sensing output during micro-mirror ringdown with nonlinear capacitance model and linear circuit dynamics, with model parameters calibrated to minimize combined feedthrough/ringdown error.

Fig. 6. Experimental and predicted capacitive sensing output during micro-mirror ringdown with nonlinear capacitance model and hypothesized nonlinear circuit dynamics as modeled with a neural net, with model parameters calibrated to minimize combined feedthrough/ringdown error.

Fig. 7. Estimated x-axis amplitude (top) and phase error (bottom) with identified nonlinear capacitance but without neural net.

Fig. 8. Estimated x-axis amplitude (top) and phase error (bottom) with identified nonlinear capacitance and neural net.
amplitude is highest, as shown in Fig. 7, with phase estimation accuracy of approximately 0.1 rad at most frequencies. Adding a neural net for additional nonlinear modeling improves high amplitude motion estimation, as shown in Fig. 8. The neural net also produces fewer large phase angle errors, though median phase estimation accuracy declines, to ~0.15 rad.

A comparison between estimated steady-state motion and measured motion at a sample frequency (22.3 kHz driving frequency) is shown in Fig. 9. Average amplitude of oscillation and timing is well-captured, though experimental measurements show visible coupling from the simultaneous slow-axis motion of the mirror. We believe this deviation to be instrumentation-related, arising from cross-axis sensitivity of the 2D photosensor array used to collect validation data. However, this limits the accuracy to which estimator performance can be measured.

V. DISCUSSION

A high-performance motion estimation system for micro-mirror scanning in endoscopy would provide high accuracy measurements of amplitude and phase delay of mirror oscillation with minimal calibration and training from data sets collected under controlled conditions. The approach proposed in this paper includes mechanisms for capturing nonlinearity of capacitance, feedthrough voltage from oscillations from imperfectly filtered voltage inputs, and nonlinear transmission need circuit dynamics, with limited need for external reference data and the ability to provide scanning amplitude estimates.

The advantage of the approach is that it can be implemented with modest training data, and appears effective in capturing behaviors in measured sensor data that are difficult to describe based on first principles capacitance or circuit modeling. System identification was performed with varying relative weighting in (10) between accuracy of ringdown modeling versus feedthrough modeling with training data (4 ms of time series data from each). Best results were obtained with modest emphasis on accurate feedthrough modeling ($\lambda_1 = 0.3, \lambda_2 = 0.7$). This reinforces the importance of high-fidelity reproduction of feedthrough effects in any capacitive sensing model with substantial parasitic voltage coupling, and proposes a means for doing so when faced with modest separation of driving frequencies and carrier frequencies during self-sensing.

Disadvantages of the proposed method include common limitations of system identification with very large numbers of model parameters. The combined model features both redundancies and potential overfitting that have not been attempted to be managed in the current work, and are unlikely to represent the physical structure of unknown dynamics. Given the importance of feedthrough effects noted above, it appears to be effective to capture observed behaviors without close attention to whether results (i.e., the neural net) represent a true physical structure. However, the large number of parameters still leads to slow genetic algorithm convergence, and occasional failure to find a close fit to sensor voltage, as in outlier data for identified phase delay at some frequencies.

For endoscopic imaging applications in clinical settings, amplitude motion estimates are sufficient to verify that a micro-mirror has achieved near-maximum parametrically-resonant amplitude, and ensure laser scanning amplitude appropriate to the target field-of-view of a given instrument. Phase delay estimates, however, are insufficient to perform image reconstruction at maximum resolution (pixel placement accuracy better than optical resolution limit) in absence of additional refinement. This can be performed through further tuning of phase delay to optimize image sharpness via autofocus-like techniques, with a final accuracy of approximately 0.4 mrad preferred for prototype endoscopy instruments [5]. Slow-axis (y-axis) phase and amplitude estimation is generally more accurate due to wider separation of driving and carrier frequencies, but it should be noted that self-sensing can only be performed in one axis at a time by proposed methods. In other words, during imaging the carrier signal can be applied to one axis or the other, periodically updating motion measurement from either axis for image registration.

In future work, the model proposed in this work might be synthesized with other approaches to partially compensate for limitations of respective methods. With additional mirror test data, model weights with substantial variation over time may be identified, and incorporated into an extended Kalman filter or other dynamic observer, for more efficient estimation and utilization of dynamic information.

VI. CONCLUSION

A combination of system identification and nonlinear neural net modeling is applied to capacitive self-sensing in a parametrically-resonant micro mirror. The model is intended to capture non-ideal behavior in transmission lines and circuitry that makes capacitive motion measurement difficult in the presence of large feedthrough effects from high driving voltages. The method is effective at estimating phase and amplitude of mirror to within approximately 0.1 rad of angular rotation and 0.15 rad of phase delay, sufficient for basic mirror diagnostic tasks and for providing initial motion estimates that can be further corrected using image processing-based metrics.
ACKNOWLEDGEMENTS

Research supported by the National Institutes of Health, awards R01-CA-249851 and R01-EB-020644.

REFERENCES


