A RESONANT EYE-TRACKING MICROSYSTEM FOR VELOCITY ESTIMATION OF SACCADES AND FOVEATED RENDERING

N. Sarkar[1,2], B. O'Hanlon[1,2], A. Rohani[1,2], D. Strathearn[1,2], G. Lee[1,2], M. Olfat[1,2] and R.R. Mansour[1,2]
[1]University of Waterloo, Waterloo, CANADA
[2]AdHawk Microsystems, Waterloo, CANADA

ABSTRACT

We demonstrate the first MEMS-based eye tracking system that operates in resonance to obtain 3300 eye position measurements per second, a 25-fold improvement over previously reported microsystems. The present system achieves at least a 10x improvement in bandwidth, volume (0.5cm^3), power consumption (15mW), and cost when compared to state-of-the-art camera-based systems, all without compromising resolution (0.4° RMS). A system with these specifications produces real-time velocity measurements within saccades, enabling on-the-fly prediction of fixations for the first time. This predictive capability will significantly reduce the power consumption of graphical processing units (GPUs) in augmented and virtual reality (AR/VR) headsets because the region of a display within the fovea may be rendered exclusively ("foveated rendering") [1].

INTRODUCTION

The human eye is capable of reaching angular velocities of 900°/sec during saccades, which are among the fastest movements that we can perform. The extraocular muscles are also precise enough to routinely generate microsaccades that subtend angles from 0.03° to 2° [2]. This combination of rapid and precise actuation serves as a "front-end" for the highest bandwidth sensory interface to the human brain: the visual system. It is therefore not surprising that the next major computing platform will rely heavily on visual inputs, a trend that is evident in the emerging Virtual and Augmented Reality (VR/AR) enterprise. A microsystem that precisely measures eye movements and gaze direction would be most welcomed in head-mounted displays, as it would dramatically reduce computational and graphical processing overhead through selective rendering, thereby increasing the bandwidth of human-computer interaction (HCI).

Conventional Eye Tracking Technology

The most widely used eye trackers today are video-based image processing systems. Such systems are bulky (>1cm3), expensive (>$100), power-hungry (>150mW), and slow (~120fps) and have therefore not been broadly adopted in mobile or wearable technology. Although medical-grade systems may obtain 500Hz tracking with 0.5° resolution, these systems are massive, may require head stabilization, and often carry price tags that exceed $50,000 USD. In addition, the resolution of camera-based systems is limited by deleterious effects related to pupil dilation and ringing in the transient response of the iris sphincter when excited by saccades. Eye-tracking sensors must be improved significantly to provide seamless HCI with connected devices in a mobile, wearable context.

Of the >50 commercial VR/AR headsets, there are none that offer un-tethered eye-tracking at the time of this writing, because of the high power consumption and computational expense of video-based methods. The present work may alleviate these constraints to enable wireless eye-tracking in mobile devices.

MICROSYSTEM FOR EYE TRACKING

The reported eye-tracking system achieves improvements in size, cost, power consumption, bandwidth and precision through the use of a simple design principle (Fig. 1). A low-irradiance (10-50 μW/cm^2), infrared (850nm), diverging (at ~50mrad) beam of light is steered by a scanner and reflected from the surface of the cornea at a glancing angle (60^0 to the normal) onto a photodiode.

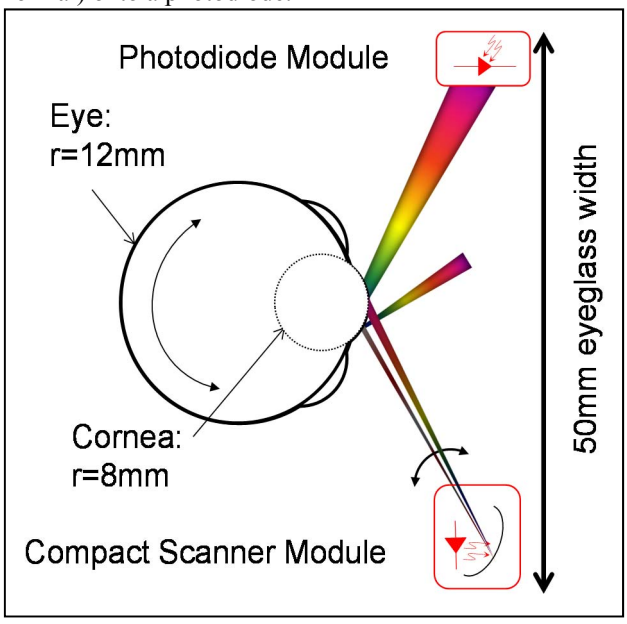

Figure 1: Eye tracker system geometry includes a MEMS scanner that sweeps a beam of light across the cornea and a photodetector that receives the reflected beam.

As the eye rotates, the system tracks the unique point on the cornea that maximizes the reflected signal received by the photodiode. A noteworthy aspect of the system is that the surface area of the photodiode acts as a spatial filter to remove high frequencies from the far-field pattern projected by the scanner. Large aperture micromirrors with flat surfaces that are ideal for imaging systems [3] are therefore not required. Instead, a small (300μm) Fresnel zone plate that projects a pattern of high spatial frequencies may be used. Prototype glasses with the hand-assembled eye tracking system are shown in Fig.2 along with a close-up of the scanner module which comprises a VCSEL and the MEMS device.

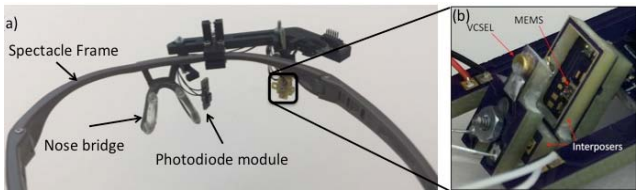

Figure 2: (a) Prototype glasses and (b) scanner module.

Several improvements to the system reported in [4,5] have augmented the robustness, speed, resolution, and ease of integration of the present system.

Device Geometry

The MEMS devices incorporate Fresnel zone plates that are mounted on 2DOF scanners (Fig.3). The CMOS-MEMS scanner utilizes alternating bimorphs and rigid beams in a serpentine pattern to achieve large angular deflections. Before singulation of the die, the proximity of the substrate gives rise to squeeze film damping and intermittent contact with the diffractive optic element (DOE) when it is operated at resonance with large angles. The die is therefore singulated with the DOE suspended over its edge. The natural frequency and offset position of the fast-axis are temperature-dependent, so the slow-axis must be operated isothermally to suppress temperature excursions. Isothermal operation has been shown to mitigate thermal coupling in scanning probe microscopes that achieve 1nm resolution in one axis over 20µm displacements in the orthogonal axes [6].

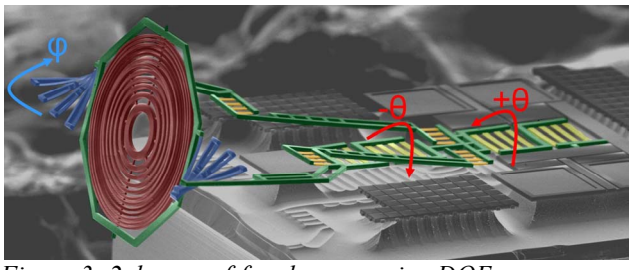

Figure 3: 2 degree-of-freedom scanning DOE.

The scanner design in Fig.3 is isothermal in one axis only. In commercial applications that must operate over large temperature ranges, the offset position and resonant frequency may suffer from thermal drift so the system may require frequent calibration. The scanner geometry in Fig.4 consists of opposed serpentine actuator pairs. Upon release, the device remains flat due to the symmetrically opposed torques applied by each pair of actuators.

To rotate the reflector, the temperature of one actuator is increased, while the opposed actuator's temperature is decreased proportionally. Drifts in the offset position of the device were not observable upon external heating to a temperature of 100°C. In addition, this design may be operated with a single PWM channel per axis, further simplifying its operation in a commercial product [5].

Experimental Results

The horizontal scan rate of the system is set to the resonant frequency of the fast-axis of the MEMS device which is at 3.3kHz. Each oscillation cycle of the zone-plate produces 2 peaks in the photodiode output that correspond to the corneal reflections captured during the forward and reverse trajectories of the scanned beam (Fig.5). A square wave acts as the drive signal for the MEMS device and is also used to reset a timer in a microcontroller. When the photodiode signal crosses a threshold, an edge-triggered interrupt captures the timer value and returns the horizontal position of the eye. The phase noise of this measurement translates to 0.4° RMS noise in the lateral position of the eye.

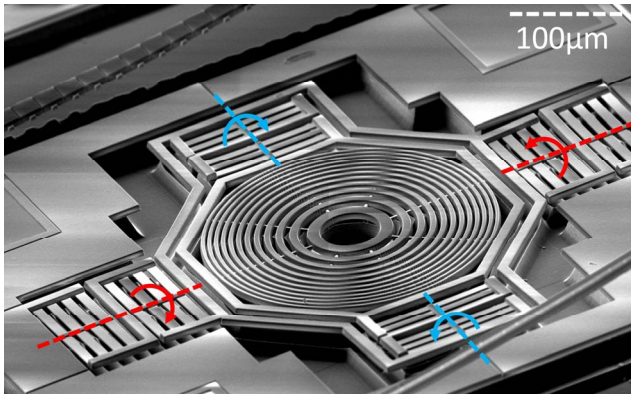

Figure 4: Scanning DOE with 2 isothermal axes manufactured in a CMOS process. CMOS-compatible voltages generate 90 degree (mechanical) deflections in both axes.

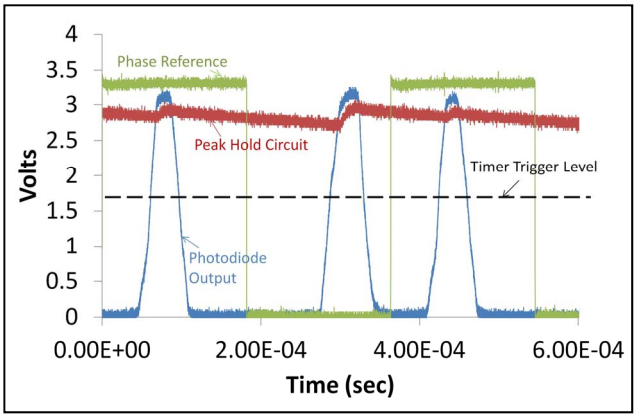

Figure 5: Oscilloscope capture shows the phase reference, photodiode output, and peak-hold circuit output used to track vertical position.

In [4] a hill-climbing algorithm was used to capture the vertical position (y-direction) of the eye. Peak amplitude measurements from the photodiode signal are susceptible to noise from a number of sources, so the present system obtains measurements at the inflection point in the vertical intensity profile (Fig.6). Because this is the highest-slope region in the photodiode's response, one obtains improved SNR when tracking this inflection point. The algorithm scans 3 lines in the vicinity of the global vertical maximum and sets the difference between the upper and lower measurements to zero in order to reveal the vertical position. The narrow beam profile of the corneal reflection in the vertical direction enables tracking with <1° resolution.

Eye movements on a number of users were captured using the system described above. The vestibulo-ocular reflex (VOR) data was taken while a user set their gaze to

a fixed point while rotating their head such that the eyes counter-rotated. The VOR produces the smoothest eye movements possible, as evidenced by the absence of saccades [Fig. 7]. Interestingly, if there is a mismatch between the VOR response and the position of the fixated object, the user may experience disorientation and nausea. This scenario may occur in a VR context if head-tracking and image refresh rates are not accurate enough to compensate for a subject's head movements. In fact, neural circuitry has been reported to adapt VOR gain upon prolonged exposure to such delays in VR, resulting in motion sickness upon re-entry into the real world. A system that constrains the position of a virtual object to the user's eye position within VOR movements may mitigate these effects.

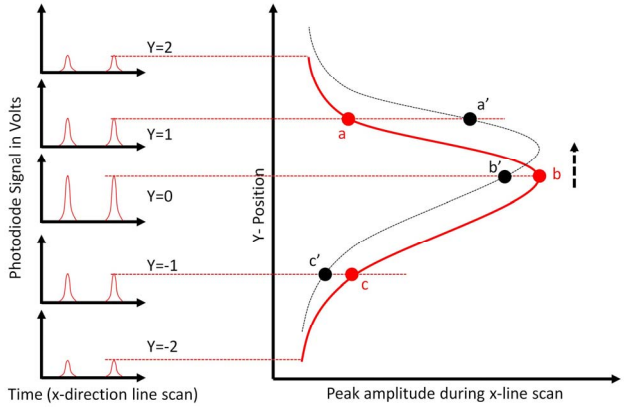

Figure 6: The envelope detected by the peak-hold circuit is tracked at inflection points to reveal vertical position.

A number of large saccades were also captured, by instructing the user to track a moving dot on a screen. The optokinetic reflex (OKR) was induced by translating a periodic pattern on the screen, producing alternating periods of smooth pursuit and saccades. The reported system samples the eye position with sufficient bandwidth to estimate velocity on-the-fly.

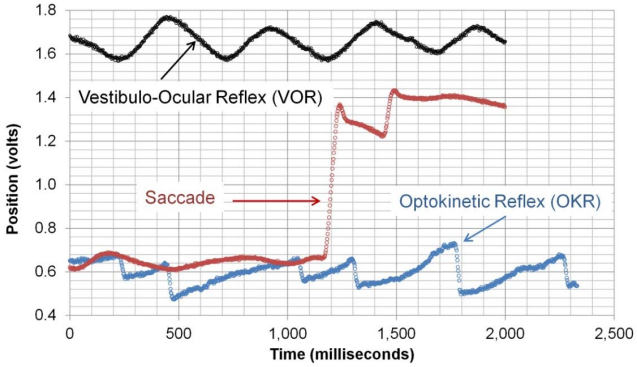

Figure 7: Various eye movement patterns captured by the microsystem with sufficient bandwidth for velocity estimation.

An important property of saccades is that their peak velocity is linearly related to their amplitude for movements up to ~20°, after which the velocity begins to plateau [7]. This is particularly convenient for the purpose of endpoint prediction. Fig.8 demonstrates that the reported system can measure velocity within saccades with sufficient resolution to predict the ensuing fixation, a feature that may dramatically reduce the power consumption of VR headsets through a technique known as foveated rendering. The key assumption here is that vision is temporarily suspended (or at least high-pass filtered) while a saccade occurs (20-200ms) and this may reduce the rendering burden in the display pipeline.

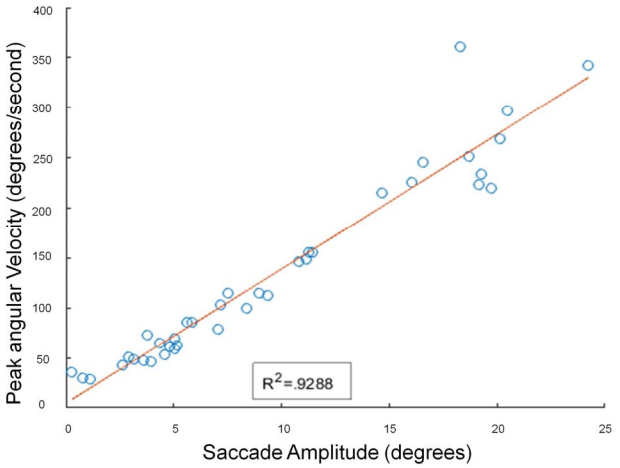

Figure 8: Saccade amplitude measured with the reported microsystem is linearly related to the peak velocity within the saccade, indicating that the post-saccade position may be predicted when the eye begins to decelerate, approximately midway through a saccade.

Practical Considerations

In a VR/AR product, several additional constraints may influence the design of the proposed eye tracking system. The field-of-view (FOV) is emerging as a key differentiator among headsets and may reach values of 110^0 in the near future. Placement of the scanner and detector modules sets the usable range of the eye tracker. Fig.9 illustrates that as the distance between modules is decreased, the usable range (where slope>0) increases at the expense of resolution (proportional to slope). Dual eye trackers may be used to increase the effective tracking range of the system for products with a fixed-focal-length display. Alternatively, an intermediate reflector may be used to expand the so-called eye box while placing the module outside of a headset's FOV.

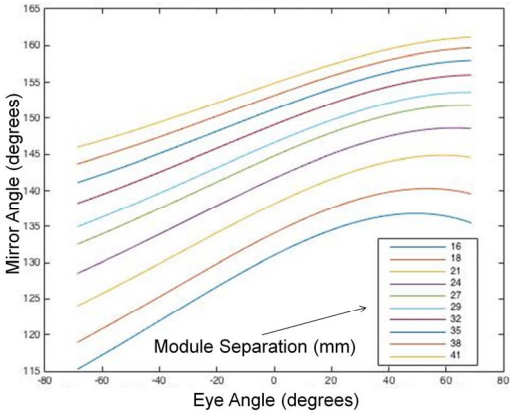

Figure 9: Plots of eye angle vs. mirror angle for various separations between scanner and photodiode modules. The slope of the line relates to the resolution of the system and ultimately constrains the tracking range.

Fig.10 is a plot of the photodiode output that is obtained with the user's gaze fixed while the system performs a full raster scan. The plot is truncated laterally

to illustrate the intensity profile of the corneal glint.

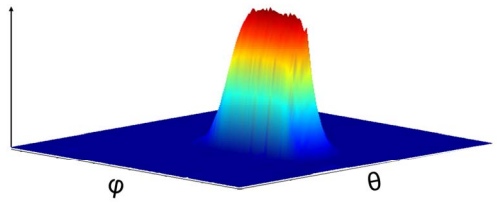

Figure 10: Intensity profile of corneal glint obtained by raster scanning the MEMS device.

The corneal position may vary widely among users of different ethnicities. The raster scan procedure was performed on 15 users of varying ethnicities and the locations of their global maxima are plotted in Fig.11.

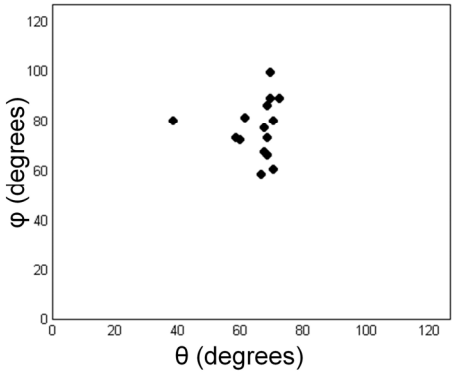

Figure 11: Corneal positions within the frame of reference of prototype glasses captured for 15 users of varying ethnicity.

In order to accommodate this variation, the scanning DOE must have sufficient range to cover most people. Once the location of the cornea is identified, the bias position of the scanner is re-centered and the range may be restricted. Tracking is performed over a ~15° range while the auto-adjust procedure may use a 90° range.

A polynomial relates the scanner position to the direction of the eye. This polynomial varies with module separation and corneal position and must be calculated for each user. Polynomial regression is performed to calibrate the measured eye position to a screen, as shown in Fig.12.

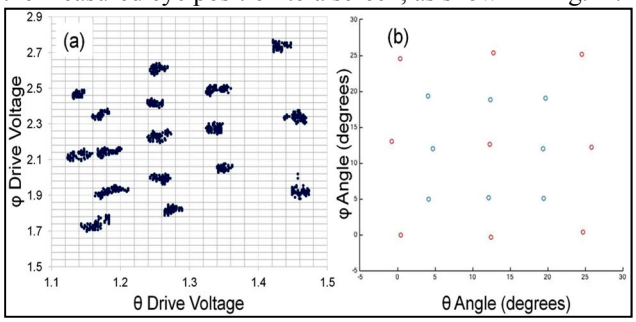

Figure 12: (a) Data captured while user's gaze was fixed on multiple points on a screen. (b) calibration is performed on inner (blue) points and post-calibration translated points are shown in red.

AR systems will be required to operate outdoors under bright sunlight. The present system has been confirmed to operate under these conditions as the photodiode is both shielded with an IR filter, and also AC-coupled to its pre-amplifier.

Small movements of AR glasses on a user's face may contribute significant drifts to the eye position measurement. Camera-based systems often use multiple imagers per eye in order to compensate for these drifts, and similar techniques may be implemented in the proposed system. Although the measurements presented here were obtained with DC illumination, modulation of the source will enable multiple modules per eye, reduce power consumption, and improve the SNR of the system.

CONCLUSION

Eye movement patterns reveal a wealth of information about neurological health, ocular health, state-of-mind, and interests. The study of microsaccadic eye movements has provided notable insight into the mechanisms of the oculomotor pathway [2] and may shed light on the pathogenesis of neural disease [8]. The peak velocity achieved in a saccade may quantify fatigue and intoxication in drivers or as a safety measure in the operation of heavy machinery. Specific eye movement patterns are associated with anxiety, attention, focus, interest, and the interpretation of emotions. In addition to these myriad applications, the reported high-speed, low-power, compact, precise and inexpensive eye tracking microsystem may serve as a non-invasive technique to obtain deeper insight into the state of users so that VR/AR hardware may deliver a more immersive experience.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the support of Ontario Ministry of Research Innovation (MRI), Canada Foundation for Innovation (CFI), the Canadian Microelectronics Corporation (CMC) and DARPA.

REFERENCES

[1] Daly, Scott. "47.3: Analysis of Subtriad Addressing Algorithms by Visual System Models." *SID Symposium Digest,* v32. no.1, 2001.

[2] M. Rolfs, "Microsaccades: small steps on a long way," *Vision Res.*, v.49, no.20, pp. 2415–41, Oct. 2009.

[3] O. Solgaard, *Photonic Microsystems*. Boston, MA: Springer US, 2009.

[4] N. Sarkar et. al, "A large angle, low-voltage, small footprint micromirror for eye-tracking and near-eye display applications," *Transducers 2015.*

[5] N.Sarkar et.al. "Scanning diffractive optic elements for untethered eye tracking Microsystems," *Solid-State Sensors, Actuators and Microsystems Workshop*, Hilton Head, SC, June, 2016

[6] D. Strathearn et. al. "A distortion-free single-chip atomic force microscope with 2DOF isothermal scanning," *Transducers 2015*, pp 2113-2116, 2015.

[7] Bahill, A., Clark, M., Stark, L., "The Main Sequence, A Tool for Studying Human Eye Movements". *Mathematical Biosciences*. 24 (3-4): 191, 1975.

[8] Anderson TJ, MacAskill M R. "Eye movements in patients with neurodegenerative disorders," *Nat. Rev. Neurol*. 9(2): 74–85, 2013.

CONTACT

*N. Sarkar, nsarkar@uwaterloo.ca