ESC499Y

Thesis Final Report

Exploring an Eye Movement Sensing Technique for Activity Recognition and Human-computer Interface

April 11th, 2019

Submitted by Cheng Lu
1000681988

Supervised by Khai N. TRUONG

Abstract

We are interested in developing a novel sensing modality for eye movement tracking and explore its applications to human activity interface. The very topic suggests this project is two-phased: a) Exploring and selecting a suitable sensing modality. b) Testing and evaluating the modality's application to human activity recognition and human-machine interface. In this document, we will show that we have completed the first phase of this research project and tentatively selected electrooculograms (EOG) for further exploration. This document seeks to demonstrate the validity of using electrooculograms to sense four different types of eye saccadic movements (left, right, up, and down) as well as discussing the pros and cons of the already explored but abandoned modalities. The document concludes by providing a discourse of the two research directions this project can pursue for the second phase: augmentation of the sensing modality and the implication of using EOG for self-monitoring.

Acknowledgments

I would like to express my gratitude toward my supervisor Professor Khai Truong for guiding me through this research project. I would also like to thank my colleagues in the Dynamic Graphics Project Lab for their support.

Table of Contents

Abstract ii
Acknowledgements iii
Table of Contents iv
Table of Figures vi
1 Introduction 1
2 Literature Review 2
2.1 **Camera-free Eye Sensing Modalities** 2
2.2 **Areas of Application** 4
3 Methods 5
3.1 **Prototyping** 5
3.2 **Data Collection** 5
3.3 **Data Analysis Methods** 6
3.3.1 **Signal Denoising** 6
3.3.2 **Evaluation of Two Signal Processing Techniques** 7
3.3.3 **Feature Extraction** 8
4 Results and Discussion 9
4.1 **Explored Prototypes** 9
4.1.1 **Microphone-based approach** 9
4.1.2 **Barometer** 10
4.1.3 **Electrooculography** 11
4.1.4 **Forehead EOG – Workout headband form factor** 13
4.1.5 **Forehead EOG – Adjustable headband form factor** 14
4.2 **Results and Analysis** 14
4.3 **Technical Limitations** 16
4.3.1 **Artifacts** 16
4.3.2 **Skin Condition** 17
4.4 **Discussion of Future Research Directions** 17
4.4.1 **Augmenting sensing modality** 17
4.4.2 **Self-monitoring** 18
5 Conclusion 18

6 Reference.....20

Table of Figures

Figure 1 Eye sensing techniques. 2
Figure 2 Data collection process 5
Figure 3. Results of signal processing using continuous wavelet transform and short-time Fourier transform.........6
Figure 4. Results of signal processing using continuous wavelet transform and short-time Fourier transform.........7
Figure 5. Feature matrix. Before and after the dimension reduction. 8
Figure 6 Microphone approach. 9
Figure 7 Barometer approach 10
Figure 8 Electrooculography 11
Figure 9 EOG Readings....................12
Figure 10 Headband EOG. 13
Figure 11 Headband EOG. 14
Figure 12 A example of EOG signal. 14
Figure 13. Results of classifying eye movements 15
Figure 14. An example of muscle movement artifact....................16
Figure 15. Acoustic sensing setup and results. 18

1 Introduction

Eye movements have been well studied, but they remain as a research interest in many fields, for two primary reasons. First of all, eye movements can convey useful information about humans. For example, continuous fixation on some objects could reflect an intense level of interest. An increase in the blink rate over time might show drowsiness[1]. Random and rapid eye movements (REM) during sleep has been proven to relate to the occurrence of vivid dreams[2].

Additionally, prior studies found that one could infer someone's daily activities from eye movements[3]. Aside from understanding humans, eye movements have well explored for its application to hands-free interactions, especially for helping people with special needs. For people with physical disabilities, eye gaze can serve as an alternative, low-physical-load way to drive an electric wheelchair[4]. For people with Amyotrophic lateral sclerosis (Also known as ALS. Patients with ALS gradually lose the ability to walk, talk, eat and swallow[5]), they can use eye gestured-based text entry interface to communicate with their loved ones[6].

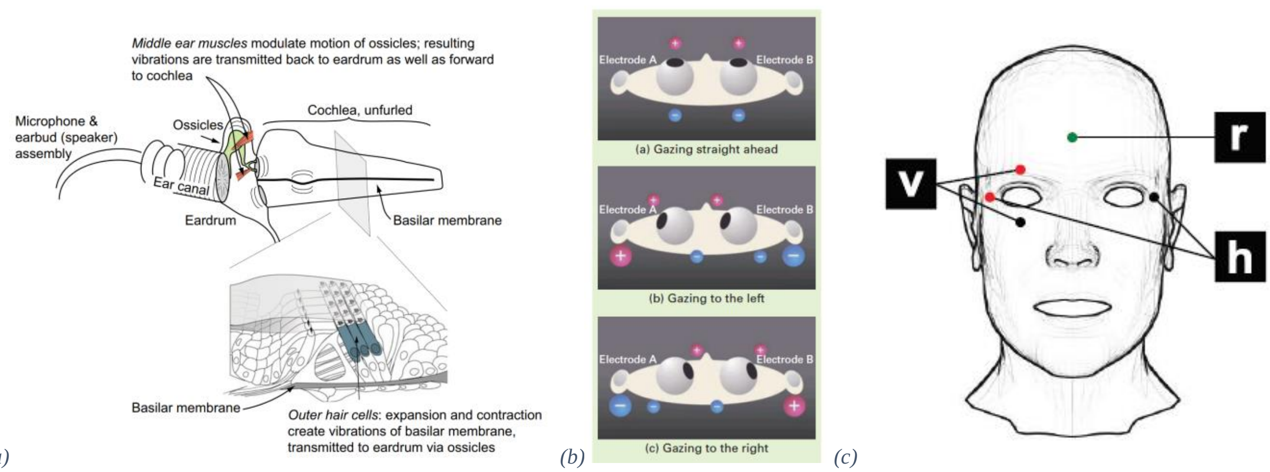
Many of the applications mentioned above heavily rely on camera as the sensing modality. The advantages of using a camera are evident: camera-based eye tracking tends to be highly accurate and robust. The downside of this modality is the steep cost and the bulky video processing unit that comes with the camera. For example, *Tobii eyeglasses 2*, a glasses-shaped eye tracker, can achieve an impressive gaze detection rate of 99% and less 5° in error. Still, its high cost, 10,000 USD, makes it only accessible to large corporations and research institutes and thus restricts its use scenario. A smartphone camera, used in [6], is a ubiquitous camera that has enormous potential for hands-free interaction. In most of its use cases, especially in mobile settings, a hand is required to operate the device, which makes smartphone camera impractical to be used for activity sensing, which is often done passively. Recognizing the limitations of camera-based sensing modalities, we proceed to investigate other types of sensing modalities. By the end of this project, we seek to develop a novel, robust, and low-cost sensing modality and explore its application to human activity recognition and human-machine interaction.

Over the first phase this two-phased project, we explored three different sensing modalities: microphone-based, barometer-based and electrooculogram-based modalities and tentatively arrived at our most current configuration (an electrooculogram-based approach) as shown in the *Results and Discussion* section, where we will justify the chosen configuration, an analysis of the data collected, and a discussion of future research direction.

The subsequent sections are organized as a *Literature Review*, *Methods*, *Results and Discussion*, and *Conclusion*. The *Literature Review* section is a comprehensive and in-depth discussion of the previously explored eye movement sensing techniques and areas of applications. The *Results and Discussion* section will summarize the eye movement sensing techniques we have attempted so far and outline the pros and cons of each technique. We will also justify the chosen configuration and an analysis of the data collected and a discussion of future research directions. The *conclusion* section will summarize the significance of my thesis work.

2 Literature Review

2.1 Camera-free Eye Sensing Modalities



(a) *(b)* *(c)*

Figure 1 Eye sensing techniques.

(a) In-ear microphone configuration in [7]*. (b) The working principle of EOG in* [8]*. Rotation of eyeballs causes one side to have higher potential than the other side. (c)Standard five-electrode EOG setup in* [9]*. "v" refers to the electrodes used for vertical eye movement sensing." "h" refers to the electrodes used for horizontal eye movement sensing. "r" is the reference electrode.*

Many works of literature [10][11][12] agreed that the human eye tends to exhibit five different types of movements: Fixations, saccades, vergence, smooth pursuit, vestibular ocular reflex. Fixation is the time when our eyes stop moving. Saccades refer to the rapid movement of an eye between two fixation points. Vergence occurs when the left and right eye move in opposite directions. Smooth pursuit allows eyes to closely and slowly follow a target. Vestibular ocular reflex takes place when eyes move in the opposite direction of the head. Fixation and saccades are dominant in static vision when we stare at static objects while keeping our heads still. Other than these five eye movements, blinks are a type of eye-related motion worth mentioning. Blinks refer to regular closing and opening of eyelids. In [7], it was discovered by

Gruters that the saccadic movements of eyes could induce distinct movement patterns of eardrums, which he termed *Eye-movement-related eardrum oscillations* (EMREOs). EMREOs could be observed in both human and monkey subjects, and the data could be collected using a very sensitive microphone placed in the ear canal. Inspired by [7], our more intuitive course of action was to sense the eardrum movements utilizing a microphone and attempt to map them to eye movements. However, after testing with a microphone prototype, as shown in **Figure 2(b)**, we were not able to identify any meaningful data, as the recording process was very susceptible to ambient noise level and peripheral muscle movements. We then investigated other means for eardrum movement sensing. Holst in [13] proposed that eardrum movement could induce air pressure changes in the canal, and air pressure was transducer for measurements. We also constructed a prototype, as shown in **Figure 3(b),** to test out this hypothesis. Unfortunately, we were not able to replicate the said results, at least not with an off-the-shelf barometer. This technique also suffered from inference problems, as the expansion and contraction of the ear canal muscle would also induce air pressure changes. Løkberg in [14] claimed that eardrum vibrations could be directly observed with a high-end television camera. Still, we did not attempt this approach because a television camera solution would be not only expensive but also too bulky to be portable.

After consulting literature and testing with prototypes, we made a tentative assumption that measuring eardrum movements was not the ideal approach to sense eye movements. Instead of measuring eardrum movements, [15][16] [17]and [18] took an alternative approach to implicitly measure eye movements: EOG (electrooculography). EOG takes advantage of the fact that human eyes are electric dipoles, with corneal (front of the eye) positively charged and retina (back of the eye) negatively charged. The rotation of the dipole will cause a change in the electric field, which can be measured by electrodes placed around the eye region, as shown in **Figure 1(b).** As standard five-electrode EOG configuration, as shown in Figure 1(c), could be challenging to put on and take off and also cause discomfort, a considerable research effort has gone into making EOG devices more portable and easier to use. Manabe in [15] instrumented a standard-sized headphone with ten electrodes and used it to detect horizontal and vertical eye saccadic movements. However, the accuracy of vertical sensing was suboptimal. Later in [16], he proposed a smaller, more portable earphone-type EOG device to plug into the ear canal, but it could only measure horizontal saccadic eye movements. Matthies in [17] attempted to detect saccadic eye movements with earplugs wrapped in the copper coil, but the accuracy was less than ideal. Favre-Félix in [18] showed that in-ear electrodes could be used to estimate horizontal eye gaze in real-time situations, and its performance could rival standard EOG setup. Heo in [19] demonstrated that the EOG signal could be reliably measured on the forehead with a headband instrumented with electrodes and mapped to saccades, blinks, and fixation. Heo further showed off a couple of simple demos, such as eye-based text input and controlling a wheelchair with blinks and saccades. It is worth pointing out that the current configuration of our system is a headband-typed forehead EOG device, as shown in Figure 6 because it satisfies the requirements of low cost and portability. We believe the forehead EOG configuration possesses an enormous potential that extends beyond the two simple demos proposed in [19].

2.2 Areas of Application

Human eye movements have long been explored for activity recognition and human-computer interface. For activity recognition, researchers would usually attempt to extract eye movement patterns unique to a particular activity class. Andreas in [9] classified five activity classes, such as copying a text, reading a printed paper, taking handwritten notes, watching a video, and browsing the Web, by analyzing data collected from users wearing a standard five-electrode EOG configuration. Years later, in [20], Andreas demonstrated the reliability of using EOG for activity recognition and eye-based human-computer interaction using a pair of EOG googles. In [3], Steil proposed an unsupervised method to discover user's everyday activities, such as reading, traveling, and eating, from their long term visual behavior by directly using an eye tracker. Instead of conventional activity recognition, Zhang in [1] proposed a forehead EOG configuration for driving fatigue detection by measuring the percentage of eye closure over time (PERCLOS). On the human-computer interface side, many works of literature have suggested interfacing eye movements with common electronics to enable hands-free interaction. [19][21][22] proposed using an eye-based interaction technique to provide accessibility to the population with special needs. [19] presented a simple demo of the text input interface and driving wheelchair using EOG, as mentioned in the previous section. Purwanto in [21]demonstrated controlling electric wheelchair used gaze direction and eye blinks, the image of which was captured from a camera mounted on the wheelchair. Tamura in [22] introduced an EOG-sEMG (surface electromyogram, a technique measuring bioelectrical signal generated by muscle movements) communication interface for patients suffering ALS (amyotrophic lateral sclerosis). ALS refers to a type of motor neuron disease, which gradually paralyzed the patient. Depending on the stage of the disease, someone will eventually lose the ability to walk, talk, eat, swallow, and ultimately breath[5]. Not necessarily designed for people with special needs, Ando in [23] showed an earphone-typed barometer device to register "face-related movements." Movements like opening/closing the mouth and turning the head left/right would cause a small deformation of ear canal muscle, which in turn caused a change in air pressure. Ando argued one could these movements could be interfaced with smart device control, such as switching music during cooking. Although he did not explicitly mention it, using the prototype we constructed as shown in **Figure 3(b),** we observed that movements like blinking eyes forcefully could also trigger a change in air pressure for the reason mentioned above.

3 Methods

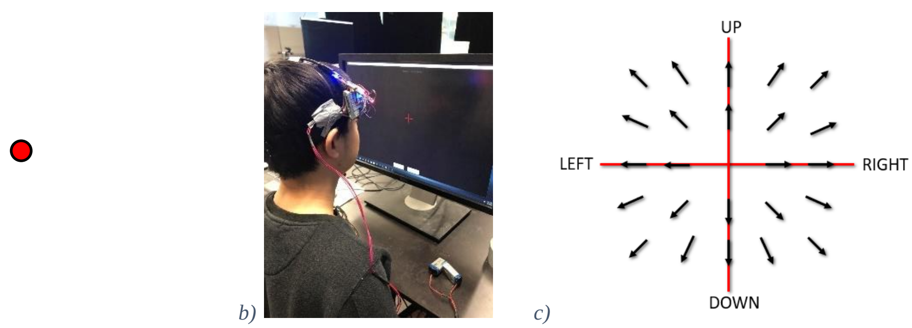
The method we implemented across this research project can be summarized into three categories: *Prototyping*, *Data Collection*, and *Data Analysis*. The following subsections will provide more explanations.

3.1 **Prototyping**

For the explorative purpose, throughout this project, we experimented with various prototypes, the details of which will be given in the *Results and Discussion* section. Although the specific criteria for evaluating a prototype might vary, we required the prototype to at least sense both the horizontal and vertical saccadic movements of the eyes to qualify for further exploration.

3.2 **Data Collection**

As mentioned in the introduction, we temporarily dwelled on our most current eye movement sensing configuration as it met the essential requirement we set out in section 3.1. For further exploration, we collected data from human subjects and attempted to evaluate its sensing robustness.



a) b) c)

Figure 2 Data collection process

(a) A moving-dot program. (b) The subject is sitting in front of the computer screen. (c) All the gaze positions.

For a preliminary evaluation of the device, we recruited five participants from the *Dynamics Graphics Project Lab*. During the data collection process, we asked the participant to wipe their forehead clean with napkins before putting on the device. Once the device was put on, the participant was allowed to adjust knob on the headband and tighten it to the level they would be comfortable with. We then asked the participant to sit still in front of a computer monitor and follow a moving dot, displayed on the screen, to their best capability. The participant was not physically constrained but was asked not to move their head.

For this particular process, there are 24 gaze positions the participant was expected to reach, including left, right, up, down, and diagonal positions, as shown in *Figure 2 (c).* Another program, running simultaneously in the background, recorded the EOG signal, and annotated the data automatically.

3.3 Data Analysis Methods

In this section, we illustrate the set of signal processing techniques we implemented to extract meaning features from EOG signals.

3.3.1 Signal Denoising

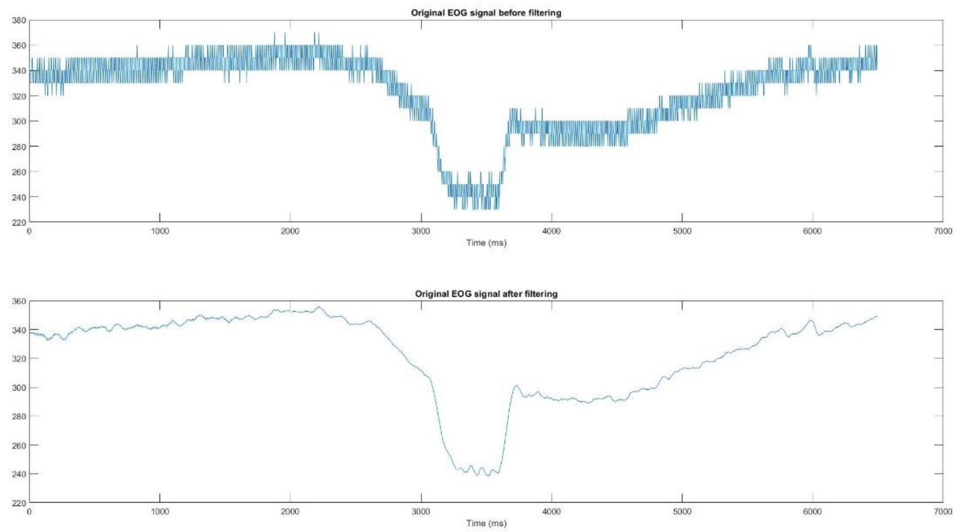


Figure 3. EOG signal before and after being denoised

For signal denoising, we applied a running average filter. We selected a window size of 30 data points through trial and error. The chose window size can smooth out most of the high-frequency noise. An example of the denoised signal is given in Figure 3.

3.3.2 **Evaluation of Two Signal Processing Techniques**

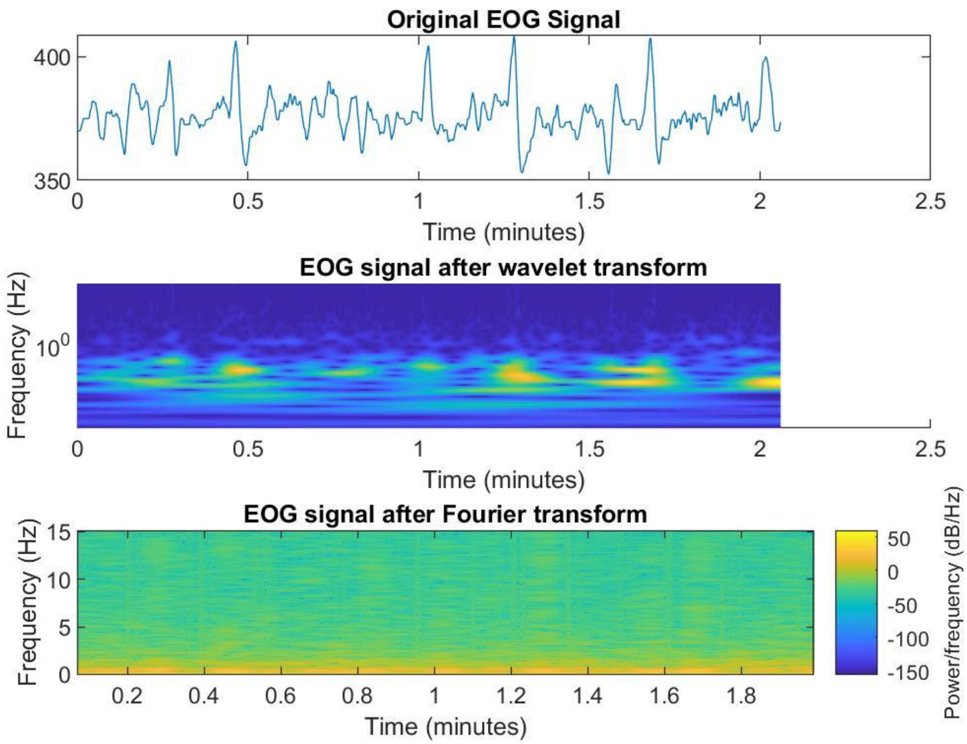


Figure 4. Results of signal processing using continuous wavelet transform and short-time Fourier transform

We attempted two techniques for processing EOG signal: short-time Fourier transform (STFT) and continuous wavelet transform (CWT). Fourier transform is a signal process technique that breaks up a waveform into a combination of cosine and sine waves and extracts features from these waves. Instead of cosine and sine waves, wavelet transform uses a set of small, transient, zero-mean waveform to extract features from a large waveform. The results obtained from the two techniques are displayed in Figure 3. As shown in Figure 4, the features are not visible in the STFT diagram, as Fourier transform typically works well with signals that are impositions of sine waves (acoustic signal, for example). For signals with abrupt local changes, like EOG, wavelet transform is a more suitable choice. As seen in the CWT diagram, the features are relatively well localized.

3.3.3 **Feature Extraction**

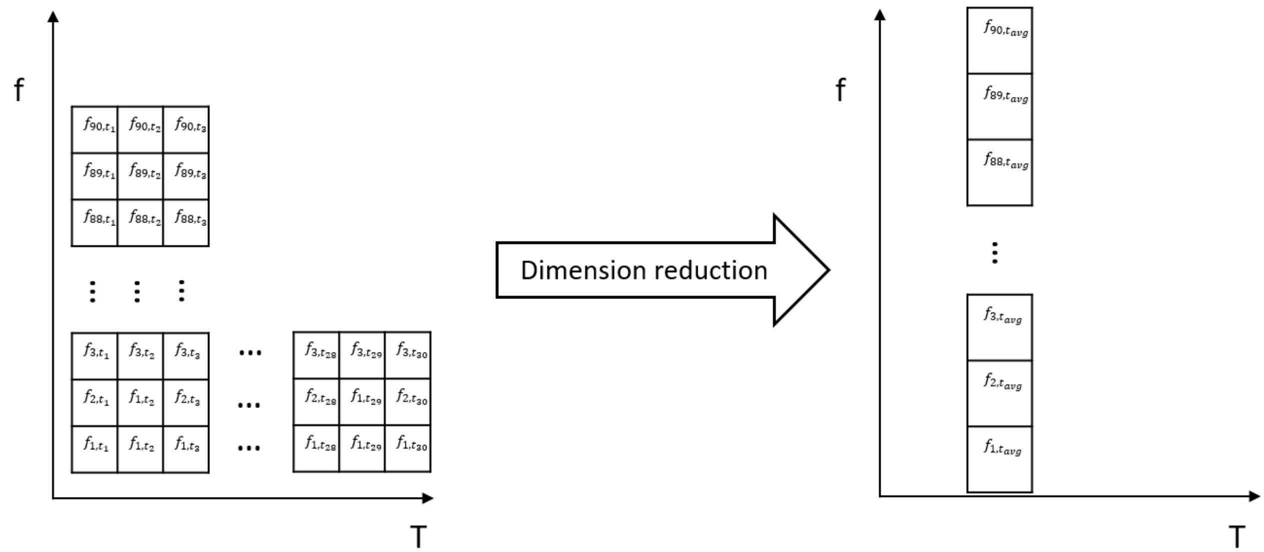


Figure 5. Feature matrix. Before and after the dimension reduction.

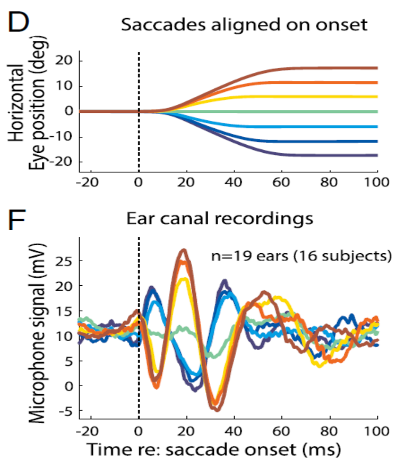
We seek to detect eye movements within a moving window on the spectrogram obtained from wavelet transform. In order to obtain meaningful features, it is important to choose the most appropriate window size. We used a wavelet transform with an analytic Morse wavelet to process the EOG data, which gave us 90 frequencies on the vertical axis. The effective sampling rate of the EOG signal was around 30Hz (around 30ms per data point). As the first attempt, we choose a 1-second window, which translates to around 30 points in the horizontal direction. In total, there are 90 x 30 = 2700 features within the window, as shown in the left image of Figure 5. To lower the computation cost, we attempted to reduce the dimension of the feature matrix. A simple solution was taking the average of the entries in the horizontal direction. Essentially, we transformed a 90x30 feature matrix into a 90x1 feature vector, as shown in the right image of Figure 5. With the feature vector generated and data annotated, we proceeded to classify the eye movement using a Support Vector Machine (SVM) classifier.

4 Results and Discussion

We have explored three different sensing modalities and five different prototypes throughout this project and arrived at our most current configuration. In this section, we will build on why we have arrived at our most current configuration and demonstrate the validity of using this approach to sense eye movements.

4.1 Explored Prototypes

4.1.1 Microphone-based approach



(a)

(b)

Figure 6 Microphone approach.

(a) Relation between ear canal recording and saccades in [7]. *(b) Picture of the prototype.*

As mentioned in section 2.1, inspired by [7], we first attempted the microphone approach and hoped to reproduce the data, as shown in Figure 2(a). A prototype was constructed for capturing the eardrum movements, as shown in Figure 6(b). The device was essentially a microphone wrapped in a rubber earmold made to fit into the ear canal. A smartphone was used for recording the microphone signal for postprocessing. Unfortunately, no meaning information could be extracted from the collected recordings. The reason was unclear, but as we further studied literature, we theorized that the environmental noise and peripheral muscle movements interfered with the data collection process. As eardrum was only able to oscillate at a very low frequency of 33Hz, it could be easily affected by environmental, typically ranging from 10Hz to 200Hz, and we lacked the means for reliable noise filtering. Additionally, the expansion and contraction of ear canal muscles, which were natural and sometimes involuntary muscle movements, could also produce sound. It might be difficult to isolate the sound patterns produced only from eardrum movements. Furthermore, [7] pointed out the data was obtained from a highly controlled environment, where participants were seated in a dark, sound-attenuating room with the head fixed on a chin rest. We believed that even if we managed to eventually collect the desired data, they might not be suitable for ubiquitous computing.

4.1.2 **Barometer**

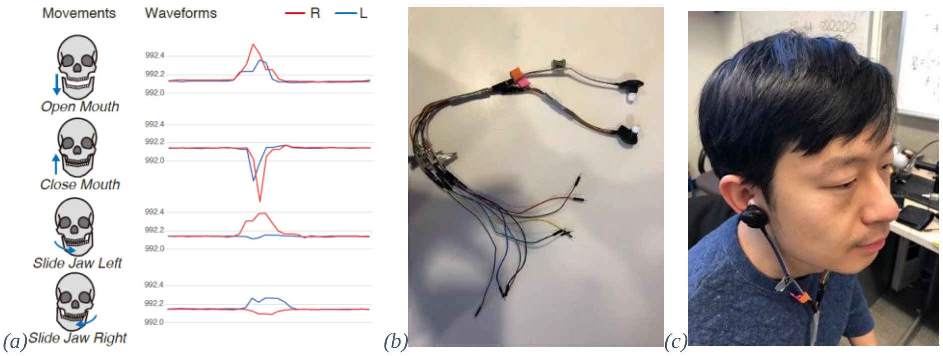


Figure 7 Barometer approach

(a) Some of the face-related movements and waveforms of the air pressure (hPa) of both ear canals corresponding to the movements in [23]. *(b) Picture of the prototype. (c) Subject wearing the device.*

The next approach we attempted was air pressure-based, inspired by [23]. According to ideal gas law, PV = nRT, as the pressure (P) inside a closed vessel changes inversely with respect to the volume (V) of the vessel. As the human eardrum is not air-permeable, if the opening is plugged with an air pressure sensor, then the ear canal will become a closed air vessel. We theorized the inward and outward movement of the air drum could cause the air inside the vessel to expand or contract. To verify the assumption, we constructed a prototype, as shown in Figure 7(b). The prototype was constructed by replacing the inner components of a pair of in-ear headphones with two BMP280 barometers. To ensure an airtight seal, the entire headphones were coated with liquid rubber except for the rubber tip portion, which would be fitted into the ear canal. This device could easily put on like a pair of normal in-ear headphones.

For data collection, we asked the subject to sit in a chair without any constraints for body motions and perform a set of gestures, as suggested in [23], such as sliding jaws and opening/closing mouth. In the end, we observed the said waveforms of air pressure like the ones shown in Figure 7(a). In addition to the gestures mentioned in [23], we also found that blinking eyes forcefully could also trigger a change in air pressure, similar to sliding one's jaw. We theorized it was because such a gesture inadvertently caused the ear canal to deform as the fascial muscle tissues are interconnected. However, beyond forceful eye blink, we were not able to detect any other eye-related movements, such as saccades and fixations. The technical limitations of the BMP280 barometer were identified as the reason. BMP280 had an accuracy to 12hPa(1200Pa) [R], but saccade-induced drum movements could only induce a peak to peak air pressure change of 42mPa($4.2x10^{-5}$)[R], which was far too small to be detected by BMP280. We attempted to replace BMP280 with more accurate barometers but could not find any off-the-shelf sensors that were as affordable and as small enough to fit into the ear canal.

4.1.3 **Electrooculography**

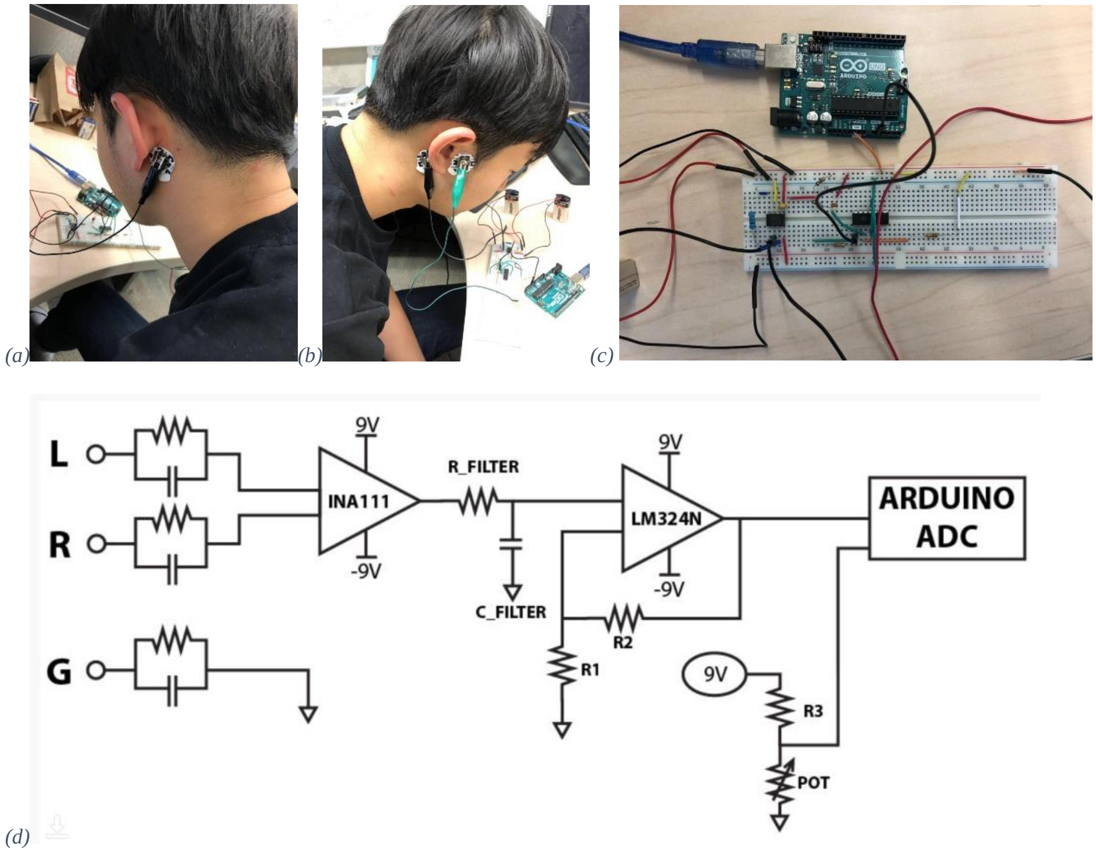
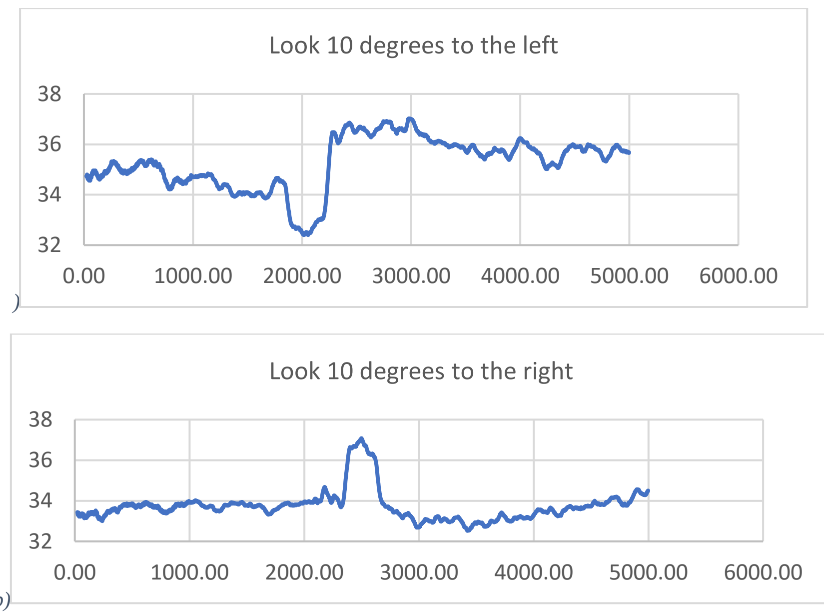


Figure 8 Electrooculography

(a) Subject wearing EOG (left side). (b) Subject wearing EOG (right side), The two black clamps are differential electrodes, while the green clamp is the reference(ground) electrode. (c) The hardware used for collecting EOG data. (d)The EOG sensing circuit. The contact point between the skin and the electrode is modeled as a capacitor in parallel with a resistor.

Electrooculography (EOG) is a technique that measures the potential changes across an eyeball due to its nature of being an electric dipole, as shown in Figure 1(b). Initially, we attempted a standard five electrode configuration as shown in Figure 1(c) and verified we could reliably collect data on eye movements. We used a very common type of AgCl gel-coated electrodes (also called “wet electrodes”). Considering the discomfort arising from sticking wet electrodes around the eyes, we explored other places for electrode placement and found that placing electrodes around ears could reduce discomfort and still sense eye movements. We then proceeded to collect data using the setup shown in Figure 8 (a) and (b). The bioelectric signals were sensed by electrodes and went through a band-pass filter(0.1Hz~1Hz) and amplification (gain = 500) circuit before eventually going into an Arduino ADC (analog to digital converter), which will map the input voltage into integer values between 0 and 1023.

We asked the subject to stick electrodes around the ears and perform a set of eye movements such as saccades and blinks. The raw signal was very noisy, so we applied an average filter with a window size of 50 data points to clear it up. Examples of the collected data are shown in Figure 9, and as one can observe, the waveform that corresponds to the eye movement is clear and distinct, although the eye blink seems to exhibit signal reading similar to saccades. This similarity is something we need to be aware of for future work.



(b)

Figure 9 EOG Readings.

Filter EOG signals (a) Look 10 degrees to the left. (b) Look 10 degrees to the right.

A limitation of this configuration was that it was incapable of sensing vertical eye movements, although the readings for horizontal movements were stable. We were unclear at this moment if activity recognition and human-computer interface could still be achieved even without the vertical eye movements, but in general, more information about eye movements should improve the performance of the system. Additionally, the conductivity of the wet electrodes would degrade over time as the gel dried up and had to be discarded. Thus, for more reliable sensing, we need to make a robust system that could sense both horizontal and vertical movements.

4.1.4 **Forehead EOG – Workout headband form factor**

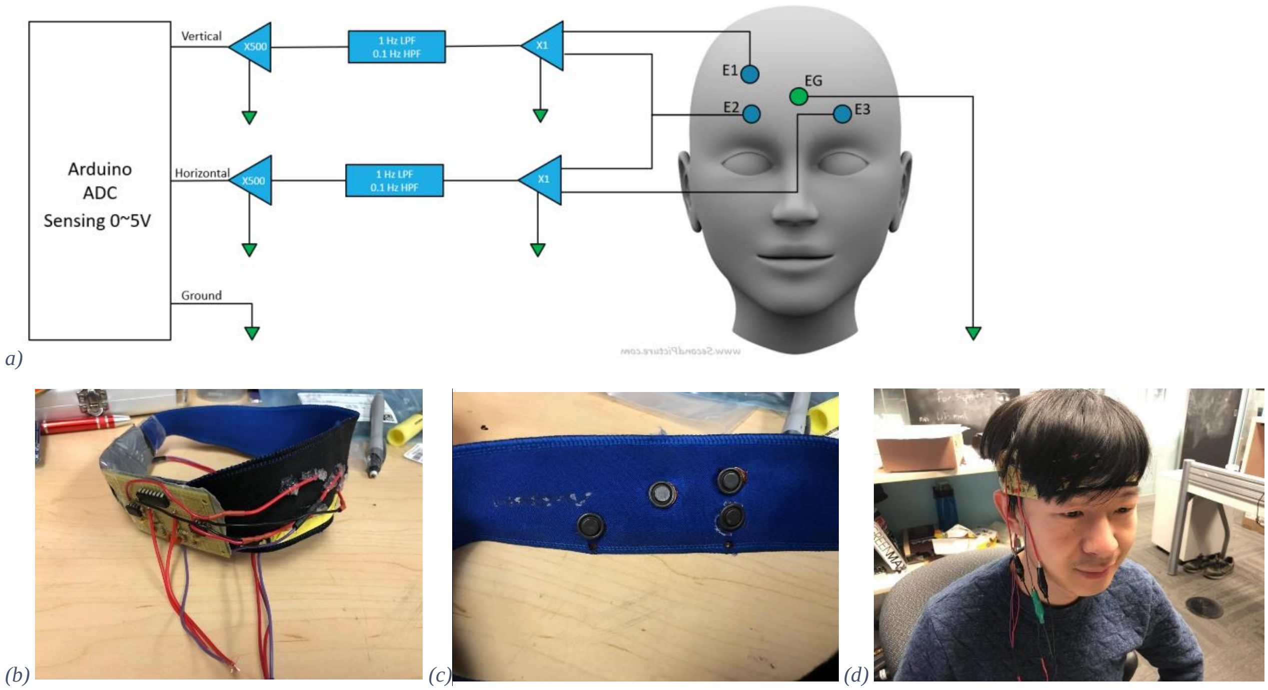


Figure 10 Headband EOG.

(a)The simplified schematic diagram of the EOG sensing circuit. (b)The front of the headband. Each electrode is connected to a band-pass filter(0.1~1Hz) and an amplification circuit (gain = 500) before going into an Arduino Nano ADC. (c) The back of the headband. There are four dry electrodes used for sensing eye movements: One reference electrode and four for sensing vertical and horizontal eye movements. It is worth noting that the vertical sensing electrodes and the horizontal sensing electrodes share one electrode. The purpose of this configuration is to simplify EOG circuitry. (d)Subject wearing the workout headband.

The headband configuration was built upon the previous EOG prototype. As mentioned in section 4.1.3, the limitations of the prototype were the inability to sense vertical eye movements and degraded performance of wet electrodes over time. To address the problems mentioned in section 4.1.3, we constructed a forehead EOG device with the form factor of a workout headband, as shown in Figure 10(a), to collect EOG data. This configuration was largely inspired by [19] and [1]. Unlike the previous prototype, we used dry electrodes, which do not degrade over time and are easy to remove and put on. Even though the device was able to sense both horizontal and vertical eye movements, there were some issues with its form factor. Some users found this headband too small to fit in and therefore were not able to use the device. Additionally, the softness of the headband material caused the workout headband to deform over time under tension, which negatively affected the area of contact between the skin and dry electrodes. As a result, we sought to explore another form factor.

4.1.5 Forehead EOG – Adjustable headband form factor

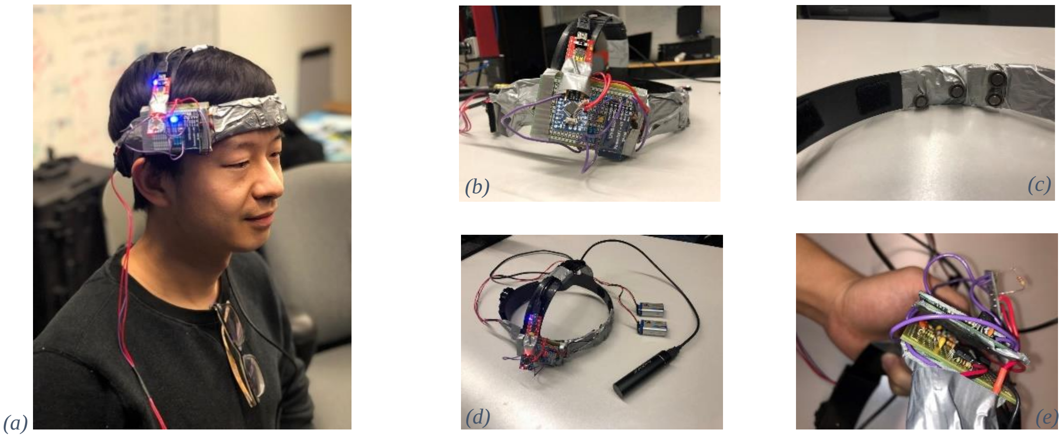


Figure 11 Headband EOG.

(a)Subject wearing the adjustable headband. (b) Arduino Pro mini along with the HC-05 Bluetooth module. (c) Four dry electrodes. (d) Power system. Two 9v batteries for powering the sensing circuit. One USB charger for charging the Arduino Pro mini. (e) The EOG sensing circuit.

Immediately following the workout headband configuration, we made several hardware improvements and arrived at our latest configuration. First of all, we instrumented the sensing circuit onto a plastic, adjustable headband, which made it possible for the device to accommodate different head sizes. The rigidity of the material also ensured the area of contact between the skin and electrodes stayed the same at all times. Additionally, we implemented a Bluetooth module as well as a portable power system. As a result, the data could be transmitted wireless and allowed usages of devices in mobile scenarios.

4.2 Results and Analysis

Figure 12 A example of the EOG signal.

Blueline indicates the horizontal EOG channel, while the orange line indicates the vertical EOG channel.

With the configuration we used in 4.1.5, we obtained some data. By visual inspection alone, we can already find that EOG signals exhibit distinct patterns in response to eye movements. As shown in Figure 12, the horizontal EOG signal will dip down in response to left saccadic movements and will rise in response to right saccadic movements. The vertical EOG signal will peak form a peak in response to downward eye movements. It will form a valley in response to upward eye movements. As we only recruited five participants for data collection at this point, it could difficult to recognize very subtle saccadic movements, so we simply divided the data into four classes: left saccade, right saccade, up saccade and downward saccade. In total, with five participants, we collected 24x5 = 120 eye movements. Within the 120 eye movements, with overlap, there are 50 left saccades, 50 right saccades, 50 upward saccades, and 50 downward saccades.

Overall misclassification rate	
Horizontal channel	Vertical channel
9.5%	10.5%

(a)

Individual misclassification rate		
	Horizontal channel	Vertical Channel
P1	1.9%	5.5%
P2	8.2%	15.0%
P3	4.4%	7.5%
P4	10.3%	14.6%
P5	21.4%	25%

(b)

Figure 13. Results of classifying eye movements

(a) Overall misclassification rate (b) Individual misclassification rate

We fed the feature vectors along with their annotations into an SVM classifier and performed 10-fold cross-validation. The misclassification rates are shown in Figure 13. From the results, one can make a few observations. The overall error is 9.5% for horizontal channel and 10.5% for the vertical channel, but the individual error rate varies considerably from person to person. This is most likely due to the fact we allowed users to tighten the headband to the extent they are most comfortable with. As a result, signal quality differs from one person to another and thus cause a discrepancy in the error rate. Additionally, it can be observed that the error rate of the vertical channel is consistently larger than that of the horizontal channel. We theorized that such a phenomenon is due to the blink artifact, as blink tends to trigger a signal response similar to that of an upward saccadic movement.

4.3 **Technical Limitations**

Even though we temporarily dwelled on our most recent sensing modality and form factor, it would be important to recognize the limitations of our prototype. Understanding technical limitations would help us identify the most suitable applications of EOG sensing.

4.3.1 **Artifacts**

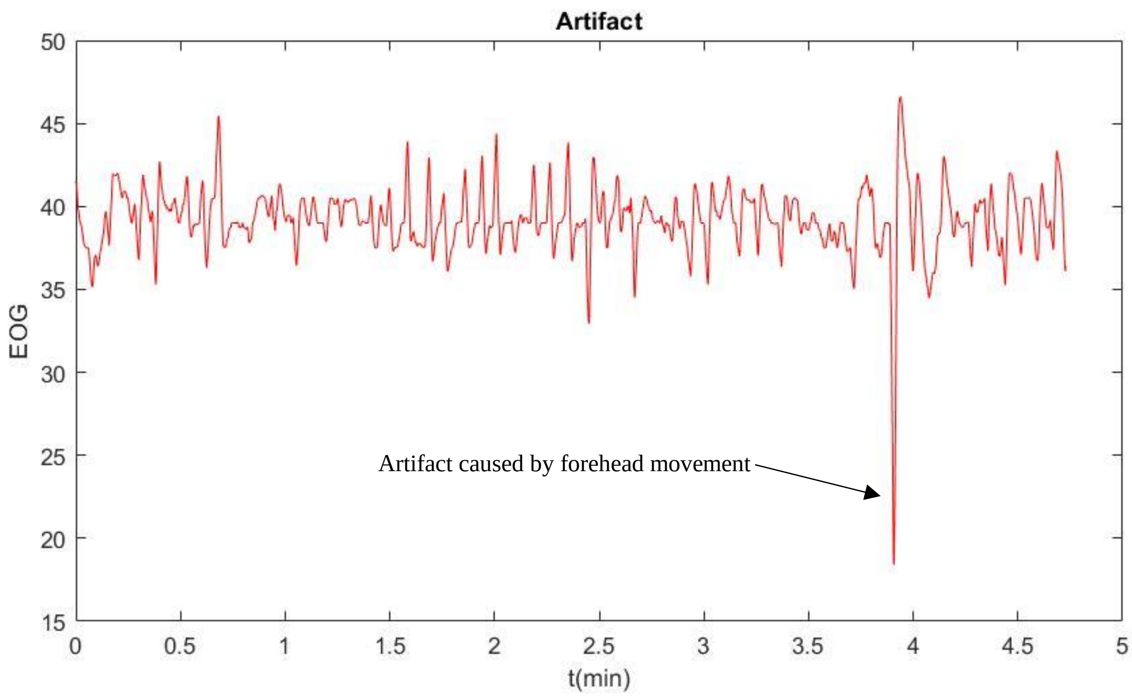


Figure 14. An example of muscle movement artifact

Although simple to implement, the EOG sensing modality is known to be susceptible to muscle movement artifacts. Since muscle movements generate bioelectricity, certain facial movements might trigger a response from the EOG sensor, depending on the specific EOG sensing configuration. Our headband configuration is especially very susceptible to forehead movements. Whereas for earphone configuration in section 4.1.3, the EOG sensing is more affected by chewing movement than by forehead movement. An example of the artifact caused by forehead movement is shown in Figure 14. Another source of artifact comes from blink, which is typically not used for activity recognition but still triggers a signal response similar to that of an upward saccadic movement. According to [16], activities like walking, speaking, yawning, and drinking might also create artifacts in the signal. However, at this point, we had only observed artifacts in our signal. In order to better understand artifacts, we might need to systematically study them in the future. Depending on the specific applications of EOG sensing, it might be possible we would need to remove artifacts from our signal or at least mitigate the effect of them.

4.3.2 **Skin Condition**

As our device reads off data from the electrodes that are closely in contact with human skin, the skin condition of the participant plays a vital role in determining the quality of the EOG signal sensed. Based on literature and also confirmed by our observation, in the most ideal case, the skin needs to be kept dry and clean. Moisture can increase the conductivity of the skins and therefore decrease the bio-potential sensed between contact points, while dirt or a strand of hair could affect the contact between the electrode and the skin and therefore produce poor readings. We believe skin condition is something we should be mindful of when designing application scenarios in the future. For example, it might be wise to avoid situations when participants would sweat excessively.

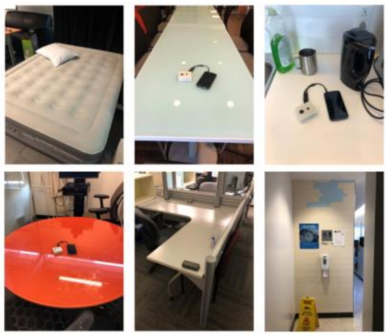
4.4 **Discussion of Future Research Directions**

The technical side of the workflow is quite straight forward, as it is essentially an iterative process of prototyping → data collection → analysis → feature extraction → eye movement recognition → activity recognition. In order to establish the contribution of this work, it is important to consider the research direction we should pursue for future work. In this section, we briefly discuss three potentially interesting research directions.

4.4.1 **Augmenting sensing modality**

Andreas mentioned in [9] that EOG sensing could sense structured eye movement patterns very well. For example, the eye movements consist of "sequences of small saccades occur while the eyes move over the words in a line of text" and "a large saccade when the eyes move back to the beginning of the next line of text." Thus, reading activity can be recognized with relative ease. Web browsing, however, is difficult is recognize because its associated eye movements tend to exhibit highly irregular patterns depending on how the contents are organized on the web page. A potential reason for this limitation is possibly the lack of input from the environment. If the environment where there the user is in is known, then the system could narrow down on the set of potential activities. For example, if one is in the living room, there is a higher chance of him or her watching TV than web browsing.

A widely used technique for sensing the environment is acoustic sensing. A sensor composed of a microphone and a speaker output s a distinct impulse wave to the environment can receive the reflected sound waves. Different environments tend to reflect the sound waves differently. Taking advantage of this fact, we can train the system to recognize environments.



(a)

	Bedroom	Dining room	Kitchen	Meeting room	Office	Washroom
Bedroom	98	0	1	0	1	0
Dining Room	0	99	1	0	0	0
Kitchen	2	1	91	0	6	0
Meeting room	0	0	1	99	0	0
Office	0	0	4	0	96	0
Washroom	0	0	0	0	0	100

(b)

Figure 15. Acoustic sensing setup and results.

(a) Six types of environments. (b) Confusion matrix generated by 10-fold cross-validation.

We experimented with this technique by placing a sensor module in six different types of environments, as shown in Figure 15 (a). For each environment, we collected 100 different acoustic samples. We then extracted the mean Mel-frequency cepstral coefficients and used them as features for training in a multilayer perceptron. We ran 10-fold cross-validation in the multilayer perceptron, and the results are shown in Figure 15 (a). Given the robustness of this sensing ability, environment sensing might be a promising direction to pursue.

4.4.2 **Self-monitoring**

Many works of literature [3][24][25]so far have mainly focused on the technical side of activities sensing. From a human-computer interaction point view, considering the low cost and portability of the device, it might be interesting to explore what users can do with the EOG sensing device and how useful they find the knowledge of the sensed activities. The potential area worth looking into is self-monitoring. For example, it is worth investigating if one's knowledge of his or her daily activities could lead him or her to a healthy lifestyle.

5 Conclusion

In this project, we have explored different sensing modalities and constructed prototypes to test out the robustness of each sensing modalities. We explained the steps we have taken to arrive at our most up-todate configuration and justified selecting adjustable headband as the current form factor of the device. We collected data from five participants and demonstrate the validity of using EOG for eye movement recognition by sensing four types of saccadic eye movements. Before we concluded the document, we briefly talked about two interesting directions we can purse for future work.

6 Reference

[1] Y. F. Zhang, X. Y. Gao, J. Y. Zhu, W. L. Zheng, and B. L. Lu, "A novel approach to driving fatigue detection using forehead EOG," *Int. IEEE/EMBS Conf. Neural Eng. NER*, vol. 2015–July, pp. 707–710, 2015.

[2] ScienceDaily, "Rapid eye movement." [Online]. Available: https://www.sciencedaily.com/terms/rapid_eye_movement.htm. [Accessed: 28-Jan-2019].

[3] J. Steil and A. Bulling, "Discovery of everyday human activities from long-term visual behaviour using topic models," *Proc. 2015 ACM Int. Jt. Conf. Pervasive Ubiquitous Comput. - UbiComp '15*, pp. 75–85, 2015.

[4] A. A. Nashte, P. S. S. Savkare, and P. S. G. Kole, "Automation of Wheelchair Using Iris Movement," pp. 33–36, 2016.

[5] ALS Canada, "What is ALS? - ALS Society of Canada." [Online]. Available: https://www.als.ca/about-als/what-is-als/. [Accessed: 28-Jan-2019].

[6] X. Zhang and M. R. Morris, "Smartphone-Based Gaze Gesture Communication for People with Motor Disabilities," pp. 2878–2889, 2017.

[7] K. G. Gruters, D. L. K. Murphy, C. D. Jenson, D. W. Smith, C. A. Shera, and J. M. Groh, "The eardrums move when the eyes move: A multisensory effect on the mechanics of hearing," *Proc. Natl. Acad. Sci.*, vol. 115, no. 6, pp. E1309–E1318, 2018.

[8] H. Manabe and M. Fukumoto, "Using Earphones to Perform Gaze Detection for Warable Interfaces," *NTT DOCOMO Tech. J.*, vol. 12, no. 3, pp. 12–17, 2006.

[9] A. Bulling, J. A. Ward, H. Gellersen, and G. Tröster, "Eye movement analysis for activity recognition using electrooculography," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 4, pp. 741–753, 2011.

[10] Tobii Pro, "Types of eye movements." [Online]. Available: https://www.tobiipro.com/learn-and-support/learn/eye-tracking-essentials/types-of-eye-movements/.

[11] "TAXONOMY: THE CLASSIFICATION OF EYE MOVEMENTS." [Online]. Available: http://www.opt.indiana.edu/v665/CD/CD_Version/CH2/CH2.HTM#Class of Movt. [Accessed: 06-Apr-2019].

[12] K. Rayner, "The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search," *Q. J. Exp. Psychol.*, vol. 62, no. 8, pp. 1457–1506, Aug. 2009.

[13] H. E. Holst, S. Ingelstedt, and U. Örtegren, "Ear drum movements following stimulation of the middle ear muscles," *Acta Otolaryngol.*, vol. 56, no. S182, pp. 73–83, 1963.

[14] O. J. Løkberg, K. Høgmoen, and O. M. Hoije, "Vibration measurement on the human ear drum in vivo," *Appl. Opt.*, vol. 18, no. 6, p. 763, 1979.

[15] H. Manabe and M. Fukumoto, "Full-time wearable headphone-type gaze detector," *CHI '06 Ext. Abstr. Hum. factors Comput. Syst. - CHI EA '06*, p. 1073, 2006.

[16] H. Manabe, M. Fukumoto, and T. Yagi, "Conductive rubber electrodes for earphone-based eye gesture input interface," *Pers. Ubiquitous Comput.*, vol. 19, no. 1, pp. 143–154, 2015.

[17] D. J. C. Matthies, B. A. Strecker, and B. Urban, "EarFieldSensing: A Novel In-Ear Electric Field Sensing to Enrich Wearable Gesture Input through Facial Expressions," *Proc. 2017 CHI Conf. Hum. Factors Comput. Syst. - CHI '17*, pp. 1911–1922, 2017.

[18] A. Favre-Felix, C. Graversen, T. Dau, and T. Lunner, "Real-time estimation of eye gaze by in-ear electrodes," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, pp. 4086–4089, 2017.

[19] J. Heo, H. Yoon, and K. S. Park, "A novel wearable forehead EOG measurement system for human computer

interfaces," *Sensors (Switzerland)*, vol. 17, no. 7, 2017.

[20] A. Bulling, D. Roggen, and G. Tröster, "Wearable EOG Goggles: Eye-Based Interaction in Everyday Environments," *Wear*, pp. 3259–3264, 2009.

[21] D. Purwanto, R. Mardiyanto, and K. Arai, "Electric wheelchair control with gaze direction and eye blinking," *Artif. Life Robot.*, vol. 14, no. 3, pp. 397–400, 2009.

[22] H. Tamura, M. Yan, K. Sakurai, and K. Tanno, "EOG-sEMG Human Interface for Communication," *Comput. Intell. Neurosci.*, vol. 2016, no. Dc, 2016.

[23] T. Ando, Y. Kubo, B. Shizuki, and S. Takahashi, "CanalSense+: Face-Related Movement Recognition and Identification System based on Air Pressure in Ear Canals," *{CHI} Ext. Abstr.*, 2018.

[24] and G. T. Andreas Bulling1, Jamie A. Ward2, Hans Gellersen2, "Robust Recognition of Reading Activity in Transit Using Wearable Electrooculography Andreas," *J. Polym. Sci. Part B Polym. Phys.*, vol. 46, no. 20, pp. 2249–2254, 2008.

[25] A. Bulling, C. Weichel, and H. Gellersen, "EyeContext: Recognition of High-level Contextual Cues from Human Visual Behaviour," *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, pp. 305–308, 2013.