Capacitive Sensing-based Eye Tracking for XR Glasses

Aidan Hanson
hanson.ai@northeastern.edu
Northeastern University
Boston, USA

Amr Kassab
a.kassab@northeastern.edu
Northeastern University
Boston, USA

Mallesham Dasari
m.dasari@northeastern.edu
Northeastern University
Boston, USA

Abstract

Eye tracking technology has become increasingly vital as human-computer interaction utilizing extended reality (XR) technologies becomes more mainstream. Traditional eye tracking systems predominantly rely on cameras. While effective, these systems often suffer from high complexity, power consumption, and significant costs. This paper proposes a novel approach to eye tracking with capacitive sensing technology. Leveraging the principles of capacitance, we envision more accessible and efficient eye tracking solutions that can be integrated into various applications. This paper outlines the process of designing and analyzing a prototype of capacitance-based eye tracking.

CCS Concepts

• **Computer systems organization → Sensors**.

Keywords

Eye Tracking, Capacitive Sensing, Extended Reality (XR)

ACM Reference Format:
Aidan Hanson, Amr Kassab, and Mallesham Dasari. 2024. Capacitive Sensing-based Eye Tracking for XR Glasses. In *The 30th Annual International Conference On Mobile Computing And Networking (ACM MobiCom '24), November 18–22, 2024, Washington D.C., DC, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3636534.3697453

1 Introduction

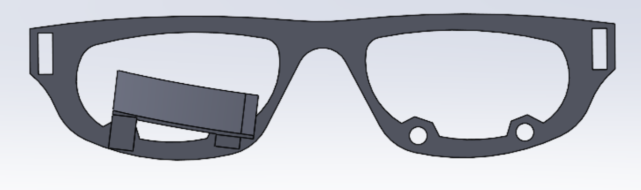

Figure 1: 3D model of our eye tracking glass prototype.

Eye tracking technologies have become pivotal in enhancing user experiences across various applications, including virtual and augmented reality (broadly extended reality— XR), human-computer interaction, and health monitoring. Currently, state-of-the-art eye tracking solutions predominantly rely on camera-based systems to capture gaze points with high accuracy [2]. However, these camera-based systems are known for their high power consumption, which poses significant challenges for integration with smart glasses that have limited battery capacity. For example, the Tobii Pro Glass 3 [1], which is one of the leading eye tracking glasses, can only last 1.75 hours with a 3400 mAh battery. When using a smaller battery, such as that of Google Glass (570 mAh), the tracking time drops to a mere 18 minutes. This limited operational time severely restricts the usability of these devices in everyday life, where continuous monitoring of gaze can provide valuable insights for applications such as mental and physical health monitoring, gaze-based input, and attention analysis.

To address the power consumption challenge inherent in camera-based systems, we propose CapTrak, a capacitive sensing-based eye-tracking solution designed for XR glasses. Figure 1 shows an example 3D model of the prototype. Capacitive sensing is known for its low power consumption, lightweight nature, and cost-effectiveness, making it an ideal alternative for prolonged use in wearable devices. Our approach leverages the principle that the human eye's movement causes variations in the capacitance around the eye region. By strategically placing four capacitive sensors around the eye—two for horizontal movement on either side of the eye and two for vertical movement above and below the eye—CapTrak can accurately infer gaze points based on the measured capacitance changes.

The CapTrak system translates the capacitance values into 2D vectors representing the pixel location on a screen where the user is looking. We propose a lightweight multivariate predictive model to process the capacitance data and predict gaze points with high accuracy. We designed our sensor based on Indium Tin Oxide (ITO) for electrode

---

ACM MobiCom '24, November 18–22, 2024, Washington D.C., DC, USA

ACM ISBN 979-8-4007-0489-5/24/11
https://doi.org/10.1145/3636534.3697453

material and custom-built a 3D-printed glass frame equipped with a Teensy 3.2 microcontroller for processing the capacitance data on the device itself. Our prototype does not need any other expensive computing capacity since our model is extremely lightweight.

Our preliminary results show that CapTrak achieves competitive accuracy (on average less than 10^o angular error §4) compared to traditional camera-based systems with a possibility of lower power consumption.

2 Related Work

Several wearable solutions were proposed to tackle eye tracking problem [2–4]. Numerous platforms, such as RealEye.io, GazeRecorder, and WebGazer.js, offer affordable solutions with acceptable tracking performance for everyday use. However, the fixed position and low resolution of webcams can impact performance, especially under varying lighting conditions, occlusions, and camera orientations. To improve accuracy and applicability, researchers have explored other non-wearable eye tracking technologies using higher resolution cameras, including RGB, infrared, and thermal cameras, as well as multi-camera systems for enhanced tracking coverage and user motion accommodation. These technologies, exemplified by products like Tobii Pro Fusion, provide high accuracy but are limited to fixed positions. Mobile device-based eye tracking allows some mobility but still requires users to hold the devices, leading to non-hands-free experiences.

To address these challenges, wearable eye tracking technologies have been developed, utilizing various sensors such as cameras, optical, acoustic, magnetic, EOG, and IMUs. Among these, camera-based wearables like Tobii Pro Glasses 3 and Pupil Labs are noted for their superior tracking performance and minimal calibration requirements. Despite their promise, current wearable solutions face limitations, particularly the high power consumption of camera-based systems, which restricts their practicality for everyday use.

3 Design of Sensor Apparatus

There are multiple ways to measure a changing capacitance [5], with differing hardware setups required to perform each. A technology readily available to perform this sensing is the hardware on a Teensy microcontroller (MC). In this research, specifically a Teensy 3.2 was used. This board has a functionality which connects one pin to one side of an internal capacitor, charges the capacitor, lets the capacitor discharge, and then repeats this cycle of charging and discharging for a short period of time. The result is that it records the number of charge/discharge cycles that were able to occur. Because the discharge time is dependent on the capacitance of the capacitor, a measurement of the number of completed cycles forms a relationship with the overall combined capacitance of the internal capacitor combined with any external features of the environment.

Noise Correction: Reading simply the raw output of the capacitive sensor provides data that has some undesirable noise. This is due to the high sensitivity of the equipment picking up extraneous unwanted minute fluctuations from the environment. To combat this and create a smoother output signal, an exponentially weighted moving average was employed. This strategy has the ultimate effect of reducing what would be perceived as noticeable jitter in the user experience. The decay constant can be tuned by the user to obtain a balance between responsiveness and noisiness.

3.1 Sensor Layout

To perform comprehensive eye tracking, the data received as input from the user's eye movements must be sufficient to distinguish vertical and horizontal movements, enabling coverage over 2D screen space or 2D angular viewing space. Because a single capacitive sensor produces just one output, obtaining a 2D final coordinate requires at least two sensors.

The initial setup involved two sensors placed to monitor one eye. The location selected for these sensors was gently making contact with the lower area of the lower eyelid. One sensor was positioned near the intersection of the lower eyelid and the nose, and the second was positioned at the edge of the lower eyelid, furthest from the nose. Both sensors were held in place using a sensor mounting platform attached to a custom frame. Figure 2 shows the overall design.

This setup was chosen for multiple factors:

- This location does not interfere with the user's view, allowing them to see easily over the sensor platform.
- The surface of the skin in this area is highly affected by eye movements due to the eye's control muscles being active in this area.

Testing this setup revealed that the sensors were much more sensitive when users looked away from the sensors and less so when they turned their gaze to the side of the apparatus where the sensors were present. The solution was to include a second set of sensors on the other eye, providing higher sensitivity where the first sensors could not. This resulted in the final sensor setup of two sensors per eye.

3.2 Interpreting Sensor Readings

From the previous discussion on the multi-sensor layout, it was known that four sensors would be used in this device, necessitating a method of mapping these values to a 2D screen coordinate. In addition to completing a system to work with 4 sensors, however, a general method for mapping an arbitrary number of sensors was developed.

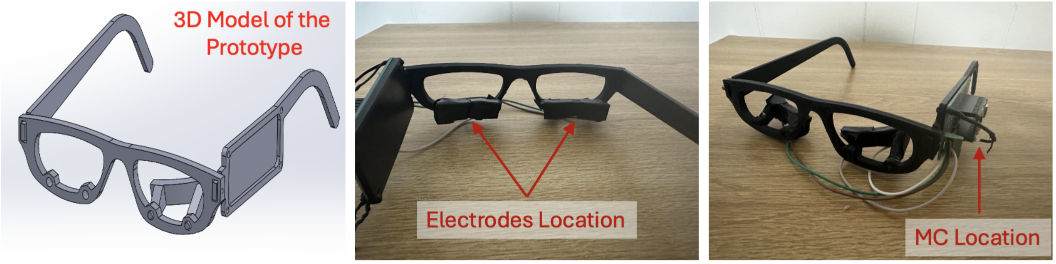

Figure 2: Design of our capacitive sensing-based prototype for eye tracking XR glass frames.

To create this mapping function, multivariate polynomial fitting was utilized. At the beginning of every trial, the user will follow an object on screen as it is moved by automation across the screen, landing at various calibration points. At each of these points, numerous samples are taken from each sensor and an average for each sensor is found. The result of this calibration period is a long list of screen coordinates and their corresponding sensor values.

Multivariate polynomial fitting, making use of matrix operations, is then utilized to calculate two sets of polynomial coefficients: one for the vertical axis and one for the horizontal axis. These coefficients can then be used to translate a set of four sensor readings into an x and y position predicting the user's gaze on the screen.

The goal of mapping sensor readings to the screen is to have two mapping functions, $M_x(s_1, s_2, ..., s_n)$ and $M_y(s_1, s_2, ..., s_n)$, where M_x takes in the n sensor values and outputs an x screen coordinate, and M_y takes in the same sensor values and outputs a y screen coordinate. As the user's eye moves and the muscles push and cause slight deformations in the skin of the lower eyelids, the sensors will have varying levels of pressures applied, causing different capacitance readings. Since the amount of pressure should experience no discontinuities as the eye muscles operate, it is reasonable to assume that as the capacitance measurements change, they do so in a continuous manner. Therefore, the mapping functions M_x and M_y can be reasonably approximated as multivariate polynomials of some degree n.

For calibration, samples are taken for many screen coordinates. Each screen coordinate will be stored alongside the four sensor values recorded for this coordinate. Using the x coordinate calculation as an example, a coordinate $M_x n$ corresponds to the four recorded sensor values $[s_{1n}, s_{2n}, s_{3n}, s_{4n}]$.

4 Preliminary Prototype and Results

The physical design of the complete device is very simple, reflecting the goal of making a simplified, low-complexity method of eye tracking. The frame is entirely 3D printed. Due to the fact that humans have different facial proportions, there cannot be a 1 size fits all approach to the sensor platforms. Because of this, the sensor platforms are built to be swappable. Different heights account for large or small nose bridges, ensuring that any user can have the sensors properly aligned. There is additionally a mounting platform for the electronics on the left earpiece.

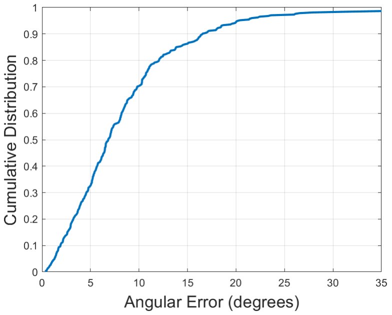

Figure 3: Distribution of Gaze angular error.

Figure 3 shows a cumulative distribution of gaze angular error. The error is measured based on a commonly used metric in eye tracking research [3]. We were able to achieve an average error of well below 10^o, which is suitable for many XR applications that do not require very fine-grained resolution. We are currently working on improving the accuracy further and making the prototype more robust for diverse users and settings.

The power consumption of this device was measured through supplying a regulated 3.3V voltage to the micro controller, while using an ammeter to measure the current being supplied to the board. The device was set to run a loop performing the same calculations it would perform at runtime, and the measured current was recorded as a steady 38.8 mA. At a voltage of 3.3V, this results in a power draw of 0.128 Watts.

References

[1] [n. d.]. https://www.tobiipro.com/product-listing/tobii-pro-glasses-3/
[2] Moritz Kassner, William Patera, and Andreas Bulling. 2014. Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction. In *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing: Adjunct publication*. 1151–1160.
[3] Ke Li, Ruidong Zhang, Boao Chen, Siyuan Chen, Sicheng Yin, Saif Mahmud, Qikang Liang, François Guimbretière, and Cheng Zhang. 2024. GazeTrak: Exploring Acoustic-based Eye Tracking on a Glass Frame. In *MobiCom*. 497–512.
[4] Tianxing Li, Qiang Liu, and Xia Zhou. 2017. Ultra-low power gaze tracking for virtual reality. In *SenSys*. 1–14.
[5] Robert Puers. 1993. Capacitive sensors: when and how to use them. *Sensors and Actuators A: Physical* 37 (1993), 93–105.