Eye Tracking for User and Environment-Aware Headset Augmented Reality

Alex Quan Xu

Thesis submitted in partial fulfillment of the requirements for Graduation with Departmental Distinction in the Department of Electrical and Computer Engineering

Duke University
Durham, North Carolina
Spring 2022

1. EXECUTIVE SUMMARY

1.1 Objective:

The objective of my work was to study eye tracking in headset augmented reality (AR) with the goal of developing user and environment-aware AR applications, specifically through two aspects:

1.) Using eye tracking metrics to detect mental fatigue in headset AR users
2.) Measuring the impact of environmental lighting conditions on headset AR eye tracking efficacy

1.2 Related Work:

Video-Based Eye Tracking: Eye tracking encompasses the process of collecting measurements for estimating where an individual is looking, the rotational position of his/her eyes are, and the state of his/her eyes [9]. The least invasive method is through video-oculography, which uses infrared imaging to detect reference points in cornea reflections [8]. This kind of video-based eye tracking is becoming increasingly accessible as this technology is made commercially available as standalone desktop devices, head-worn devices, or integrated into wearables such as Microsoft Hololens, MagicLeap One, and Google Glass [8, 21, 26]. Previous works have considered algorithms that implement and improve pupil-detection in video-based eye tracking, generation of synthetic eye datasets for training machine learning-based detection models, and the accuracy and precision of standalone desktop and head-worn eye trackers [1, 2, 3, 6, 7]. However, these works do not cover how the efficacy of video-based eye tracking that is integrated into AR headsets may be impacted by environmental conditions.

Fatigue Detection via Eye Tracking: Mental fatigue refers to the feeling of tiredness people may experience during or after cognitive activities [12]. Previous works have studied psychophysiological measures that reflect mental fatigue, including electroencephalography (EEG) and eye tracking [12]. Previous studies have also focused on the use of unobtrusive methods (primarily eye tracking) to track mental fatigue during tasks such as driving as well as during natural viewing conditions for individuals of different age groups [12, 14, 16]. These studies have focused on standalone desktop or head-worn eye trackers. While there has been some work investigating the feasibility of using eye tracking to detect mental fatigue in head-mounted displays, especially virtual reality (VR) devices, no study has yet investigated mental fatigue detection via eye tracking in headset augmented reality applications [17].

Context-Aware AR: Context-awareness is one of the ways that would allow an AR application to adjust to environmental, human, social, and system factors [18]. Some goals of context-aware AR include making applications more ergonomic, reliable, and resource-efficient under a variety of diverse and potentially changing conditions. As AR applications become more ubiquitous, frequent context changes are expected as users move around to use their devices in different environments and for different purposes. Previous work has considered frameworks for developing context-aware mobile AR, as well as user interfaces that adapt to background or environment content [10, 19, 36]. These works do not yet consider how tracking performance may be impacted by environmental conditions, especially in headset AR.

1.3 Motivation:

Workers in many industries suffer fatigue at the workplace, impacting productivity, well-being, and safety. As AR technology advances, its applications will be more widely adopted by a variety of workplaces and users. While this is an exciting development with the potential to transform how humans interact with technology, this also presents the opportunity and challenge of understanding user fatigue in AR. Past studies have shown that eye tracking is an unobtrusive way to monitor mental fatigue [12-16]. Modern commercially-available headset AR systems are equipped with eye tracking functionality, but there has not yet been work focused on using eye tracking for monitoring users' mental fatigue while using AR applications [17].

Furthermore, AR has the potential to impact a wide-range of industries, ranging from industrial applications to consumer use. In order for AR to achieve wider mass-market adoption, it must perform reliability in a variety of diverse conditions [9, 12] . This includes eye tracking, a key functionality implemented in many commercially-available AR systems that enable not only detection of cognitive attributes such as engagement, mental workload, and fatigue (as demonstrated in the first study presented here), but also greater immersion through gaze-based user interactions [4, 14, 21]. Therefore, it is important to understand the impact of environmental conditions on eye tracking efficacy, as a first step to developing more reliable, context-aware AR. Past studies have shown that factors such as the users' age, room environment, eye color, contact lenses, and operating distance have an impact on the performance of desktop and mobile eye tracking systems [4, 6, 12]. However, no studies have yet to consider how environmental conditions impact eye tracking performance on headset AR systems.

My work focuses on eye tracking in headset AR, in particular using the MagicLeap One system. To the best of our knowledge, this is the first study that investigates the detection of user fatigue through eye tracking in headset AR, as well as the impact of environmental lighting conditions on eye tracking efficacy in headset AR.

2. BACKGROUND

2.1 Eye Tracking:

"Eye tracking" is a term that encompasses the process of measuring where an individual is looking (point of gaze), where his/her eyes are, and the state of his/her eye (including pupil size) [9]. Eye tracking has been used extensively for decades in the fields of cognition, medical, and psychological research [4]. Recent advances in technology have enabled more accurate eye tracking. There are three classes of existing eye tracking solutions:

Video-oculography (VOG): VOG is a video-based approach to eye tracking that is the most popular and widely adopted in commercial systems [8]. This method utilizes an infrared light emitting diode (LED) to transmit and reflect light off a subject's cornea, as well as an infrared-sensitive camera to capture images of his/her eyes [3]. The positions of these corneal reflections are almost constant during eye movements, so they are utilized as reference points for measuring pupil movement [3, 8]. Eye movement can be estimated by the change in the positions between the pupil center and corneal reflections [8]. VOG technology is often found commercially in three form factors: as standalone stationary trackers mounted beneath computer screens or on car dashboards; as standalone head-mounted devices (e.g. Pupil Labs Core); as integrated trackers in head-mounted displays [3, 4, 6, 8, 11, 14, 22].

Electro-oculography (EOG): Electric potential on the skin around the eye changes when the eye moves [8]. EOG utilizes electrodes placed on the skin to measure changes in electric potential, and these readings can be used to estimate angular position of the eye [3]. While this method is more computationally efficient, it suffers from poor ergonomics and accuracy due to interference from facial muscles and head movements [3, 8].

Scleral search coils (SSC): SCC requires the user to wear contact lenses embedded with magnetic field sensors [23]. A magnetic frame is then placed around the user, and the magnetic field sensors provide readings that can be used to estimate eye movement [24]. While SCC is the most accurate tracking approach, it is also the least ergonomic as it requires the use of topical anesthetics [8, 23, 24]. Thus it is impractical to use SCC in typical sensing scenarios in a commercial or consumer environment.

2.2 Eye Tracking in Augmented Reality:

Recent technological advances in infrared cameras that can produce high-quality images, consume a low amount of power, and are small in size have enabled the integration of video-based eye tracking functionality in head-worn devices such as smart glasses and AR/VR [5, 9, 21]. Eye tracking functionality integrated in AR has already enabled improvements in ergonomics, computing resource management, and power consumption [9]. For example, tracking where a user is looking allows an AR application to use a selective graphics architecture that only renders visuals in the area that can be seen, thus alleviating computational requirements, reducing power consumption, and enabling greater flexibility in device thermal management [9, 10]. State-of-the-art eye tracking in AR also allows for user intent prediction, leading to the design of intuitive user experiences that utilize this data [9-11].

Thus, eye tracking has become a core feature in modern headset AR systems including the MagicLeap One and Microsoft Hololens [21, 25]. The user-centered data that eye tracking provides also plays an important role in the development of *context-aware AR*, which aims to have AR applications that adapt to user, environmental, social, and system factors [18].

2.3 Gaze Estimation:

In video-based eye tracking techniques, infrared images of the eye are gathered and processed through an eye tracking pipeline that includes appearance-based (i.e. eye part segmentation) and geometric components (cornea, pupil, and gaze estimation) [3, 8]. Traditionally this is done using computer vision techniques that utilize eye landmark detection heuristics and geometric eye models [1, 3]. Recent studies have investigated the use of deep neural networks to estimate eye gaze, specifically from images captured from off-axis cameras, which is the case for many integrated eye trackers in AR headsets such as the MagicLeap One [1, 2]. There have also been studies that focused on alternative approaches, such as utilizing multiple low-resolution eye images combined with a learning-based gaze estimation network, or directly mapping 2D pupil positions to 3D gaze directions [3, 27]. The majority of previous work has focused on the use of standalone stationary VOG trackers [4, 6, 12]. In recent years there has been more research into the use of head-worn and integrated AR eye trackers for gaze estimation [12].

Furthermore, the state of a person's gaze may be classified as either fixation or saccade [28, 29]. Fixation is the process of maintaining gaze on a single location, while saccades are rapid movements of the eye [28, 29]. Regular eye movement consists of both fixation and saccades, often alternating between the two states [28, 29]. Gaze metrics available from the MagicLeap One headset include gaze points, gaze direction, and gaze confidence [21]. However, fixations and saccades are not available and require specialized eye trackers.

2.4 Gaze Estimation Quality Metrics:

There are four common components to gaze estimation data quality: spatial accuracy, spatial precision, temporal accuracy, and robustness [4]. Spatial accuracy refers to the difference between the true gaze point and the recorded gaze point [4]. Spatial precision is the deviation between repeated samples of gaze point location when the true gaze point is held constant [4]. Temporal accuracy refers to the difference in timing of recorded gaze events and when they truly occurred [4]. Lower differences indicate higher precision and accuracy, which is more desirable. Finally, robustness refers to the amount of recorded data relative to invalid, corrupt, or lost data during the measurement process [4].

Poor quality in gaze estimation can lead to numerous issues, especially when eye tracking is deployed in an AR application. For example, poor spatial accuracy creates errors in determining the true gaze point, leading to potential problems in AR interaction that rely on it [4, 9]. Poor spatial precision leads to noisy data and can negatively impact the performance of fixation/saccade detection algorithms or adaptive rendering architectures (as described previously) that rely on it [4, 9]. Low temporal accuracy can lead to delay or "lag" in these same features, greatly hindering application usability and performance [4, 9]. Finally, poor robustness directly reduces the usefulness of eye tracking, depending on the proportion of data lost [4, 9].

Previous studies have calculated spatial accuracy as the average distance between a target object where a participant is looking (ground truth) and the recorded gaze location [7]. For spatial precision, standard deviation was used along individual axes as well as in multiple dimensions [4, 7]. Past work has also shown that factors such as testing room environment, eye color, contact lenses, and operating distance may negatively affect these quality metrics [6, 12].

2.5 Fatigue Detection

Previous research in fatigue detection has mainly focused on two paradigms: (1) inferring mental fatigue through monitoring fatigue-correlated metrics such as task performance; (2) using psychophysiological measures [12]. The first may involve the usage of specific "fitness-for-duty" tests that assess a user on executive functions like attention, hand-eye coordination, or memory [32]. Or, alternatively no additional tests are required and mental fatigue may be inferred during regular activity using similar metrics [31]. While these methods have shown success, more recent studies have suggested that mental fatigue may not always impact certain fatigue-correlated metrics, such as task performance [12].

Thus, the use of psychophysiological measures have been suggested to be the more desirable approach [32]. These measures can be further divided into two categories: electroencephalography (EEG) and eye tracking [12]. EEG works by recording changes in brain waves (theta and alpha activity) that occur when an individual grows fatigued [32]. EEG has been shown to be highly accurate for fatigue detection, but this technique requires the use of bulky sensors that are time-consuming to apply and uncomfortable to wear in real-world daily use scenarios [12]. Also considering that eye tracking technology is already implemented in commercially-available AR headsets, it is therefore the more desirable option for fatigue detection.

Previous works have found several fatigue-correlated eye tracking measures: pupil metrics, blinking behavior, and oculomotor-based metrics [12]. Specifically, increased fatigue has been found to lead to reduced pupil size, longer and more frequent blinking, and slower saccade velocities, respectively [12, 33]. Of these, previous research has found that blinking was the best indicator of fatigue [12].

Blinking and pupil size can be easily tracked using commercially available eye trackers, including those integrated into AR headsets like the MagicLeap One [21]. Like the MagicLeap One, Microsoft Hololens also provides gaze tracking data, but it does not detect pupil diameter, which is why it was not chosen for the user studies presented [25]. However, saccades can not be measured by these relatively lower-cost eye trackers [12, 28]. More specialized trackers capable of high sampling rates (i.e. 1 kHz - 2 kHz) are required due to the rapid movement of the eyes during saccades [12. 28]. Eye trackers like the one built-in to the MagicLeap One are only capable of up to a 30 Hz sampling rate [21]. Thus, saccadic and other oculomotor metrics are difficult to use and not currently feasible for application to AR.

2.6 Implications of Mental Fatigue

Mental fatigue is a growing health issue that is associated with a high public health cost [12, 34]. Past studies have found that mental fatigue can negatively impact safety, behavior, and cognitive performance in the workplace [12]. Some estimates suggest that fatigue-related accidents have cost over $31 billion in the United States [12, 34]. Moreover, mental fatigue is associated with dangerous accumulations of stress and cognitive decline later in life, severely impacting an individual's well-being [12, 35]. Thus, it is important to study mental fatigue and investigate ways in which it can be detected and alleviated. Given that AR applications have been proposed for a wide variety of industries, this presents the opportunity to examine whether fatigue detection can be built into AR.

3. METHODS

I conducted two user-based studies, both approved by Duke University's Institutional Review Board. The first investigated the use of headset AR eye tracking metrics to detect mental fatigue while the second focused on the impact of environmental lighting conditions on headset AR tracking efficacy. The hypotheses were as follows:

User Fatigue Detection via Headset AR Eye Tracking:

1. Eye tracking data quality from the MagicLeap One headset will be sufficient in tracking fatigue-correlated metrics, as described in (2).
2. Based on results from prior studies, a decrease in pupil diameter and increase in blink duration and frequency is expected as participants become mentally fatigued.
3. Task performance will be negatively impacted by increases in mental fatigue.
4. The activity of watching a video may induce mental fatigue.

Environmental Lighting Impact on Eye Tracking Efficacy:

1. Eye tracking data quality from the MagicLeap One headset will be sufficient in tracking gaze metrics, as described in (2) and (3).
2. Gaze estimation accuracy and precision will be highest/best at moderate levels of ambient light, worst in low light, and intermediate in bright ambient light.

3. Gaze estimation accuracy and precision will be highest at points closer to the center of an individual's field-of-view (FOV). As points move farther away from this center, accuracy and precision will decrease.

3.1 User Fatigue Detection via Headset AR Eye Tracking

3.1.1 Participants and Location

Eight adults from Duke University aged 18-24 (6 male, 2 female) were recruited to participate in this experiment. All user trials were conducted in the Intelligent Interactive Internet of Things (I^3T) Laboratory located in Room 443 of the Wilkinson Building on Duke University's campus.

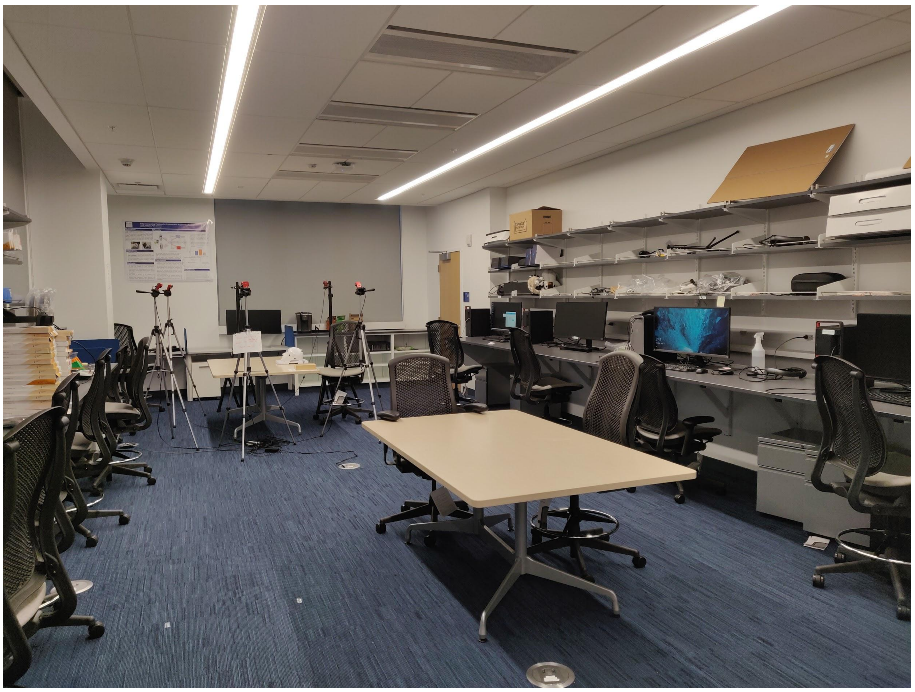

Figure 1. Wilkinson Building Room 443, approximately 6 m x 15 m

3.1.2 Entry Survey

All individuals were asked to confirm their informed consent and complete a short survey prior to their participation. This survey asked participants for sight-related health conditions, familiarity with AR/VR, amount of sleep received the night prior to the experiment, and demographic information. This survey also asked the participant to rate their current level of mental fatigue on a 1 - 10 scale, with 1 being not fatigued at all and 10 being extremely fatigued.

3.1.3 Experimental Setup

Each participant completed a single session of the user study, which consists of seven phases conducted in the following order:

1. Entry questionnaire
2. First AR user task
3. Post-task questionnaire
4. Fatiguing activity
5. Post-activity questionnaire
6. Second AR user task
7. Exit questionnaire

The entry questionnaire (1), post-task questionnaire (3), post-activity questionnaire (5), and exit questionnaire (7) each asked the participant to rate their level of mental fatigue at that given time on a 1 - 10 scale, with 1 being not fatigued at all and 10 being extremely fatigued. The two AR user tasks (2, 6) asked the users to complete an AR-based activity that simulated a warehouse inspection task which may be commonplace in an industrial setting (see section 3.1.2). The fatiguing activity phase asked the user to complete a task designed to mentally fatigue him/her (see section 3.1.3). This experimental setup was designed based on similar procedures from previous studies on eye tracking for mental fatigue detection [12-16].

3.1.2 Augmented Reality Task

For each of the two AR user tasks, participants were asked to perform a simulated warehouse inspection task that was designed to mimic an AR application that may be used in a real-world industrial setting. The AR tasks were completed by participants using a MagicLeap One AR headset. Participants were instructed on how to use the headset prior to each session. A visual calibration was performed prior to each session as well to ensure comfort and eye tracking accuracy. I developed the custom AR application used in this phase with Unity 3D development platform (version 2021.3.0f1) and the MagicLeap Lumin software development kit (version 0.26.0).

Figure 2. MagicLeap One AR headset in use

The AR task is centered around locating ten items (via a six-digit *item ID* number) based on AR inventory lists (*bins*) placed evenly spaced around the room. Each bin is an AR object with text indicating the *bin number* (1 to 10) and which *item IDs* it contains. Each *bin* contained five *item IDs* with no duplicates. Each *item ID* appeared exactly once in each trial. All *item IDs* were generated, partitioned for each *bin*, and selected to be located using a pseudorandom number generator prior to each session.

Prior to each task, the participant was given a clipboard, pen, and sheet of paper printed with a list of ten *item IDs*. The participant was then asked to locate each of the *item IDs* amongst ten the *bins*, and record the *bin number* on the paper. Each participant was instructed to wear the AR headset at all times during the task. A total of five minutes was allotted to complete the task, but users were allowed to finish early. Participants were allowed to walk around the room, but were instructed to look for items in a consecutive fashion according to bin number. Participants were not allowed to “go back” to a certain bin after they have moved on to the next one. The purpose of this was to eliminate variation in data due to differences in task completion strategy and movement around the room.

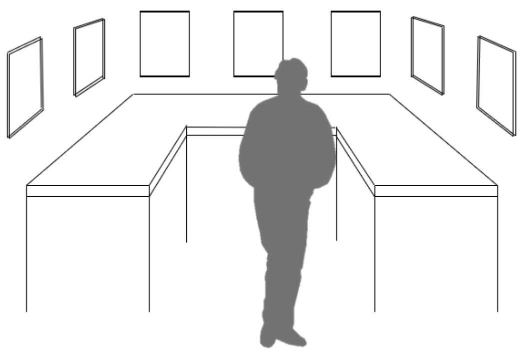

Figure 3. Layout of AR elements in the simulated warehouse inspection task

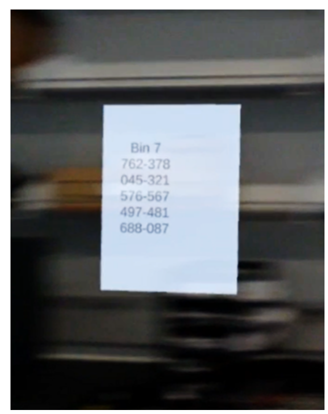

Figure 4. Screen capture of an inventory list from custom AR app

3.1.3 Fatiguing Activity

The fatiguing activity asked users to watch a warehouse safety training video that was 21 minutes in length [30]. This phase of the experiment was designed to mentally fatigue the user. The content of this video was selected for this purpose due to its use of corporate jargon, repetitive nature, and slow pacing. The video was played on a widescreen computer monitor, and participants were not allowed to control the playback of the video in any way. No “progress bar” was visible such that participants could easily determine what proportion of the video had been played at any given time. Participants were instructed to wear the MagicLeap One AR headset during video playback. The purpose of this was twofold: to allow for the collection of eye tracking metrics from the headset during this phase (see section 3.1.4); and to ensure consistency in each participant’s experience throughout all phases of the experiment.

3.1.4 Data Collection

Eye tracking metrics were recorded for the duration of the fatiguing activity and both AR user task phases using the MagicLeap One AR headset. The custom AR application I created recorded blink metrics (time, rate, duration) and pupil diameter (left, right, average) at a 30 Hz sampling rate.

3.1.5 Task Performance

Task performance was measured using two metrics: task completion accuracy and time to complete. Accuracy was determined by the percentage of correct *item ID* to *bin* matches. Time to complete was measured by taking the difference of two manually recorded timestamps at when the AR user task begins and ends.

3.1.6 Data Analysis

Data analysis and visualization were conducted via Python script. Each of the eight sessions had recorded eye tracking metrics for the first AR user task, second AR user task, and fatiguing activity. Data for pupil diameter and blink metrics (times, count, and duration) were considered in 30 second intervals over the duration of each phase. 30 second intervals allow for enough time such that randomness in pupil size and blinking is reduced between samples and is short enough such that trends may be observed over the duration of each phase. Each eye tracking metric was plotted with respect to time. Lastly, changes in fatigue levels and task performance were computed.

3.2 Environmental Lighting Impact on Eye Tracking Efficacy

3.2.1 Participants and Location

Four adults from Duke University aged 18-24 (4 male) were recruited to participate in this experiment. All user trials were conducted in the Intelligent Interactive Internet of Things (I^3T) Laboratory located in Room 443 of the Wilkinson Building on Duke University's campus.

3.2.2 Entry Survey

All individuals were asked to confirm their informed consent and complete a short survey prior to their participation. This survey asked participants for sight-related health conditions, familiarity with AR/VR, and demographic information.

3.2.3 Experimental Setup

Each participant's session was divided into three phases:

1. Pre-session survey, training, and calibration
2. Experimental task and data collection
3. Post-session feedback

During the pre-session phase (1), participants were asked to take the entry survey. Special attention was made to questions regarding the participant's familiarity with AR/VR and any vision or health related issues. Then, participants were instructed on how to wear and use the MagicLeap One AR headset. A visual calibration procedure was run next to ensure participant comfort and accurate data collection. Lastly, the experimental task procedure was explained, with opportunity for the participant to voice questions or concerns.

Next, during the experimental task and data collection phase (2), participants were instructed to look at the center of targets that appeared on an AR-based "screen" (1 meter X 0.667 meters), which is an AR object on which a 2D animation is played (see section 3.2.4). Participants stood at a consistent distance (2.5 meters) away from this AR object and were instructed to refrain from moving their head during the trial. Each trial lasted for the duration of the animation. In addition, participants were given control of when each animation began, by clicking the bumper button on the MagicLeap controller. This procedure was repeated five times for each of three ambient lighting conditions:

Trial Number	**Lighting Condition**
1 to 5	Low (25 Lux)
6 to 10	Medium (30 Lux)
11 to 15	High (72 Lux)

Table 1. Fifteen trials and their corresponding lighting condition

Thus, a total of fifteen trials were conducted with each participant. Between each trial, the animation was reset and the participant was given a short break of two minutes. The lighting condition of the room was only adjusted after trial no. 6, trial no. 10, and trial no. 15.

Overhead lights in the room were adjusted to create the lighting condition specified in Table 1 according to readings from an Adafruit BH1750 ambient light sensor mounted in the position the participant would stand, facing the AR "screen". The light sensor was only present between and after trials when the lighting condition was being adjusted, and it was purposefully removed during each trial to reduce distractions for the participant. Light levels were recorded from the sensor using an Adafruit Feather RP2040 microcontroller board connected to the serial monitor of a Windows PC.

3.2.4 Animation and Target Features

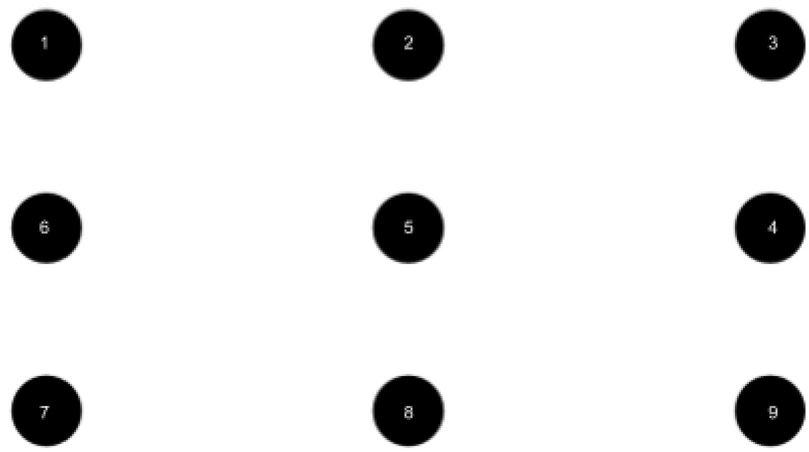

Figure 5. Layout of 3 X 3 grid where targets appear. Order of appearance is labeled in white font, which is for illustration purposes only and not part of the target itself or seen by participants prior

The animation consists of nine black circular targets presented on a white background. Each circular target was of even size (58 pixels X 58 pixels) and distributed evenly in a 3 X 3 grid configuration (see Figure 5). The distance between each target center was 0.364 meters horizontally (X axis) and 0.385 meters vertically (Y axis). The entire animation had a resolution of 640 pixels X 360 pixels, which was scaled down from the original resolution of 1980 pixels X 1080 pixels due to limited rendering capabilities of the MagicLeap One AR headset.

Figure 6. Screen capture of starting screen in custom AR application, before animation begins playing

At the beginning of the animation, all targets are visible for three seconds. Then, targets are presented individually in a consecutive fashion, from 1 to 9 (see Figure 5). Each target remains on-screen for 10 seconds during which valid eye tracking data is collected (see section 3.2.5). To help with transitioning between targets, a 0.5 second fade animation is applied when targets first appear. In addition, the targets first appear as an orange color (R: 255 X G: 160 X B: 0) for two seconds before the color is cross-faded to black. No valid data is collected during this transition consisting of the fade animation and orange color state. This transition was designed to guide the participant's gaze to the next target and allow the participant to make necessary adjustments and "settle" their gaze, thus reducing variation that may arise in the data due to these factors.When the animation ends, the AR "screen" is cut to a dark color, removing it from the MagicLeap One AR headset's field of view.

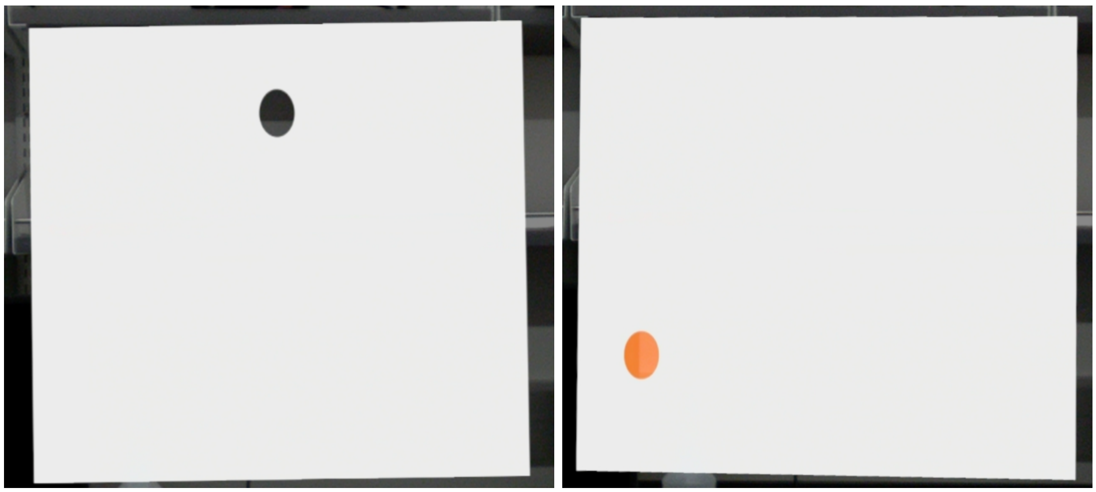

Figure 7. Comparison of two screen captures from custom AR application, with example of black-colored target (left) and example of orange-colored target (right)

3.2.5 Data Collection

Data collection was achieved via the same custom AR application I created that plays the animation on the AR "screen". Gaze metrics (gaze point location and gaze confidence) were collected via MagicLeap's MLEyes application programming interface (API). Pupil diameter and blink metrics (times, count, and duration) were also collected as information that may help inform explanations for variations in gaze. Data was collected for the duration of each trial.

3.2.6 Data Analysis

Data analysis and visualization were conducted via Python script and Microsoft Excel 365 spreadsheet software. Gaze point data was first used to generate plots along the X, Y, and Z axes as well as three dimensionally. Then, accuracy and precision metrics were calculated for each participant's trial and target point. Specifically, standard deviation (precision metric) in gaze location was computed along each of the X, Y, and Z axes for each of the nine target points. In addition, standard distance (accuracy metric) was computed two-dimensionally for the X and Y axes. Standard distance is computed by the following equation, where S_D is standard distance, X_i is a particular data point in the X dimension, $\bar{X}_c$ is the mean value along the X axis, Y_i is a particular data point in the Y dimension, $\bar{Y}_c$ is the mean value along the Y axis, and n is the total number of data points.

$$S_D = \sqrt{\frac{\sum (X_i - \overline{X}_c)^2 + \sum (Y_i - \overline{Y}_c)^2}{n}}$$

Computed precision and accuracy metrics were then aggregated by ambient light levels (low, medium, high) and target point (numbered 1 to 9).

4. RESULTS

4.1. User Fatigue Detection via Headset AR Eye Tracking

4.1.1 Entry Survey Results

From the entry survey, none of the participants required glasses for short distances, had any eyesight related health conditions, or were sensitive to flashing images. Six of eight participants were undergraduate students. One participant indicated no familiarity with AR or VR, while other participants reported varying degrees of familiarity, ranging from <1 to 10 hours of prior usage. For the night prior to the experiment, two participants reported receiving 4 - 6 hours of sleep, four participants reported receiving 6 - 8 hours of sleep, and two participants reported receiving more than 8 hours of sleep.

4.1.2 AR User Tasks

When comparing eye tracking data from the second AR user task with the first AR user task, the results showed a net decrease in mean blink rate (Δ -1.35 blinks per minute) and increase in mean blink duration (Δ 63.16ms per blink). There was also a net decrease in mean pupil diameter (Δ -0.57 mm).

4.1.3 Fatiguing Task

Eye tracking metrics collected during the fatiguing task exhibited trends consistent with the hypothesis, as shown below in Figure 8 and Figure 9. There was a direct negative correlation between time and average pupil diameter over 30 second intervals, while there was a direct positive correlation between time and blink count over 30 second intervals.

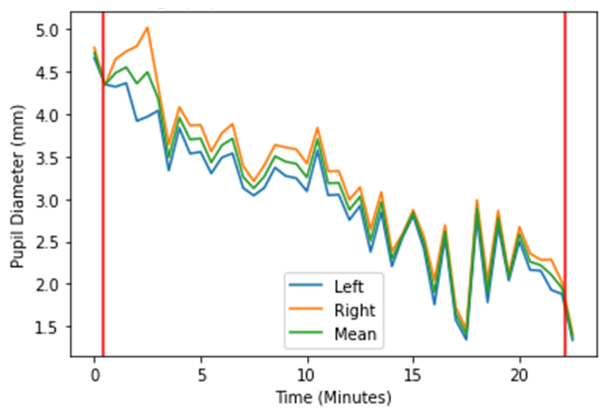

Figure 8. Pupil diameter during fatiguing task

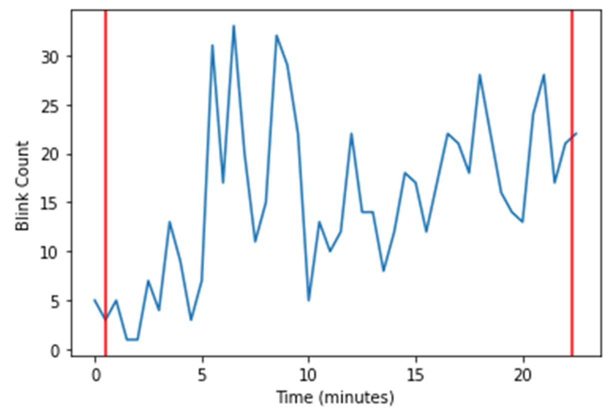

Figure 9. Blink rate during fatiguing task

4.1.4 Reported Fatigue

All participants reported an increase in fatigue (mean difference of 2.5 on a 1-10 scale) as a result of the fatiguing activity. Performing the AR user task reduced reported user fatigue by a mean difference of 0.125 for the first AR user task, and 1.0 for the second AR user task.

4.1.5 Task Performance

2 of 8 participants completed the second AR user task 10% more accurately than the first user task, while there was no difference in accuracy for the other participants. Half of the participants completed the second user task an average of 17.75 seconds faster, while the other half completed the second user task on average 25 seconds slower.

4.2 Environmental Lighting Impact on Eye Tracking Efficacy

4.2.1 Entry Survey Results

From the entry survey, none of the participants required glasses for short distances, had any eyesight related health conditions, or were sensitive to flashing images. All four participants were undergraduate students who had moderate familiarity with AR/VR (approximately 10 hours of prior usage).

4.2.2 Precision

Standard deviation for each target point was compared across low, medium, and high ambient lighting conditions. See Figure 10 for this data visualized by axis. The results showed that the low light condition resulted in the worst precision (standard deviation of X: 0.0912 meters, Y: 0.0713 meters, Z: 0.5215 meters) followed by the high lighting condition (standard deviation of X: 0.0708 meters, Y: 0.0747 meters, Z: 0.3247 meters). The medium lighting condition resulted in the best precision (lowest standard deviation at X: 0.0479 meters, Y: 0.0379 meters, Z: 0.2546 meters). This is better seen in Figure 12. When examined per point and per axis, notable exceptions include Point No. 2, Point No. 7, and Point No. 9 which all had the largest standard deviation in the medium lighting condition for two axes.

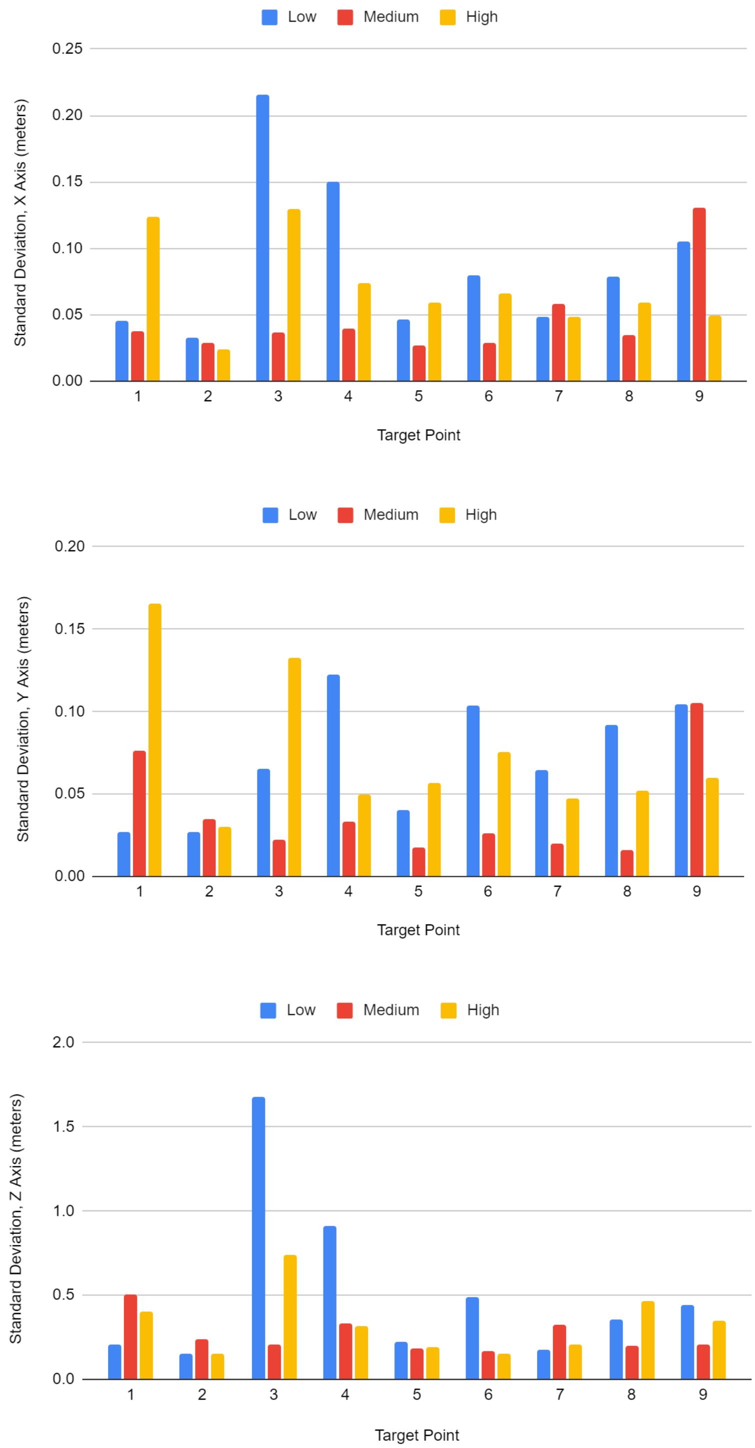

Figure 10. Precision (standard deviation) for each target point, measured across low, medium and high ambient lighting conditions for X (top), Y (middle), and Z (bottom) axes

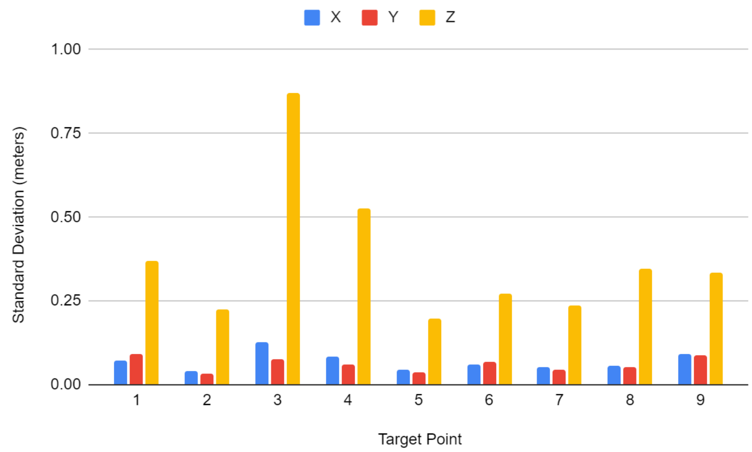

Figure 11. Standard deviation for each point, averaged across low, medium, and high ambient lighting conditions

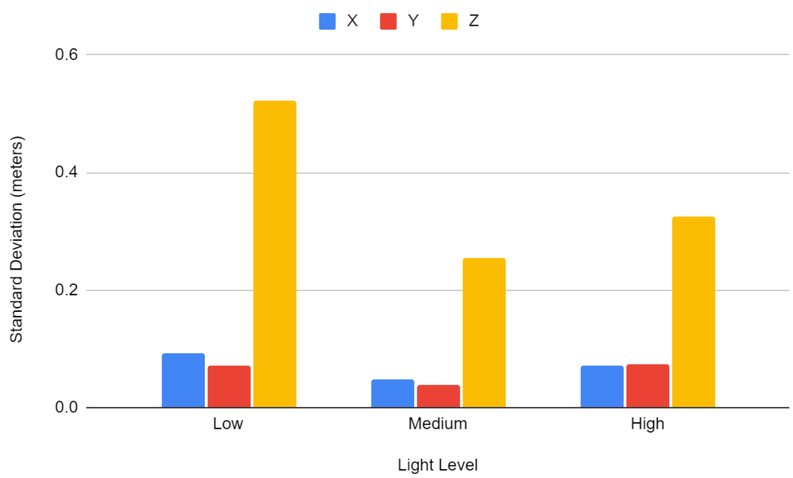

Figure 12. Average standard deviation per light condition, measured across all points

Another trend is that points farther away from the center of the AR screen or the user's FOV tend to have lower precision. This can be observed in Figure 11. Point No. 5 was at the center of the screen and had the highest precision (standard deviation of X: 0.0449 meters, Y: 0.0389 meters, Z: 0.1994 meters). Points No. 1, 3, 7, 9 were farthest from the center and had low precision. In fact, Point No. 3 had the highest standard deviation of all the points. Points No. 2, 4, 6, and 8 had varying degrees of precision in-between the extremes.

4.2.3 Accuracy

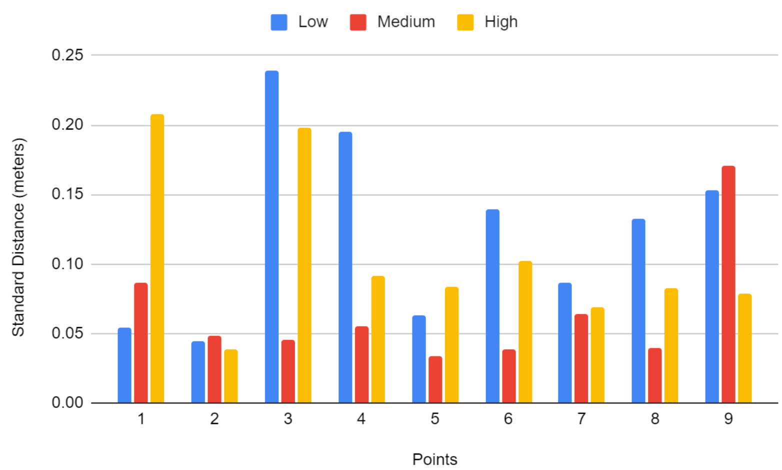

Figure 13. Standard distance for each point and light level

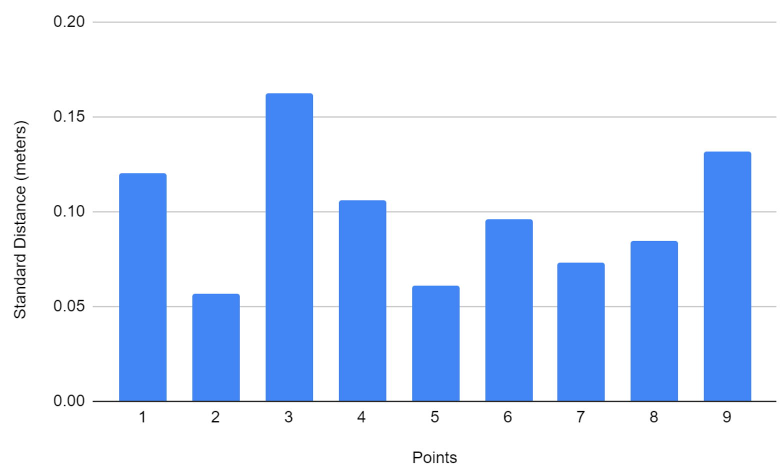

Figure 14. Standard distance averaged across light levels, for each point

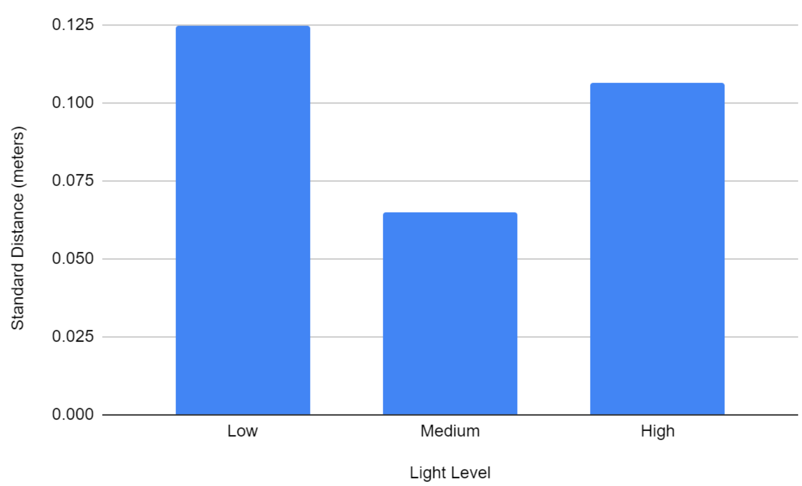

Figure 15. Standard distance averaged across points, for each light level

The results for standard distance (accuracy) are similar to that of precision. When averaged across points for each light level (Figure 15), the results show that accuracy is the worst in low-light conditions (2D standard distance of 0.1249 meters), followed by high light levels (2D standard distance of 0.1065 meters) and medium light levels (2D standard distance of 0.0648 meters). However, unlike that of precision, when this data is separated by point (Figure 13), the impact of light levels on accuracy is less clear. Lastly, the accuracy results for each point, averaged across all light levels (Figure 14) mirror that of precision: points closer to the center of the AR screen tend to have greater accuracy.

5. DISCUSSION

5.1 User Fatigue Detection via Headset AR Eye Tracking

The results from this user study showed that first, the MagicLeap One headset was capable of gathering eye tracking metrics of sufficient quality for the purpose of finding connections between fatigue-correlated measures. Furthermore, there was a decrease in pupil diameter and increase in blink time correlated with increased fatigue, which was consistent with the hypothesis. This study has also shown that the video watching-based mentally fatiguing activity was effective in inducing mental fatigue amongst participants. However, a correlation between task performance and mental fatigue was not shown, primarily due to unforeseen limitations in the design of the experimental task. Instead, this study has also shown that certain activities (i.e. the AR user tasks) may reduce user fatigue by changing the cognitive load placed on the user. This suggests that by varying the nature of an activity on a regular basis, users' mental fatigue may be reduced by keeping the user attentive. This finding is consistent with findings from past studies on attention and fatigue [12 - 16].

5.2 Environmental Lighting Impact on Eye Tracking Efficacy

The results from this second user study confirm the hypotheses. Gaze estimation accuracy and precision was correlated to both lighting condition and target point position. Specifically, accuracy and precision were both highest in medium lighting conditions and performed the worst in low-light. This could be due to the MagicLeap One headset utilizing room light for environment/depth mapping. In low-light conditions, environment mapping may be erroneous, and therefore gaze estimation that is reliant on computed gaze vectors intersecting with a virtual mesh of the surrounding environment will also be inaccurate and imprecise. For bright lighting conditions, the error may arise from the fact that pupils constrict when exposed to more intense light. Thus, pupil center detection may be less accurate and precise, depending on the image captured by the eye tracking camera. Moreover, bright lighting conditions may also introduce infrared noise which may interfere with the eye tracking system. Therefore eye tracking efficacy is highest when lighting conditions are intermediate.

Variation in gaze estimation quality due to target position could be caused by the angles the eye must be in to focus on a target farther from the center of an individual's FOV. The eye tracking cameras may not be able to capture a clear image of the eyes/pupil when it is positioned at these angles, thus reducing measurement quality. These results are consistent with findings from previous studies using standalone eye tracking systems [6, 12].

6. CONCLUSION

The findings of these two user studies show that eye trackers embedded in headset AR devices may be used for user fatigue detection and that eye tracking quality is affected by lighting conditions of the environment as well as target position. This has implications in context-aware AR, specifically:

1. Headset AR applications may want to adapt content to reduce mental fatigue in the user (e.g. simplifying user interface or reducing number of virtual objects). This may be done through the eye tracking system already embedded in AR devices.
2. When user interface (UI) elements rely on gaze estimation for interactivity, changing light levels and target positioning may make it difficult for the user to focus on the intended gaze location. Thus, AR applications may want to adapt UI content (e.g. make targets larger or smaller) or modify the environment (e.g. brighten or dim an internet-enabled light bulb located nearby) to maintain a good user experience by ensuring eye tracking quality.

The main limitation of the studies presented are that they are limited in participants and scope. Nonetheless, they can serve as pilot studies for future research into these areas. In particular, further work is required in developing predictive models to detect user fatigue from eye tracking metrics. In addition, more work is needed to better understand how to optimize UI interfaces or internet-of-things (IoT) systems to adapt to changing eye tracking quality. As AR technology advances and its applications become more ubiquitous, AR features need to work toward promoting human well-being while working reliably under all conditions at all times in order to gain greater appeal. The work presented here may help serve as a starting point to achieving this goal.

ACKNOWLEDGEMENTS:

Special thanks to Tim Scargill and Professor Maria Gorlatova, Ph.D. for mentorship and guidance. This work has been completed as part of two semesters of undergraduate independent study (Fall 2021 and Spring 2022) in the Intelligent Interactive Internet of Things (I^3T) Lab of the Electrical and Computer Engineering Department at Duke University.

REFERENCES:

[1] Z. Wu, S. Rajendran, T. van As, J. Zimmermann, V. Badrinarayanan, and A. Rabinovich, “MagicEyes: A Large Scale Eye Gaze Estimation Dataset for Mixed Reality,” arXiv:2003.08806 [cs, eess], Mar. 2020, Accessed: Apr. 10, 2022. [Online]. Available: http://arxiv.org/abs/2003.08806

[2] Z. Wu, S. Rajendran, T. Van As, V. Badrinarayanan, and A. Rabinovich, “EyeNet: A Multi-Task Deep Network for Off-Axis Eye Gaze Estimation,” in 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), Oct. 2019, pp. 3683–3687. doi: 10.1109/ICCVW.2019.00455.

[3] M. Tonsen, J. Steil, Y. Sugano, and A. Bulling, “InvisibleEye: Mobile Eye Tracking Using Multiple Low-Resolution Cameras and Learning-Based Gaze Estimation,” Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., vol. 1, no. 3, pp. 1–21, Sep. 2017, doi: 10.1145/3130971.

[4] K. A. Dalrymple, M. D. Manner, K. A. Harmelink, E. P. Teska, and J. T. Elison, “An Examination of Recording Accuracy and Precision From Eye Tracking Data From Toddlerhood to Adulthood,” Front. Psychol., vol. 9, p. 803, May 2018, doi: 10.3389/fpsyg.2018.00803.

[5] T. Santini, W. Fuhl, and E. Kasneci, “PuRe: Robust pupil detection for real-time pervasive eye tracking,” Computer Vision and Image Understanding, vol. 170, pp. 40–50, May 2018, doi: 10.1016/j.cviu.2018.02.002.

[6] A. Clemotte, M. Velasco, D. Torricelli, R. Raya, and R. Ceres, “Accuracy and Precision of the Tobii X2-30 Eye-tracking under Non Ideal Conditions:,” in Proceedings of the 2nd International Congress on Neurotechnology, Electronics and Informatics, Rome, Italy, 2014, pp. 111–116. doi: 10.5220/0005094201110116.

[7] M. Thibeault, M. Jesteen, and A. Beitman, “Improved Accuracy Test Method for Mobile Eye Tracking in Usability Scenarios,” Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 63, no. 1, pp. 2226–2230, Nov. 2019, doi: 10.1177/1071181319631083.

[8] G. Lan, B. Heit, T. Scargill, and M. Gorlatova, “GazeGraph: graph-based few-shot cognitive context sensing from human visual behavior,” in Proceedings of the 18th Conference on Embedded Networked Sensor Systems, Virtual Event Japan, Nov. 2020, pp. 422–435. doi: 10.1145/3384419.3430774.

[9] R. Cavin et al., “Eye Tracking for VR and AR .” Facebook Reality Labs. [Online]. Available: https://scontent-iad3-2.xx.fbcdn.net/v/t39.8562-6/240670137_1014428422651912_7953772329359113423_n.pdf?_nc_cat=108&ccb=1-5&_nc_sid=ad8a9d&_nc_ohc=YqA9erLnsJAAX-PPsxB&_nc_ht=scontent-iad3-2.xx&oh=00_AT-Uj0fu4DqPcZIavyKyFRPl-WV7GMblzlvbpfLT5o5N6Q&oe=625B5100

[10] S. Krings, E. Yigitbas, I. Jovanovikj, S. Sauer, and G. Engels, “Development framework for context-aware augmented reality applications,” in Companion Proceedings of the 12th ACM SIGCHI Symposium on Engineering Interactive Computing Systems, New York, NY, USA, Jun. 2020, pp. 1–6. doi: 10.1145/3393672.3398640.

[11] M. Mikhailenko, N. Maksimenko, and M. Kurushkin, “Eye-Tracking in Immersive Virtual Reality for Education: A Review of the Current Progress and Applications,” Frontiers in Education, vol. 7, 2022, Accessed: Apr. 13, 2022. [Online]. Available: https://www.frontiersin.org/article/10.3389/feduc.2022.697032

[12] Y. Yamada and M. Kobayashi, “Detecting mental fatigue from eye-tracking data gathered while watching video: Evaluation in younger and older adults,” Artificial Intelligence in Medicine, vol. 91, pp. 39–48, Sep. 2018, doi: 10.1016/j.artmed.2018.06.005.

[13] M. I. Ahmad, I. Keller, D. A. Robb, and K. S. Lohan, “A framework to estimate cognitive load using physiological data,” Pers Ubiquit Comput, Sep. 2020, doi: 10.1007/s00779-020-01455-7.

[14] S. Benedetto, M. Pedrotti, L. Minin, T. Baccino, A. Re, and R. Montanari, “Driver workload and eye blink duration,” Transportation Research Part F: Traffic Psychology and Behaviour, vol. 14, no. 3, pp. 199–208, May 2011, doi: 10.1016/j.trf.2010.12.001.

[15] R. Schleicher, N. Galley, S. Briest, and L. Galley, “Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired?,” Ergonomics, vol. 51, no. 7, pp. 982–1010, Jul. 2008, doi: 10.1080/00140130701817062.

[16] L.-R. Derick, G.-S. Gabriel, L.-S. Maximo, F.-D. Olivia, C.-S. Noe, and O.-R. Juan, “Study of the User’s Eye Tracking to Analyze the Blinking Behavior While Playing a Video Game to Identify Cognitive Load Levels,” in 2020 IEEE International Autumn Meeting on Power, Electronics and Computing (ROPEC), Ixtapa, Mexico, Nov. 2020, pp. 1–5. doi: 10.1109/ROPEC50909.2020.9258693.

[17] A. D. Souchet, S. Philippe, D. Lourdeaux, and L. Leroy, “Measuring Visual Fatigue and Cognitive Load via Eye Tracking while Learning with Virtual Reality Head-Mounted Displays: A Review,” International Journal of Human–Computer Interaction, vol. 0, no. 0, pp. 1–24, Sep. 2021, doi: 10.1080/10447318.2021.1976509.

[18] J. Grubert, T. Langlotz, S. Zollmann, and H. Regenbrecht, “Towards Pervasive Augmented Reality: Context-Awareness in Augmented Reality,” IEEE Trans. Visual. Comput. Graphics, vol. 23, no. 6, pp. 1706–1724, Jun. 2017, doi: 10.1109/TVCG.2016.2543720.

[19] D. Hallaway, S. Feiner, and T. Höllerer, “BRIDGING THE GAPS: HYBRID TRACKING FOR ADAPTIVE MOBILE AUGMENTED REALITY,” Applied Artificial Intelligence, vol. 18, no. 6, pp. 477–500, Jul. 2004, doi: 10.1080/08839510490462768.

[20] T. Tahara, T. Seno, G. Narita, and T. Ishikawa, “Retargetable AR: Context-aware Augmented Reality in Indoor Scenes based on 3D Scene Graph,” in 2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Recife, Brazil, Nov. 2020, pp. 249–255. doi: 10.1109/ISMAR-Adjunct51615.2020.00072.

[21] “Developer Portal | Magic Leap.” https://developer.magicleap.com/en-us/learn/guides/design-eye-gaze (accessed Apr. 13, 2022).

[22] “Pupil Core - Open source eye tracking platform - Pupil Labs.” https://pupil-labs.com/products/core/ (accessed Apr. 13, 2022).

[23] L. J. Bour, D. van’t Ent, and J. Brans, “Peak Saccade Velocities, Short Latency Saccades and their Recovery after Therapy in a Patient with a Pineal Tumor,” in Studies in Visual Information Processing, vol. 6, J. M. Findlay, R. Walker, and R. W. Kentridge, Eds. North-Holland, 1995, pp. 291–302. doi: 10.1016/S0926-907X(05)80025-X.

[24] “Scleral Search Coil Technique - an overview | ScienceDirect Topics.” https://www.sciencedirect.com/topics/medicine-and-dentistry/scleral-search-coil-technique#:~:text=The%20scleral%20search%20coil%20method,32%2D7). (accessed Apr. 13, 2022).

[25] “Eye tracking overview - Mixed Reality.” https://docs.microsoft.com/en-us/windows/mixed-reality/design/eye-tracking (accessed Apr. 13, 2022).

[26] J. Calderone, “Eye Tracking in Google Glass: A Window into the Soul?,” Scientific American. https://www.scientificamerican.com/article/eye-tracking-in-google-glass-a-window-into-the-soul/ (accessed Apr. 13, 2022).

[27] M. Mansouryar, J. Steil, Y. Sugano, and A. Bulling, “3D Gaze Estimation from 2D Pupil Positions on Monocular Head-Mounted Eye Trackers,” Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications, pp. 197–200, Mar. 2016, doi: 10.1145/2857491.2857530.

[28] “Learn about the different types of eye movement - Tobii Pro,” Aug. 06, 2015. https://www.tobiipro.com/learn-and-support/learn/eye-tracking-essentials/types-of-eye-movements/ (accessed Apr. 13, 2022).

[29] D. Purves et al., “Types of Eye Movements and Their Functions,” Neuroscience. 2nd edition, 2001, Accessed: Apr. 13, 2022. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK10991/

[30] Warehouse Safety Video. Accessed: Apr. 13, 2022. [Online]. Available: https://www.youtube.com/watch?v=xjiBeKLA6t8

[31] D. Dawson, A. K. Searle, and J. L. Paterson, “Look before you (s)leep: Evaluating the use of fatigue detection technologies within a fatigue risk management system for the road transport industry,” Sleep Medicine Reviews, vol. 18, no. 2, pp. 141–152, Apr. 2014, doi: 10.1016/j.smrv.2013.03.003.

[32] C. Ahlstrom et al., “Fit-for-duty test for estimation of drivers’ sleepiness level: Eye movements improve the sleep/wake predictor,” Transportation Research Part C: Emerging Technologies, vol. 26, pp. 20–32, Jan. 2013, doi: 10.1016/j.trc.2012.07.008.

[33] “The dependence of pupil size upon external light stimulus under static and variable conditions,” Proc. R. Soc. Lond. B, vol. 121, no. 823, pp. 376–395, Dec. 1936, doi: 10.1098/rspb.1936.0072.

[34] V. Shahly et al., “The Associations of Insomnia With Costly Workplace Accidents and Errors: Results From the America Insomnia Survey,” Arch Gen Psychiatry, vol. 69, no. 10, p. 1054, Oct. 2012, doi: 10.1001/archgenpsychiatry.2011.2188.

[35] S. E. Hardy and S. A. Studenski, “Fatigue and Function Over 3 Years Among Older Adults,” The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, vol. 63, no. 12, pp. 1389–1392, Dec. 2008, doi: 10.1093/gerona/63.12.1389.

[36] D. Lindlbauer, A. M. Feit, and O. Hilliges, “Context-Aware Online Adaptation of Mixed Reality Interfaces,” in Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology, New Orleans LA USA, Oct. 2019, pp. 147–160. doi: 10.1145/3332165.3347945.