5-1-2020

Impact of context switching and focal distance switching on human performance in all augmented reality system

Mohammed Safayet Arefin

Impact of context switching and focal distance switching

on human performance in all augmented reality system

By

Mohammed Safayet Arefin

A Thesis
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in Computer Science
in the Department of Computer Science and Engineering

Mississippi State, Mississippi

May 2020

Impact of context switching and focal distance switching

on human performance in all augmented reality system

By

Mohammed Safayet Arefin

Approved:

J. Edward Swan II
(Major Professor)

Shahram Rahimi
(Committee Member)

Cindy Bethel
(Committee Member)

T. J. Jankun-Kelly
(Graduate Coordinator)

Jason M. Keith
Dean
Bagley College of Engineering

Name: Mohammed Safayet Arefin

Date of Degree: May 1, 2020

Institution: Mississippi State University

Major Field: Computer Science

Major Professor: J. Edward Swan II

Title of Study: Impact of context switching and focal distance switching on human performance in all augmented reality system

Pages of Study: 79

Candidate for Degree of Master of Science

Most current augmented reality (AR) displays present content at a fixed focal demand. At the same time, real-world stimuli can occur at a variety of focal distances. To integrate information, users need to switch eye focus between virtual and real-world information continuously. Previously, Gabbard, Mehra, and Swan (2018) examined these issues, using a text-based visual search task on a monocular AR display. This thesis replicated and extended the previous experiment by including a new experimental variable *stereopsis* (stereo, mono) and fully crossing the variables of context switching and focal distance switching, using AR haploscope. The results from the monocular condition indicate successful replication, which is consistent with the hypothesis that the findings are a general property of AR. The outcome of the stereo condition supports the same adverse effects of context switching and focal distance switching. Further, participants have better performance and less eye fatigue in the stereo condition compared to the monocular condition.

Key words: augmented reality, context switching, focal distance switching, accommodation, fatigue

DEDICATION

Dedicated to:

My parents: Mohammed Abdus Sobhan & Jahanara Begum,

My wife: Farzana Alam Khan,

My younger brother: Mohammed Safayet Jamil

ACKNOWLEDGEMENTS

First, I would like to thank and express my gratitude to my adviser Dr. J. Edward Swan II for his encouragement, direction, and guidance towards my research. Before joining the Spatial Perception and Augmented Reality (SPAAR) lab at Mississippi State University, I did not have any prior knowledge about augmented reality research. Dr. Ed is the first person who introduced me to this fantastic research area and also acquainted me with several well-known researchers in the field of AR. He left no stone unturned to inspire and motivate me in every single thing I intend to in my research. Besides, his guidance on my program of study directed me to take courses from other departments too (e.g., physics and psychology), which helped me to enrich my cross-disciplinary knowledge. Finally, I strongly hope and believe that the unconditional inspiration and guidance of Dr. Ed will lead me to the completion of my Ph.D., as well as the development of my career.

I want to give a very special thanks to Dr. Alexander Plopski, who helped and guided me thoroughly during the experimental development of my research. Throughout my research, I did not hesitate to ask him any research questions at any time. No matter how simple it was, he never disappointed me. I still remember the days when we talked about our research at midnight or very early morning due to different time zones.

I would also like to thank Dr. Joe Gabbard for his support and guidelines for my research. In this thesis, I replicated his previous work using a different AR display. All the feedback from Dr. Gabbard helped me a lot conduct my research more appropriately and accurately.

I also want to thank Nate Phillips for his support in my research. He is one of the first persons who showed me the calibration procedure of AR haploscope along with Dr. Ed. We spent many days working with the AR haploscope till midnight. In addition, I want to thank all the members of the Spatial Perception and Augmented Reality (SPAAR) lab of Mississippi State University for their many exciting and enlightening discussions as well as suggestions.

I acknowledge the National Science Foundation (NSF) for their support on this research. This research is based upon work supported by the National Science Foundation, under awards IIS-1937565 and IIS-1320909, to J.E. Swan II.

I am grateful to my beautiful family, especially to my parents and my younger brother, for their endless support from halfway across the globe, in Bangladesh. From my elementary school to my Ph.D. program, I have received endless love, support, and prayers from my parents, and I have no words to explain how blessed I am to count them in my everyday life.

I acknowledge my amazing wife, Farzana, for her unconditional love and support. We got married in the middle of my Ph.D. program. I think the initial period of marital life is challenging for a graduate student while he is in the middle of the graduate program. But, the support and inspirations I got from Farzana helped me a lot to overcome that initial period. More specifically, she did not make me feel any pressure and distraction in this period. I am grateful to her.

Finally, and most importantly, I acknowledge the Almighty for showering positive vibes and energy in me to complete my Master's thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS, ABBREVIATIONS, AND NOMENCLATURE

AR: Augmented Reality

HMD: Head-Mounted Display

OST: Optical See-Through

VST: Video See-Through

CHAPTER 1

INTRODUCTION

Augmented Reality (AR) is a modern technology that superimposes computer-generated graphics on the view of the real world to enhance users' real-world view [3]. The current scenario of AR is built on its prolonged history, which started when Ivan Sutherland created the first three-dimensional display in 1968, called "The Sword of Damocles," to display an image that moves with the movement of users [66]. In 1975, Myron Krueger established an artificial laboratory called "videoplace," where he created the first virtual reality interface which responded to the users in real time [30]. Many research groups around the world from academic laboratories and industries continued their research during the 1970s and 1980s to put their step on the next level of three-dimensional display technology. However, in 1990, Boeing researcher Thomas P. Caudell coined the term "augmented reality," for the first time in history, after developing the system that combined head position sensing and real-world registration system to superimpose computer-generated graphics in front of the users view [7]. The rapid transition and development of AR happened at the beginning of the 20th century. Kato et al. [34] introduced an open-source software library named ARToolKit in 2000 for helping researchers to develop fiducial marker-based AR applications. The innovation of Google Glass [64] brought a revolutionary change in the AR field in 2014. Moreover, in 2016, a more advanced version of an AR display was unveiled by Microsoft, named the Microsoft HoloLens [38], which added more dimensions in the progress and development of AR. Due to the dramatic progress of research [12] and the advancement of commercial AR devices, AR is rapidly progressing to provide an unprecedented user experience in various applications as diverse as manufacturing, repair, military, healthcare, education, entertainment, navigation, and others [69, 60].

According to Azuma [3], an AR system has three characteristics: "(1) Combines real and virtual (2) Interactive in real-time (3) Registered in 3D." This definition implies that in an AR system, the virtual contents are rendered and displayed in the real world with a 3D coordinate system, and the user can interact with the AR contents in real-time. That means, in an AR system, information is distributed between real-world and graphical contents, which often appear at different distances from the user. As a result, users are required to perform rapidly transition between fixating on the graphical content presented through the AR display and fixating on the real-world content. Additionally, to integrate information, users need to shift their eye focus from one particular focal distance of virtual content to another focal distance of real-world content and vice versa.

Unfortunately, most current AR displays have a fixed focal distance (e.g., Microsoft HoloLens, Google Glass) to place virtual information, whereas real-world objects have a range of different focal distances. Continuously shifting eye focus between different contexts and distances has been shown to cause significant differences in task performance, reduce comfort, increase fatigue, and eye strain [19, 52, 40]. It is strongly believed that these effects will also occur in AR, but this hypothesis needs to be verified.

Figure 1.1: Example of the switching problem in AR, frames taken from a 2012 Google Glass concept video on YouTube [1]. To the left, the AR text is in focus, while the background text is out of focus. To the right, there are two frames, which show the focus changing from the AR symbology (upper right) to the building (lower right). Throughout this video, only the AR symbology or the background is in focus at any one time, and the focus constantly switches between them.

1.1 Problem statement

When using OST AR displays, such as Google Glass, Microsoft HoloLens, or Magic Leap One, interacting with the virtual content requires the observer's eyes to focus on the optical depth of the display. However, sometimes, the user's task requires the eyes to focus on real-world content, which may be located at a different focal distance. Consider a real-life scenario (Figure 1.1), taken from a Google Glass concept video on YouTube. In this video, the user is walking around an urban setting, and as shown in Figure 1.1, only the AR symbology or the background is in focus at any one time. These two are never in focus at the same time, and throughout the video, the focus continually switches between AR symbology and the background.

If the user's task requires them to integrate information between the real world and virtual content, they must repeatedly switch context and refocus the eyes. Here, *context switching* refers to switching the visual and cognitive attention between the real world and virtual information [19]. On the other side, *focal distance switching* refers to accommodating (changing the shape of the eye's lens) to see, in sharp focus, information at a new distance [19]. Further, both context switching and focal distance switching have a strong correlation with two of the dominant components of the human visual system, are accommodation and vergence. Any mismatch between accommodation and vergence eventually reduces user performance, increases cognitive load, and creates eye strain among the users [46, 16, 27, 28, 44]. In addition, changing accommodation takes time. Up to age 20, the human eye requires 360 milliseconds to accommodate from far to near and 380 milliseconds to accommodate near far [65, 19]. Further, after the age of 20, time requires to accommodate near to far is constant, but time for accommodating far to near has increased [33]. This changing accommodating time implies that changing accommodating can impact users' task performance and accuracy.

1.2 Motivation

As AR technology has been rapidly growing in many sectors, it is essential to consider the current limitations and issues of the AR system from the human perspective. The motivation of this thesis is to utilize the power of AR entirely by considering all the factors and variables which have significant effects on the difficulty and efficacy of context switching and focal distance switching. Without considering each of these, it is difficult to come to a complete understanding of context switching and focal distance switching. Further, it is impossible to say whether there is an inherent cost that persists across environments and applications. To the best of my knowledge, to date, only Gabbard et al. [19] have explored the impacts of context switching and focal distance switching in AR. However, this research has certain limitations that might distort or bias the results. This previous research did not consider the effects of certain perceptual variables, including angular size and stereopsis, which might affect experimental results. Further, it is possible that instead of being a general AR phenomenon, their findings are specific to their used AR display in the experiment. For this reason, I am motivated to explore these two crucial AR interface issues to determine the separate effects of context switching entirely and focal distance switching on user task performance, fatigue, and cognition.

1.3 Research tasks

The research tasks of this thesis can be divided into three parts:

- First, this work successfully replicates the experiment conducted by Gabbard et al. [19] using a text-based visual search task that integrates the information both from the real world and AR world on AR haploscope. The outcome of this section broadly generalizes the impact of context switching and focal distance switching issues in AR user interface design.

- After successful replication of the previous work, this research extends the experiment conducted by Gabbard et al. [19], including the variable *stereopsis,* (stereo, mono) and fully crossing the variables of context switching and focal distance switching. This section provides novel empirical findings that illustrate the existence of similar negative effects of context switching and focal distance switching on human performance in the stereo AR system.

- After successful replication and extension, this thesis demonstrates the empirical comparison of monocular and stereo AR systems in terms of the impact of context switching and focal distance switching on human performance. This section demonstrates how much human performance and fatigue varied between the monocular and stereo AR system.

CHAPTER 2

RELATED WORK

2.1 AR and applications of AR

When a person observes the real world with his own eyes, he does not see anything extra or special except the real objects. However, if we add graphical content to augment his view of the real world, then it would not be an ordinary reality; it would be augmented reality(AR). AR involves the amalgamation of real and virtual objects by superimposing graphical objects in the real environment, which increases the visual knowledge of an observer. On the other hand, if real-world objects do not exist in the person's view, but he can see and interact with graphical content in the immersive world, then reality is known as virtual reality (VR). Paul Milgram [45] showed the "virtuality continuum" where the real environment was shown in one end, and the virtual environment was shown at another end of the continuum. From the left side of the continuum, the real environment indicates viewing the real-world objects that can be viewed directly without any display device or head-mounted displays (HMDs). On the right side of the continuum, there is a virtual environment that only contains the graphical contents in the virtual environment.

AR has numerous fields where development and technological progress has already been applied. Researchers and developers give their best efforts to spread AR technolo-

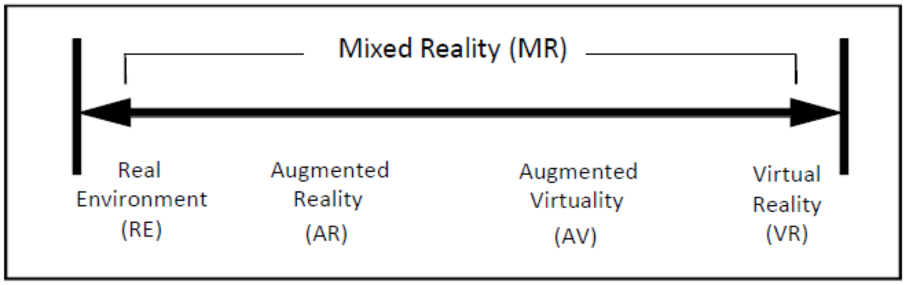


Figure 2.1: "virtuality continuum" by Paul Milgram [45]

gies in different sectors. Medical, military, industries and forms of entertainment such as gaming are the various sectors where AR is applied. Some are discussed below:

Medical sector AR has put tremendous efforts in the medical sector to help the doctors to visualize the patients in more detailed ways and adequately prepare for surgeries. According to Azuma [3], it is possible to render and combine 3D datasets of a patient in real-time through non-invasive sensors such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT) scans, and ultrasound imaging with a view of the patient's body. Moreover, this leads to the concept of "X-ray vision," which indicates seeing through the non-transparent objects, here mentioned as patients. This technology might be helpful for surgical operations such as Laparoscopy, where surgeons limit the size and number of cuts or incisions that need to make. These types of surgeries are known as minimally invasive surgery, and AR technology can help the surgeons by providing an internal view without the need for larger incisions [3]. Besides, MRI scans using AR HMD for delivering the aspects of tool manipulation hidden beneath the tissue [69] help surgeons find a practical approach to do the treatment. For critical surgeries like brain surgery or biopsy, doctors need to continuously observe the patient's condition and results from CT scans or MRIs. AR makes it easier so the doctors can access both types of context simultaneously and perform the operations safely. Further, researchers found a way to visualize the 3-D representation of the fetus inside the womb through AR display by conducting several trials of scanning a pregnant woman's womb with an ultrasound sensor [3].

Navigation AR in indoor and outdoor navigation systems has been tested and applied in the real world. Navigation apps within an AR device provide additional information while hovering over the real-world objects. Using fiducial markers for position tracking with a hand-held camera can be an example of navigation for indoor use [69]. Navigation in the outdoor provides meaningful information such as a precise location, the best route for a destination in real-time, highway exits, fuel prices, and so on [51, 61].

Manufacturing & Maintenance Integrating AR in multiple industries and manufacturing companies has added a new dimension. AR applications make the production process easier and faster by providing instructions efficiently for the workers as well as manufacturers. For example, it would be easier to get the guidance of actual equipment in 3D drawings rather than manuals with texts and pictures [3]. In the automotive section, Doshi et al. [13] used a projector-based AR system to improve the precision and accuracy of the manual spot welding task. In addition, several branded car companies such as BMW experimented with AR to enhance the welding process of their cars. Besides, Volkswagen used AR in construction to analyze interfering edges, plan production lines, and workshops [15, 69]. Additionally, Echtler and Klinker [15] proposed an intelligent welding gun system. This display helps the operator by showing three-dimensional stud location on the car frame relative to the current welding gun position (see figure 2.2).

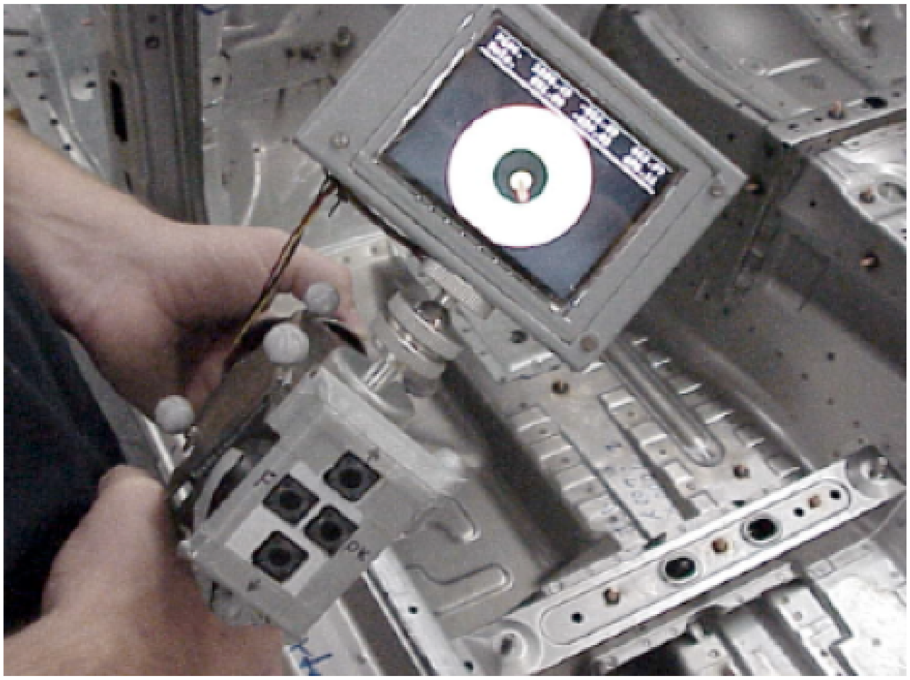


Figure 2.2: "Intelligent Welding Gun" proposed by Echtler and Klinker [15]

AR is equally useful for maintenance purposes. AR system can be used for the inspection of power plants. Besides, AR intends to support the electrical troubleshooting of vehicles. Reported in the survey paper of Krevelen et al. [69], some vehicle companies such as Honda and Volvo are using AR technology to help the technicians with vehicle history and repair information.

Education One of the significant areas of AR application is education. In the education sector, AR could be an effective system for increasing the learning capabilities among children. Incorporating AR content in the classroom can make the learning process more efficient and enjoyable. Billinghurst et al. [5] investigated the impact of AR application in the learning system in both elementary and high school classrooms. Both their research results and classroom reports supported the idea of using AR as a teaching tool. Therefore, AR technology gets the attention of the students in the classroom and makes learning easier [5, 9]. Visual learners can incorporate theoretical knowledge into real scenarios that are easier to grasp.

Entertainment AR displays provide options for entertaining people around the world; more specifically in the gaming sector. Pokémon Go is referred to as one of the first games in AR. Since then, AR has taken the game development to a whole new level. Now, numerous types of games are using AR technology, such as angry birds, puzzle games, etc. People do not have to depend on TV or computers for watching movies or hearing songs; they can get that in the real environment using AR displays. One can decorate the surrounding AR environment using different holograms.

2.2 AR displays

In this section, different characteristics and designs of AR displays will be discussed. The AR display system can be divided into many categories such as optical see through(OST) vs. video see through(VST), monocular vs. stereo, etc.

Optical see through(OST) vs. video see through(VST) In OST HMDs, optical combiners are placed in front of the user's eyes. These optical combiners are not opaque but semi-transparent and allow the users to view the real world [50, 49]. The optical combiners are also partially reflective, which enables the users to see virtual objects superimposed on the real world. A conceptual diagram of an OST is given in figure 2.3.

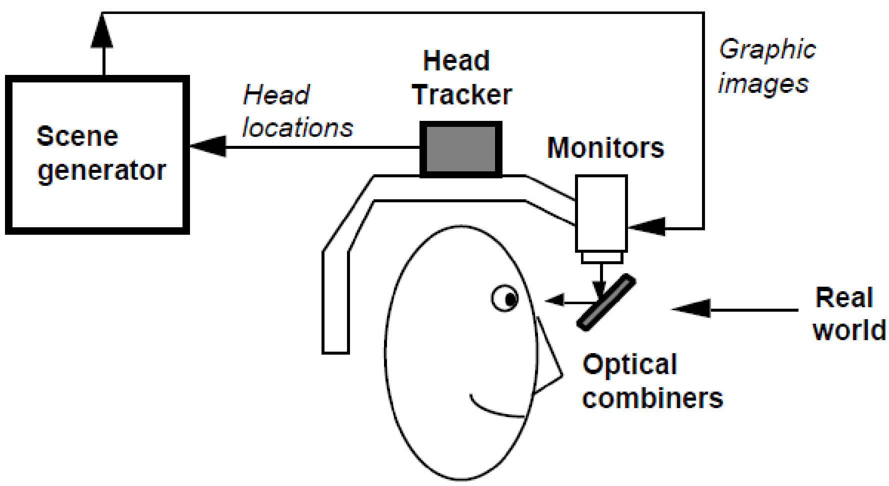


Figure 2.3: A conceptual diagram of an optical see-through display [3]

On the other side, in VST display, one or two head-mounted video cameras are mounted on the HMDs to capture the view of the real world. A monitor is placed in front of the user's eye to feed the video of the real world combined with the AR elements [50]. In this case, users' eyes can not see the view of the real world directly. A conceptual figure of video see-through HMD is given in figure 2.4.

Monocular vs. stereo AR HMDs can be categorized into monocular and binocular, also known as stereo vision. In the monocular display, the image from the real world is shown to one eye only while the other eye is deprived of seeing the view. It requires one single image source and one set of optics for viewing through only one eye [23]. Sometimes, based on

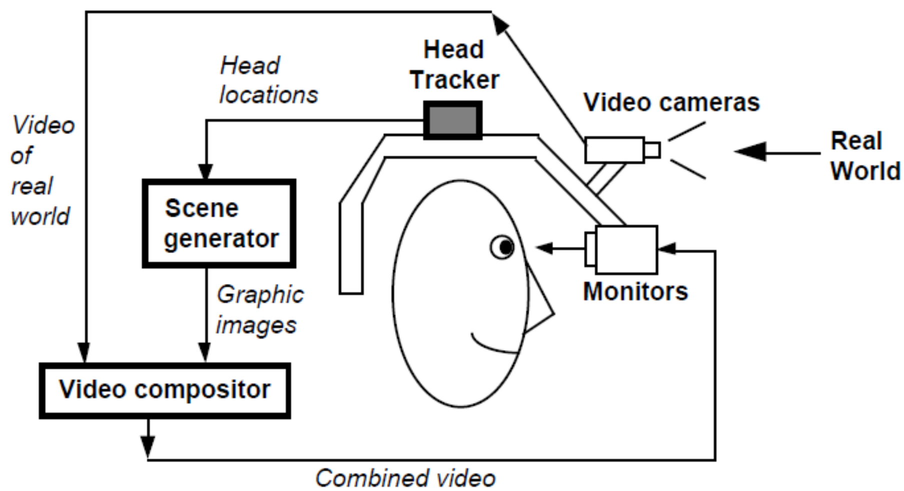


Figure 2.4: A conceptual diagram of an video see-through display [3]

the application, people can observe the real world with both eyes where monocular display optic is positioned a little top of the left or right eye. This kind of monocular display (e.g., Google Glass Enterprise Edition 2) is used to provide extra information to accelerate the users' task performance and speed in many fields, including health care, manufacturing, etc. On the other hand, in stereo condition, both left and right eyes are allowed to observe the real world as well as AR elements. Both eyes require individual sets of optics to see two separate sets of images. Therefore, the configuration and methodology of the monocular display are quite more straightforward than binocular. The monocular display is also lightweight and easy to use. In both displays, users feel some level of comforts and discomforts because the configuration and operation procedure of the conditions are different from one another.

2.3 Context switching and focal distance switching

Although switching of context and focal distance in AR is frequent, very few research considered human performance impact, in our concern, only three published papers considered the context switching and focal distance switching in AR. Huckauf et al. [28] first examined the AR context switching through a switching task. They found that switching between devices cost in visual performance and reduced the user performance up to 10% in their conducted experiment. Gabbard et al. [19] have experimentally considered the impact of both context switching and focal distance switching on human performance in an AR system, using a text-based visual search task with a monocular display against a black background. They found that user performance was better at near and medium distances than at far distances. They further found improved task completion and accuracy when participants did not have to re-focus on another distance.

In 2019, Eiberger et al. [17] evaluated human performance by conducting a visual search task on a joint OST HMD-body proximate display systems at 30 cm. In this system, OST HMD is combined with a smartphone or smartwatch. In their experiment, they displayed information at a uniform depth layer by presenting a smartwatch through an OST HMD in one condition. In this condition, participants did not need to switch focus during the task. In another situation, they considered two depth layers (smartwatch and focal distance of the head-mounted display), and participants had to switch focus to integrate information from both displays. The results of their research showed that participants had higher task completion time with a high error rate in two depth layers condition compared to the uniform depth layer condition.

2.4 Text based AR

In the AR system, various types of graphical contents can be overlaid in a real-world environment. Text is one of the most primary graphical contents in the AR system. Short AR Textual information has been widely used in many AR applications not constrained in maintenance, education, navigation and driving, and others. In general, textual information adjoins with photos, illustrations, flyers, and other sort of outdoor background content. In addition, human cognition and perception have a direct influence on the text readability [11]. However, text readability on the computer-based displays is different from the AR displays as text representation on AR displays related to the display technologies, text style, text color, and background. In the last few years, extensive research has been conducted on the text drawing style, background, text readability, text color, and text-based user studies in the AR research field.

In AR systems, one of the major challenges is putting the textual information in the real world outdoor environment. Due to uncontrollable conditions such as lighting, objects, and other factors, it becomes challenging to put textual information precisely and clearly on the real-world natural background. Cho et al. [8] showed that natural image properties do not follow the specific properties of the text-based images. They also presented the log-scale gradient histogram comparison of the natural image and text-based image. The comparison clearly showed that natural image is significantly different from a text image (see figure 2.5). For this reason, they provided three properties of text-based images. In essence, textual characters have high contrast against the background, each

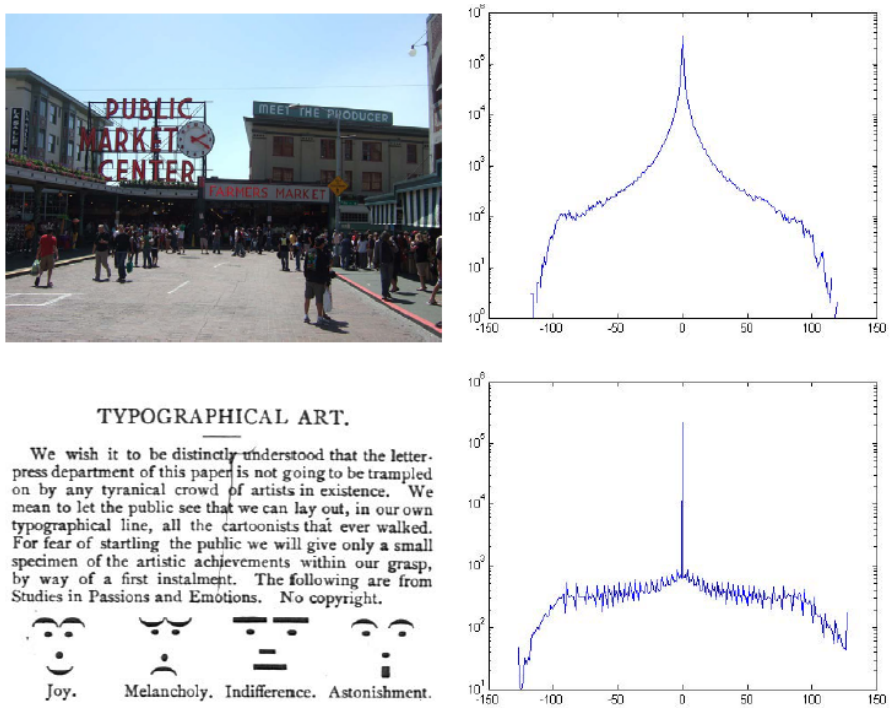


Figure 2.5: Log-scale gradient histogram comparison of the natural image and text based image [8]. Upper row shows the natural images with the corresponding gradient histogram and bottom row shows the text based image with its gradient histogram.

character has a near-uniform color and text background should follow the characteristics of natural image [8].

Text style and color on the background is an essential factor considering the visibility and legibility of texts in the AR system. Sometimes, due to the impact of brightness and color of the background, the virtual text becomes invisible. In order to explore this issue, Gabbard et al. [21] conducted an empirical user-based study using an OST HMD to find out the effects of outdoor background textures, lighting and text drawing styles on user performance while performing a text identification task. In their experimental setup, they considered six outdoor background textures, six text drawing styles, and three distances (near, medium, and far). In their research, they found a piece of clear empirical evidence that user task performance in an AR system is affected by the background texture, text drawing styles, and their interaction. Results from their study also suggested that user-preferred and performed better in the billboard and green text drawing styles. They mentioned that participants text-based reading task could be affected by the ambient illumination.

In 2007, an extension of the previous work, Gabbard et al. [20] conducted another empirical study to explore the effects of text drawing styles for outdoor AR by employing a visual search task. In the study, they considered four real-world outdoor background texture (brick, building, sidewalk, sky), four text colors, three text drawing styles, and two text drawing style algorithms as experimental variables. Their experimental findings showed that the maximum brightness and contrast algorithm performed better than any other algorithm. Besides, participants performed most accurately on the building background and made the most error on the brick background.

Later in 2010, Jankowski et al. [31] investigated the effects of text drawing styles, image polarity, and background styles on text readability. They considered reading speed and accuracy to evaluate the users' performance and recorded the subjective evaluation of the participants. In their study, they used four text drawing styles, two kinds of image polarity, and two backgrounds. Their results showed that participants performed best in the billboard text drawing styles. Based on their findings, they suggested using a billboard text drawing style to maximize readability for many text-related applications.

In 2014, Debernardis et al. [11] explored readability on two head-worn devices (optical and video see-through), two backgrounds (light and dark), five colors (white, black, red, green, and blue), and two text styles (plain text and billboarded text). They conducted a text-based visual search task, similar to the Gabbard et al. [19] to measure human performance on text readability in AR systems. Their research revealed that participants preferred the OST device on the dark background and VST device on the light background. Based on their research and findings, they suggested that white text color with any mandatory background color is suitable for increasing the user performance in the AR system.

2.5 Accommodation and vergence

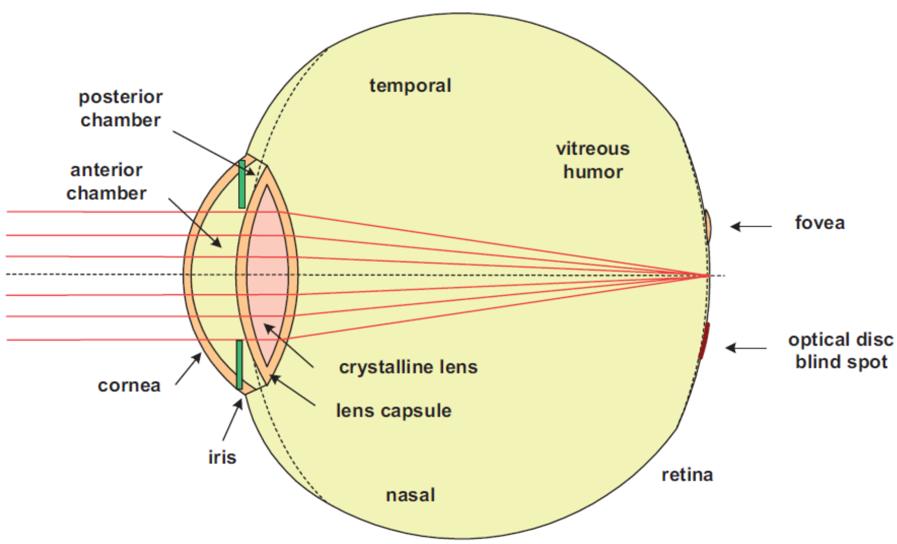


Figure 2.6: Geometry of the human eye [22]

The eye is one of the involved organs of the human body with a size of approximately 24mm [22], and the only medium to visually observe and perceive the surroundings around us. The primary optical function of the human eye is to form an image on the retina (retinal image) while observing an object at a specific range of distances. The retina consists of the fovea and blind spot. The fovea is a tiny part of a retina where humans observe the most precise vision, and resolution is maximized; on the other hand, no light is detected in the blind spot [22]. Light from an image enters into the eye through the cornea, passes through the lens which behaves like a convex lens, and forms an image on the retina (figure 2.6). The human eye can focus at a particular object within a certain range. However, the distance between objects and the eye is regularly changing. The focus length of the eye lens is also adjustable by the ciliary muscle based on the object distance so that the image is always formed on the retina in a normal situation. The ability of the eye to adjust its focal length is known as *accommodation* [70]. A person with normal vision can see objects clearly at distances ranging from 25 cm to essentially infinity [22]. Eye accommodation is necessary for focusing on an object both in the monocular and stereo vision. Along with accommodation, eye movements are required to focus an object in the binocular vision as the axes of both eyes are not parallel while viewing an object with two eyes [22]. In theory, three types of eye-movements are found [22] (see figure 2.7): 1. If the axes of the two eyes rotate inwards horizontally for a common object point, then it is called *convergence*, which is needed for binocular vision. 2. If two eyes axes are rotated outwards horizontally, then it is called *divergence*. 3. If the two eyes moved in the vertical direction with a reversed sign, then it is called *dipvergence*.

While viewing an object at any distance human visual system requires a short amount of time (0.2-0.6 seconds) to fuse the image with an appropriate vergence angle. Vergence

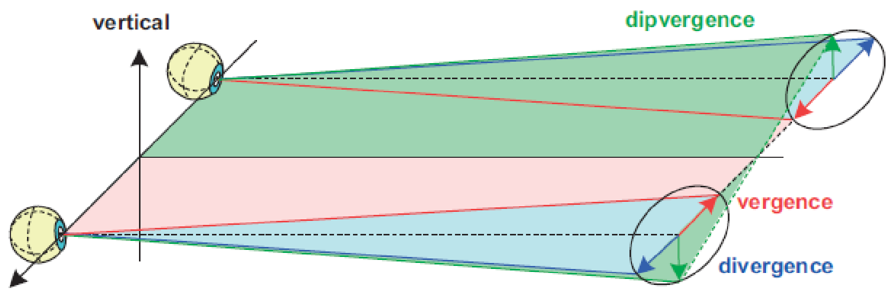


Figure 2.7: Three differenct cases of axes orientation of human eye [22]. *red* color shows the *convergence* of the eye, *blue* color shows the *divergence*, and *green* color denotes the *dipvergence* in this figure

angle is related to the accommodation demand and interpupillary distance of the eye [22, 70]. Therefore, we can say that accommodation and vergence are coupled with each other in binocular vision. In a monocular vision, the imagery is fixed, and only accommodation cue enables the human visual system to focus on an object. In this situation, vergence becomes an open loop as information is presented to only one eye [55]. Besides, Gabbard et al. [19] mentioned that the human visual system comfortably overrides the vergence-accommodation linkage. This linkage eventually helps the user to successfully fuse virtual contents displayed at different focal distances, which are different from the fixed focal display of AR displays.

In our research, during context switching and focal distance switching, users need to change accommodation and vergence to integrate information continuously. As accommodation and vergence are strongly correlated with each other, any mismatch between these can degrade task performance, decrease speed, reduce comfort, and also increase fatigue and eye strain among the users [48, 29, 46, 27, 44, 43, 37]. Mon-williams et al. [48] experimented to find the impact of stereoscopic depth in VR displays. They found that stereoscopic depth does not cause vision problems for a short period (10mins), but when a user continuously switches fixation point in-depth, a conflict between vergence and accommodation may cause deficits on stereo vision. However, this vergence-accommodation conflict distorts perceived depth and size as the human visual system creates the size of an object in the retina by considering the information of accommodation and vergence [55, 16, 39]. In order to reduce eye-strain and discomfort due to vergence-accommodation mismatch, Patterson et al. [55] recommended to display the virtual objects through the HMD in such a way that users do not need to change convergence angle significantly.

In the last few decades, many experiments have been conducted to resolve and minimize the conflict between vergence and accommodation in stereo displays. MacKenzie et al. [44] proposed a solution named 'depth filtering' where a sum of images was presented on several discrete focal planes and distributing the image intensity across planes based on the focus depth. The main goal of their experiment was to measure the maximum image plane separation that yields an accurate and reliable stimulus to accommodation. They considered three focal planes with five focal plane separation distances in the monocular condition. Results from their study showed that accommodation to depth filtered images was more accurate and correct when image planes have the separation of one diopter. After two years of this experiment, MacKenzie et al. [43] conducted another similar type of study by including stereoscopic condition (vergence demand) to measure the maximum image-plane spacing required for the perfect accommodation to binocular depth filtered images. Comparing with their previous study, they found that maximum image place separation could be 0.6D for binocular condition, whereas in monocular condition, maximum image plane separation could be 1.1D for accurate accommodation. They also mentioned that depth filtered images could solve many vergence- accommodation related issues, but as the image plane separation increased, the contrast, sharp and details of the images decreased. Besides, different approaches (e.g., image-based, ray based, retinal display based, and so on) to minimize vergence-accommodation conflict were briefly discussed in [37].

Accommodation and vergence of the human visual system also bring another essential factor, age. Several studies have found that the accommodative ability of the human visual system decreases with increasing age [14, 24, 25] (figure 2.8). Duane [14] also believed that ciliary muscle of the eye might get weaken with age. These findings are not surprising as the ciliary muscle adjusts the change of human eye lens for different accommodation demands. When the accommodation range of the human visual system decreases to less than 4 diopters, then the condition is known as age generating farsightedness or presbyopia [14, 22]. As measured by Duane [14], the presbyopia begins by the age of 12; during the early 30s, the accommodation loss is not high, but after that, the amplitude of accommodation falls is accelerated. Based on the human eye's accommodative ability, we can say that age would have adverse effects on the perceptual task, meaning the performance of older people could be worse than the performance of younger people [42, 59].

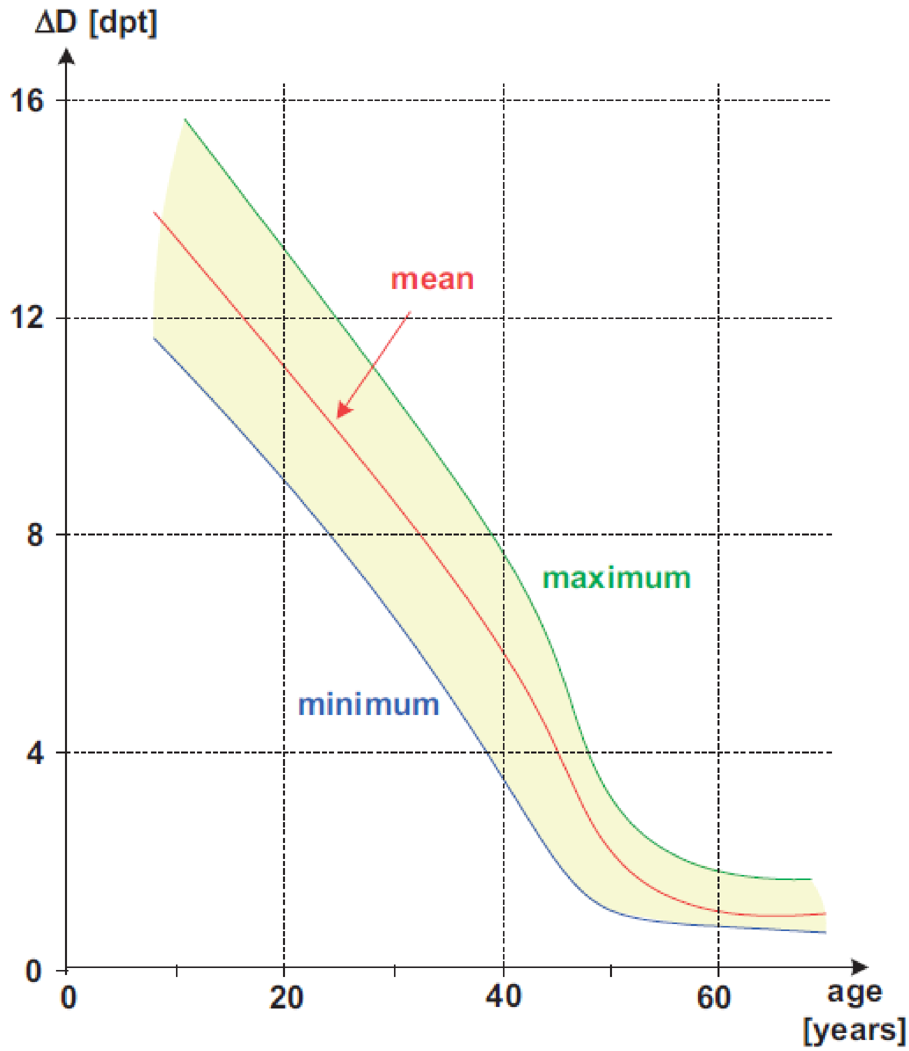


Figure 2.8: Decrease in the accommodation range with age [22]. The figure shows that in the young age (10-15 years old), human has accommodation range from 10 diopter to 15 diopter, but during the very old age (above 60 years old), the eye's accommodation range decrease to 2 diopter to 3 diopter.

2.6 Visual fatigue

According to Lambooij et al. [40], visual fatigue can be defined as "physiological strain or stress resulting from excessive exertion of the visual system." Visual fatigue has e a wide range of visual symptoms such as eye strain, blurred vision, difficulty in focusing, ache around the eyes, soreness around the eyes, among others [68]. There could be various reasons behind this visual fatigue: using eyes for more extended periods on a computer screen, reading under inadequate lighting, reading poorly printer text, and many others. According to Ukai et al. [68], one of the main reasons of visual fatigue is the accommodation and vergence conflict. Various previous studies found that changing the resting stage of accommodation and vergence increases eye fatigue and eye strain [54, 68, 27]. Previous studies ([62, 19]) directly related to our experiment (context switching and focal distance switching) showed that participants observed significant visual fatigue after completing the experiment. This observation is not surprising as the experiment generally takes a long time, and the participants hardly get any time to relax and rest their eyes, which gradually creates eye fatigue.

Measurement of visual fatigue categorizes in subjective measurement method and objective measurement method [40]. In a subjective measurement method, visual fatigue and discomfort can be measured in three ways: exploration studies, psychophysical scaling, and questionnaires. Among these three methods, questionnaires have been extensively used [54], but a general questionnaire for determining visual fatigue and discomfort has not been established [40]. On the other hand, in an objective measurement method, visual fatigue, and visual discomfort can be quantified. In this method, different optometric devices are used: autorefractor, PowerRefractor, stereo eye tracker and so on. Most of the devices are costly and not available in commercial markets.

Lambooij et al. [40] recommended that in order to measure the degree of visual fatigue accurately, reliably, and validly, both objective and subjective measurement techniques, need to be combined. In 2019, Hirota et al. [26], considered both subjective and objective measurement visual fatigue in their study. In the experiment, they measured participants' subjective and objective visual fatigue before and after performing a visual task. The results of their research showed that the objective and subjective evaluation of visual fatigue were not significantly different. Another interesting way of measuring visual fatigue is to consider the pupil size of the participant's eye before and after the experiment. Jaschinski et al. [32] mentioned that pupil size changes according to the change of eye fatigue; persons suffering eyestrain had smaller pupil sizes compared to the pupil size in normal vision with no eye fatigue.

CHAPTER 3

EXPERIMENTAL METHODS

This study aims to measure human performance and eye fatigue by empirically exploring two AR display interface issues: context switching and focal distance switching. The study has three goals. First, replicating the previous experiment of Gabbard et al. [19] using a custom-built AR haploscope. Second, extending the previous experiment for stereo condition by fully crossing the variables of context switching and focal distance switching. Finally, comparing the results of the monocular and stereo condition. To conduct this experiment, we needed some apparatus such as AR haploscope, standard monitors, tracking system, and keyboard. Besides, for analyzing the issue of context switching and focal distance switching, we needed to choose a task that distributes information both in the real world and augmented world. Therefore, we selected a text-based visual search task according to the previous study. [19]. Besides, we required to calculate the letter size and side by side distance between text blocks more precisely and accurately. To evaluate the whole study, we defined five independent variables, three dependent variables, and within the subject experimental design in the study. This chapter provides a details explanation of different apparatus, experimental tasks, experimental variables, and experimental design.

3.1 Apparatus

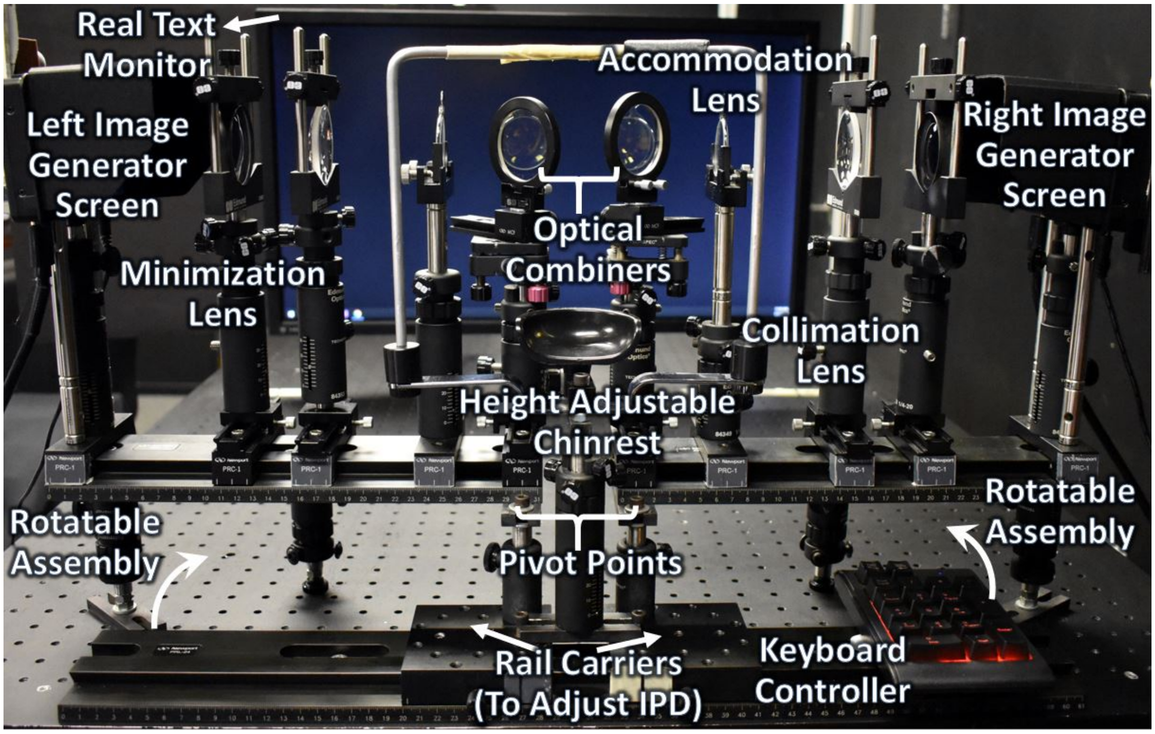


Figure 3.1: The AR haploscope. This system features fully rotatable arms, adjustable focal demand, and adjustable IPD, along with other visual parameters.

AR haploscope There are many open questions related to perception in AR; these issues range across topics like the effect of fixed device focal distances on user perception, the underestimation problem for near-field depth judgments, and the effect of various cues on user perception. To research and analyze these unresolved issues, many perception related work [56, 57, 63] were conducted in the SPAAR (Spatial Perception And Augmented Reality) lab using the AR haploscope previously. Following the previous works, a custom made AR haploscope (see figure 3.1) is used to experiment with this research. A haploscope is a tabletop augmented reality device that presents controlled augmented information both monocularly and binocularly to the user though an optical system.

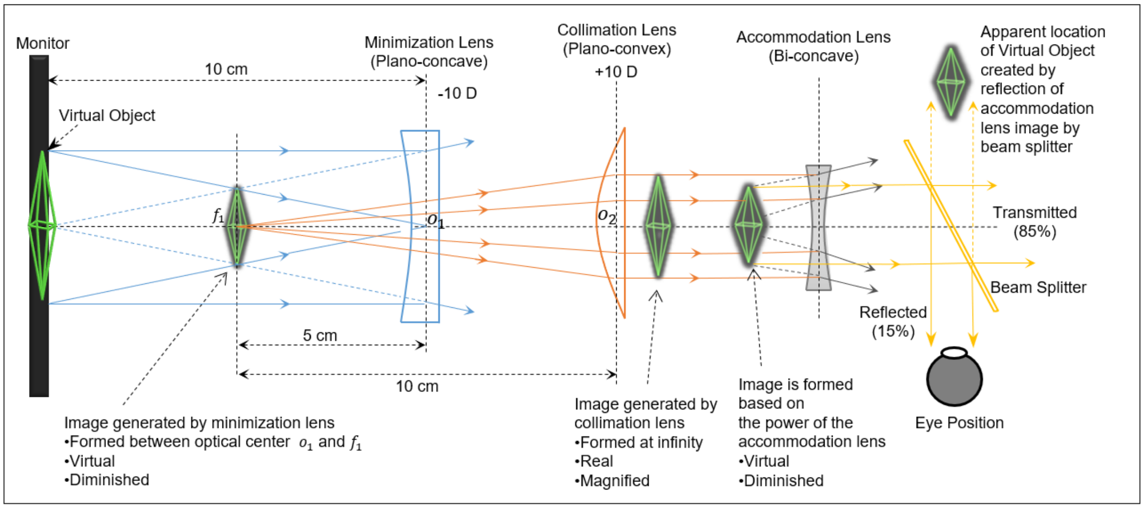
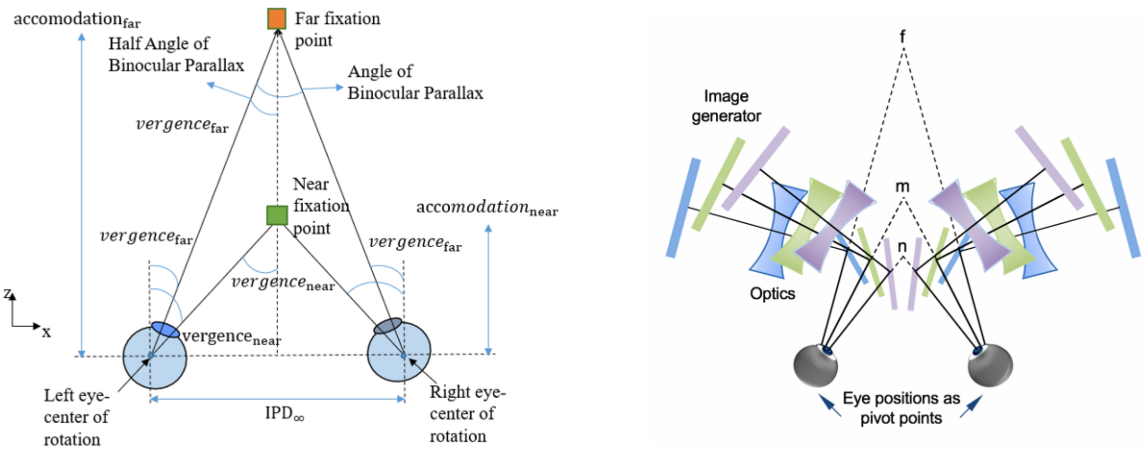


Figure 3.2: Ray diagram of the AR haploscope [57]. This diagram shows how rays pass through the AR haploscope optical system for displaying the information.

AR haploscope was built based on the design of Singh et al. [63] and the design, assembly, calibration and measurement were briefly described by Phillips et al. [56, 57]. The haploscope with a base of 61cm was mounted on the end of an optical workbench, supported by a custom build aluminum table. It has five main components (image generator monitor, minimization lens, collimation lens, accommodation lens and optical combiners) at each side of the wing. For displaying AR information, we have used 4K full HD 5.7 inch monitor with display resolution of 1920×1080 (see figure 3.4b). It has a brightness of 460 cd/m^2 and contrast ratio of 1400:1. AR haploscope has two plano-convex lenses for collimating the graphical contents coming from the displays. To resize the collimating images, a set of cylindrical concave lenses are positioned between the monitors and collimaitons lens. A set of accommodation lens (Bi-convex lens) are used in the haploscope for displaying the image through the optical combiners at a particular distance. By manipulating the accommodation lens, the AR haploscope is able to display virtual information at various accommodation demands. Finally, the rays hit the optical combiner. 85% of the light rays transmitted through the combiners and 15% are reflected back to the human eye. A ray diagram of the AR haploscope is shown in figure 3.2. Therefore, with a controlled vergence angle and accommodative demand, this device allows us to adjust each AR information and perform the experimental task monocularly as well as binocularly.

The AR haploscope is designed based on the eye model shown in Figure 3.3a. It is designed in a way that both the left and right optical apparatuses can rotate freely about a pivot point that is coincident with the estimated center of user eye rotation. The distance between the left eye center of eye rotation and the right eye center of eye rotation is known as the interpupillary distance (IPD) of the haploscope. Figure 3.3b shows that how the haploscope's assemblies rotate inward and outward to properly match with the convergence angle for a particular focal demand. According to this principle, for all the accommodation and vergence demand, the centers of the user's two eyes always stay in line with the principal axis of the optical lens system of the haploscope [63].

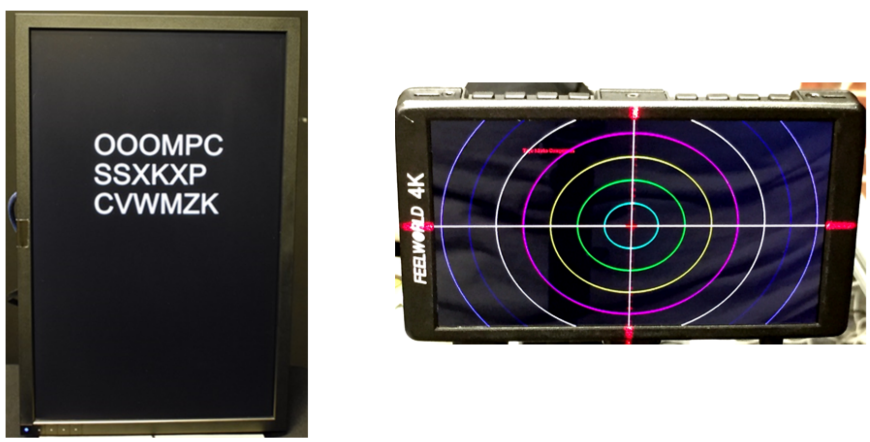
Tracking system In our experiment, we needed a tracking system to track the haploscope wings so that we could provide the appropriate vergence demand to display the virtual image at any arbitrary accommodation distance. For observing an object at a particular distance, human eyes need to rotate inward or outward. Similar to the human eye, for displaying virtual information through the haploscope at a specific distance, we rotate the haploscope's wings to a particular angle (angle of binocular parallax). It allows the user to



a) AR haploscope's eye model

b) Eye converging in AR haploscope

Figure 3.3: a) Eye model of the AR haploscope. b) This figure shows how haploscope's optical system rotates to match with the accuracte convergence angle values for different accommodation demand [63].



a) Standard PC monitor

b) AR haploscope's display

Figure 3.4: a) Standard PC monitor for displaying real information. b) AR haploscope's monitor to display virtual information.

observe the virtual information without binocular disparity. The angle of binocular parallax can be calculated by the formula:

$$\text{Angle of binocular parallax} = \arctan \frac{\text{object distance}}{(\text{IPD/2})}, \tag{3.1}$$

where IPD is the user's interpupillary distance focusing at infinity.

A tracking system is required to measure the value of the angle of binocular parallax based on the above formula. After getting the value from the tracker, we need to manually rotate each of the haploscope's wing to the tracker's calculated value. In our experiment, we used the OptiTrack tracking system (V120:Trio), which has three tracking cameras in line with 6DoF object tracking capabilities (see figure 3.5). With the help of three 640×480 VGA sensors, this tracker is capable of tracking markers down to sub-millimeter movements with high accuracy.

Figure 3.5: OptiTrack tracking device (V120:Trio)

Standard PC monitor Two standard PC monitors with adjustable brightness used in the experiment for displaying the real information (see figure 3.4a). Each of them was placed vertically during the experiment for representing information. Standard PC monitors are identical and have a resolution of 1920×1080, which is similar to the resolution of monitors used in the AR haploscope. Therefore, all the monitors used in the experiment for displaying information have the same consistent resolution.

Numeric keypad A numeric keypad was used for getting the responses from the user (see figure 3.6). Participants had the flexibility to place the keypad at their convenient position. As the experiments were conducted in a dark room, the keypad's backlight also helped the users to see the keypad's button during the experiment. Responses from the numeric keypad were stored into a data file for further analyzing the results of the experiment.

Figure 3.6: Numeric keypad

3.2 Experimental Setup

Subtask and Task In order to examine the effect of both context switching and focal distance switching empirically, a text-based visual search task that integrates the information both from the real world and augmented world was employed, previously applied by Gabbard et al. [19]. In this task, participants observed two side-by-side text blocks, *Left text block* and *Right text block*. In all conditions, the left text block was presented on a monitor, but based on the experimental conditions right text block was displayed on a monitor or displayed through the AR haploscope (see Fig 4.2). Each text block contained three text strings where each text string contains six letters. We ignored the letters 'i', 'j' and 'l' from our experiment as these letters are almost similar both in upper and lower case, Gabbard et al. [19] did the same in their experiment. Participants were instructed to focus on the left text block at first and find out the *target letter*. Target letter was side by side identical letters where the first one was the upper case, the second one was the lower case and vice versa (e.g., "Aa" where 'A' was the target letter and "bB" where 'B' was the target letter) (see Fig. 3.7). After determining the target letter, participants were instructed to focus on the right text block and count the number of times the target letter appeared. The letter appears any number of times between 0 and 3, with an upper limit of once per text string. Finally, participants gave their answers by pressing a key on the numeric keyboard.

Each full task presented up to 5 sub-tasks, and a maximum of 25 seconds was given to the participants to finish them all. Participants got a new target letter each time after providing the answer. After completion of 5 subtasks within 25 seconds or finishing 25 seconds time limit, both left and right text block became blank for 3 seconds, and then a new set of left and right text blocks appeared. In the interval between two tasks, a calibration step where the left and right text blocks were presented to the participants to ensure that left and right text blocks were observed correctly by the participants during the experimental task.

Participants completed the whole experiment in a dark room to observe the text strings without any environmental factors as Kangsoo et al. [36] found that dark mode in the AR system significantly increased the users' visual acuity during the experiment. We used sans serif font (Arial) for displaying left, and right text in our experiment as sans serif provide good readability for print media as well as preferred by the participants during the reading task from the computer display [4, 47]. In addition, all the textual information in our experiment were white and presented on a black background as Debernardis et al. [11] suggested that white text color with any mandatory background color is suitable for increasing the user performance in the AR system.

	Left text	Right text
	sKvKuS mUpKuP sOoMsP	POXCSK SZSXMM VKUKPC
Gabbard et al. [19]	AR or real world	real world
Our Experiment	real world	AR or real world

Figure 3.7: Example of the text-based task that requires integrating information presented in both the real world and AR. Participants identified the doubled target letter "O" in the left text, then counted the number of target occurrences in the right text; here the correct answer is "1."

Letter size Letter size is one of the major factors for displaying textural information in any display. If the letter size is too small or extensively big, then it creates difficulties in the human eye to capture the textual information. The size of each letter is highly dependent on the relationship of viewing angle and distance between the position of human eye and text position [18]. According to FAA human factors [18], the preferred visual angle size for text legibility is 20-22 arcminutes (0.33-0.37) degree for each letter. Previously, Gabbard et al. [19], considered relative size cue with a visual angle of 22 arcminutes in their experiment while displaying the textual information. Therefore, the retinal size of textual information is physically similar in all distances, and participants observed too big textual information at a near distance. However, according to Cutting et al. [10], in relative size cue, the size of an object can not be too large and too near. Therefore, in this study, a constant visual angle of 22 arcminutes or 0.37 degrees was used for each letter in both left and right text blocks. The letter size calculation from a constant visual angle was done using the following formula:

$$visual\ angle = 2\arctan\frac{\frac{Letter\ size}{2}}{Distance\ from\ the\ observer} \tag{3.2}$$

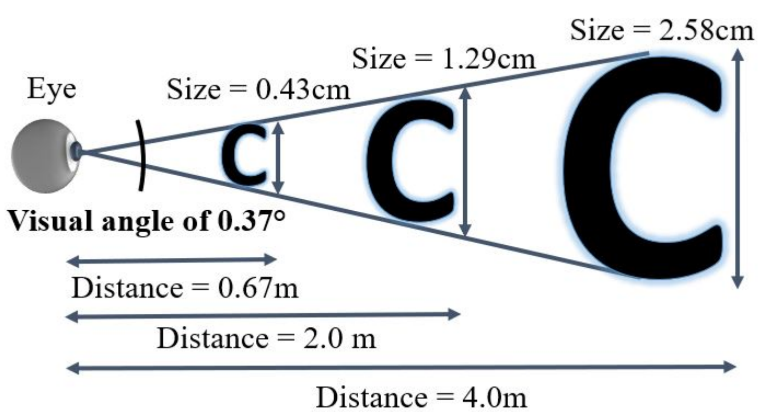


Figure 3.8: A constant visual angle of 22 arcminutes or 0.37 degree was used for each letter at all three distances. This figure shows the each letter size observed by the participants at near (0.67m), medium (2.0m) and far (4.0m) distances.

Side by side distance calculation between two text blocks In our experiment, participants observed two side by side text blocks to integrate information from both text blocks. The side by side position of the left text block and right text block was calculated based on the geometrical calculations to standardize the horizontal eye scanning [23]. To maintain this, we empirically measured the angle between the center of the left text block and the center of the right text block with respect to the eye.

At near (0.67m) distance, there is no gap between two monitors. That's why the distance between the center of the two text blocks is 33cm (empirically measured).

$$\theta = \arctan \frac{0.33m}{0.67m} \tag{3.3}$$

Therefore, angle(θ) between the center of the left text block and center of the right text block with respect to eye at near distance is 26.22 degree. As the AR haploscope has the ability to put the information perpendicularly relative to the participant eyes, we moved the left text block during the experiment such a way that angle value of 26.22 degree was maintained in each trial throughout the experiment (see figure 3.9). Accordingly, the distance between the center of the left text block and center of the right text block with respect to eye at medium and far distance are 98.5cm and 197cm.

3.2.1 Independent Variables

In our experiment, we have considered five independent variables: *Stereopsis:mono, stereo*, *Context Switching:on, off*, *Left Text Distance:0.67m (near), 2.0m (medium), 4.0m*

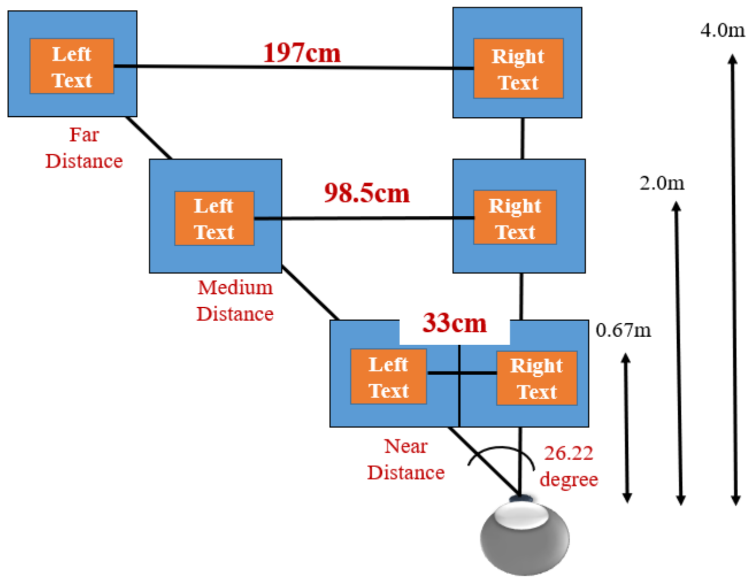


Figure 3.9: Side by side distance calculation between left and right text blocks.

(far), *Right Text Distance:0.67m (near), 2.0m (medium), 4.0m (far)* and *Repetition: 1,2,3,4,5*. Summary of the independent variables is shown in table 3.1's upper part.

Stereopsis In this study, participants participated in both monocular and stereo experimental conditions. In monocular conditions, participants completed the experiment by covering the non-dominant eye with an eye patch, and one side of the haploscope is blanked out. In *stereo* condition, participants completed the experiment with both eyes open, stereoscopically.

Context Switching In this study, when *context switching = on*, participants observed the left text on the left monitor, and the right text was presented through the AR haploscope. On the contrary, when *context switching = off*, the right text was presented through the right monitor, and left text was seen from the left monitor by the participants. Based on the context switching on/off condition, the experimenter added/removed the right monitor (see figure 3.10).

Table 3.1: Summary of the experimental variables

Independent Variables		
Stereopsis	2	mono , stereo
Context Switching	2	on , off
Left Text Distance	3	0.67m (near), 2.0m (medium), 4.0m (far)
Right Text Distance	3	0.67m (near), 2.0m (medium), 4.0m (far)
Repetition	5	1,2,3,4,5

Dependent Variables	
Subtask Completion	0,1,2,3,4,5 (times)
Subtask Accuracy	0,1,2,3,4,5 (times)
Eye Fatigue	1,2,3,4,5,6,7 (low to high)

Left Text Distance *Left Text Distance* was the distance between the left text block and the participant's eye position. Left text block denoted the real text and represented through a left monitor. Based on the literature of the AR head mounted display, the space around us can be categorized into three main levels: arm's length distance or near distance (0.7m), medium (2.0m) distance, and optical infinity (around 6.0m) [19, 29, 35, 53]. Therefore, following the previous work, we have considered three different distances of left text block in our experiment: near (0.67m), medium (2.0m), and far (4.0m). Figure 3.9 and figure 3.10 show the considered left text distances in the experiment. For each distance, the experimenter required to adjust the left monitor to the appropriate position.

Right Text Distance In our study, the *Right Text Distance* defined the distance between the right text block and the participant's eye position (see figure 3.9). Depending on the experimental condition, the right text block displayed real text or AR text, either through the right monitor or AR haploscope. In our experiment, we have considered three right text distances: near (0.67m), medium (2.0m) and far (4.0m). The experimenter needed to re-position the right monitor or change the optical accommodation lens of the haploscope to present the right text at appropriate experimental distance (see figure 3.9 and figure 3.10).

Repetition The experimental setting for each combination of *Left Text Distance* and *Right Text Distance* was repeated five times.

3.2.2 Dependent Variables

The three measured dependent variables in our study include *sub-task completion(0-5)*, *sub-task correctness (0-5)*, and *user eye fatigue rating (1-7)*. Summary of the dependent variables is shown in table 3.1's bottom part.

Subtask Completion In our study, each participant was given five sub-tasks to complete a set of full task withing 25 seconds time limit. *Sub-task Completion* means number of subtasks completed by a participant within the 25 seconds time pressure. Time limit of 25 seconds was considered according to the experiment of Gabbard et al. [19]. Therefore, each participant's number of subtask completed ranges from 0 to 5.

Subtask Accuracy In our experiment, we have considered the following error metric to evaluate the accuracy of each participant.

$$error = participant\ target\ count - correct\ target\ count \tag{3.4}$$

$$error \begin{cases} < 0, & \text{undercount error} \\ = 0, & \text{no error} \\ > 0, & \text{overcount error} \end{cases} \tag{3.5}$$

The error was calculated based on the equation 4.1, where each target count ranged from 0 to 3. We divided the error into three parts, when error = 0, the participant's response was correct, meaning *no error*. But, when error $\neq 0$, participant's response was incorrect, error could be *undercount* or *overcount* (equation 3.5). In *undercount error*, participants' answers were smaller than the actual number of targets and ranged from -1 to -3. On the other side, in *overcount error*, participants' answers were greater than the actual number of target letters, ranged from 1 to 3.

Eye-Fatigue After completing each task in all conditions, participants were asked to subjectively rate the condition of their eyes using a seven-point bipolar rating scale. The rating scale was displayed on the real text monitor, which ranged from 'very rested' to 'very fatigued.' Participants responded to their subjective eye-fatigue rating by pressing the key on the numeric keyboard.

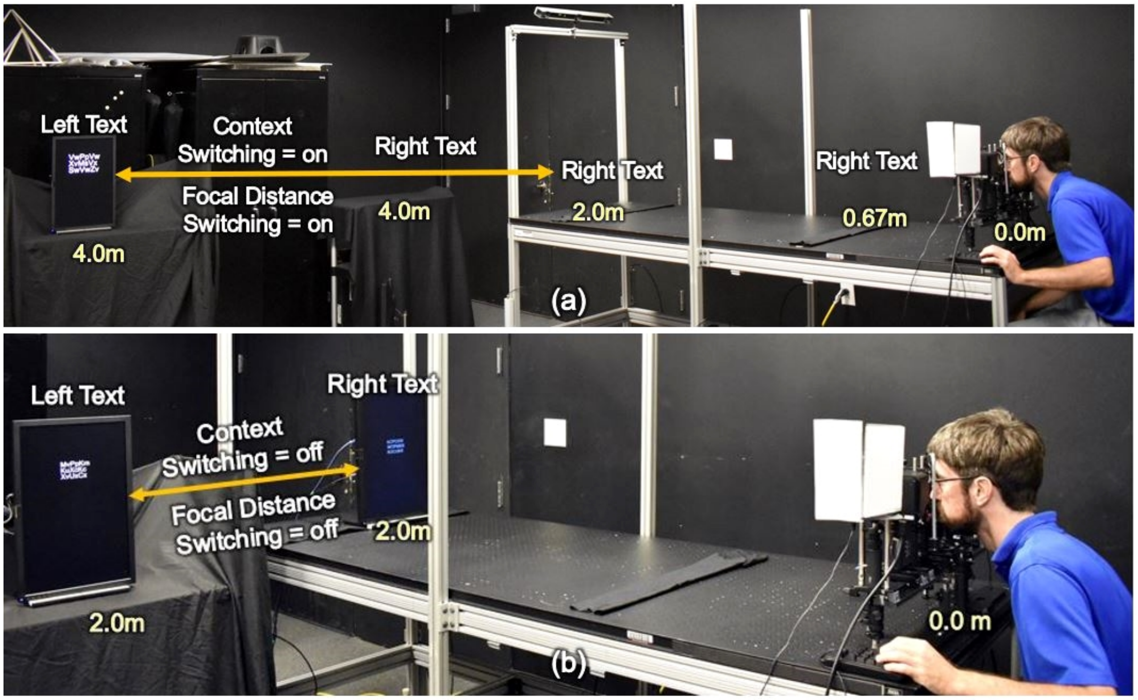


Figure 3.10: Side view of the experimental setup. A participant is performing the visual search task looking through the AR haploscope. A participant is observing the left text on the left monitor placed on a movable cart, and the right text could be positioned at three different distance levels by changing the accommodation lens of the AR haploscope (a). A participant is also performing the task when both information is presented on the two physical monitors (b). The side by side distance between the left text and right text is calculated based on the geometrical calculations to standardize the horizontal eye scanning [23]

3.3 Design

The experimental design of our study is shown in table 3.2. The table has two parts: the upper part shows the *Real World to Real World Conditions (context switching: off)*, and the lower part represents the *Real World to AR conditions (context switching: on)*. Both parts have 9 possible distance combinations, and each participant experienced all of them. Therefore, each participant observed (9+9) = 18 distance combinations in total, where all of them were unique. Further, we consider *Real World to Real World Conditions* *(context switching: off)* as a control condition in our study to fully crossing the variables of context switching and focal distance switching. The highlighted cells in the table denote the condition where left text distance and right text distance are equal; means no focal distance switching is required on those cells conditions.

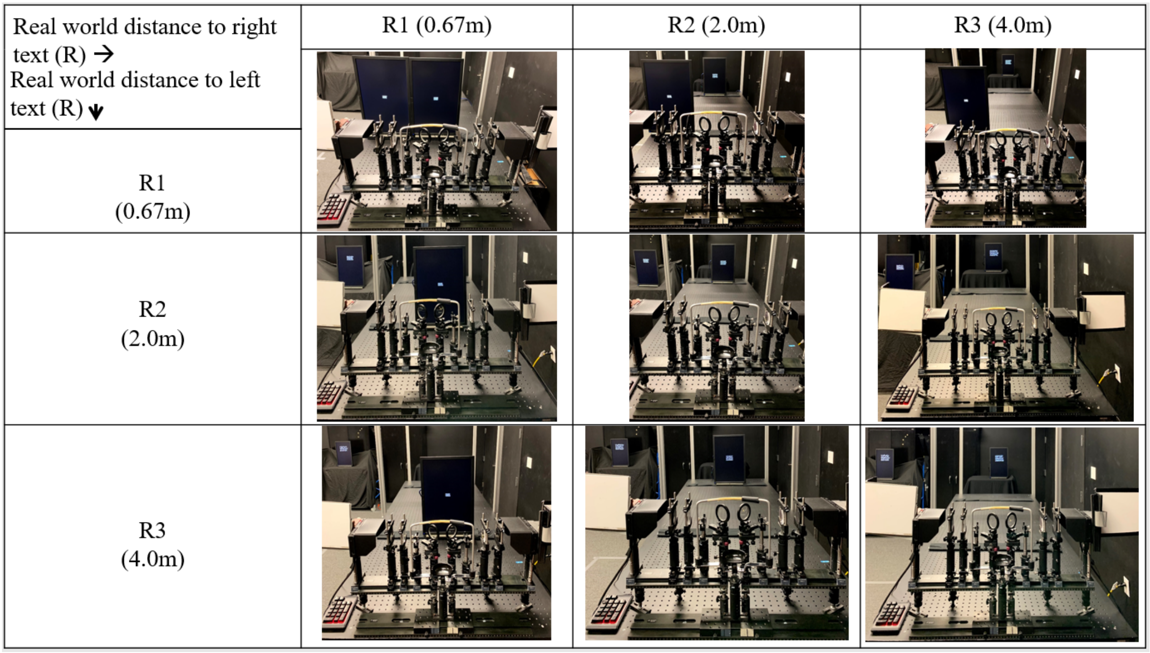
The within-subject experimental design was considered in our study so that each participant can observe all different levels of the independent variables. The presentation order of stereopsis was counterbalanced by using 2×2 Latin Square. For this, half of the participants observed the monocular condition followed by the stereo condition, and half of the participants observed the stereo condition followed by the monocular condition. Further, the presentation order of context switching was also counterbalanced by using 2×2 Latin Square. This means half of the participants observed the *context switching* = *on* condition at first than *context switching* = *off* condition, and the remaining half saw the opposite order. The remaining independent variables: Left Text Distance and Right Text Distance were counterbalanced within-participant by random permutation. Each experimental condition was repeated 5 times.

In this study, considering all the experimental variables each participant observed: $2(\textit{Stereopsis}) \times 2(\textit{Context Switching}) \times 3(\textit{Left Text Distance}) \times 3(\textit{Right Text Distance}) \times 5(\textit{repetitions}) = 180$ tasks. As mentioned earlier, each task will have up to five sub-tasks. Therefore, a total number of (180×5) = 900 sub-tasks will be completed by each participant in this study.

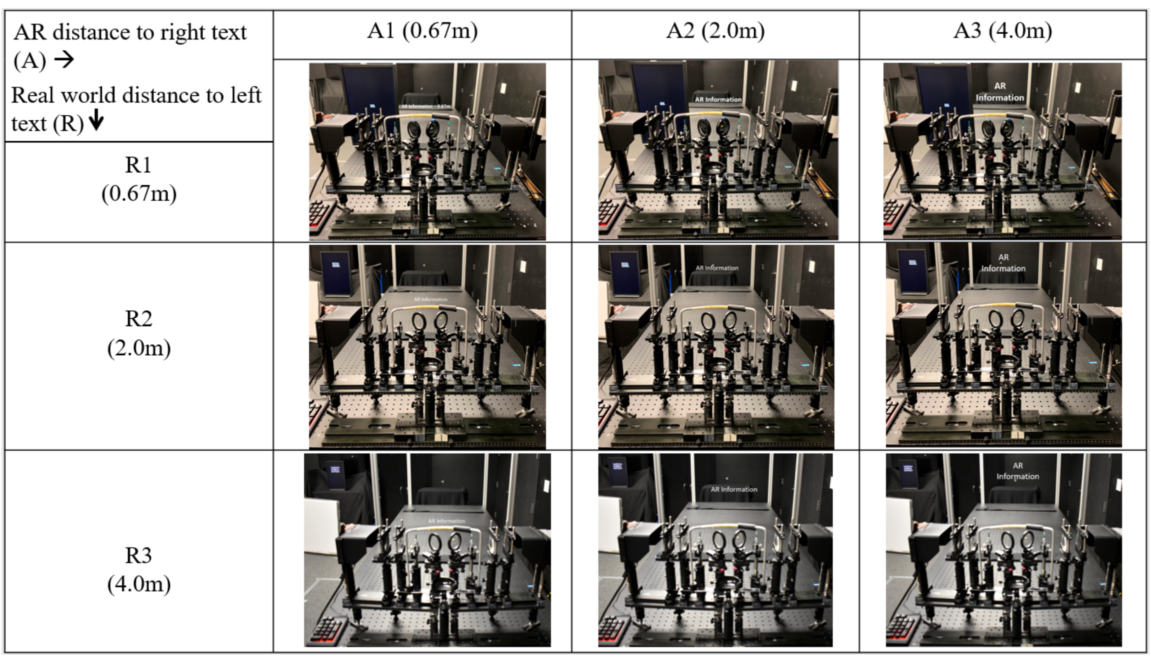
Table 3.2: Experimental design

Real World to Real World Conditions			
Real-world Distance to Left Text (R)	Real-World Distance to Right Text (R)		
	R1(0.67m)	R2(2.0m)	R3(4.0m)
R1(0.67m)	R1R1	R1R2	R1R3
R2(2.0m)	R2R1	R2R2	R2R3
R3(4.0m)	R3R1	R3R2	R3R3

Real World to AR Conditions			
Real-world Distance to Left Text (R)	AR Distance to Right Text (A)		
	A1(0.67m)	A2(2.0m)	A3(4.0m)
R1(0.67m)	R1A1	R1A2	R1A3
R2(2.0m)	R2A1	R2A2	R2A3
R3(4.0m)	R3A1	R3A2	R3A3



(a) Experimental setup based on real world to real world design conditions.



(b) Experimental setup based on real world to AR design conditions.

Figure 3.11: Experimental setup based on experimental design table 3.2. Here, each cell represents the experimental setup of each cell of the upper and lower part of the experimental design table 3.2.

3.4 Participants

We recruited 24 participants (12 male and 12 female) from Mississippi State University for this experiment with a collective mean IPD (inter-pupillary distance) of 63.1 mm. The mean age of the participants is 22.9 years, and age ranges from 18 to 31 years. There is no restriction of age and corrective lens/glasses while recruiting the participants. In this study, 17 participants were right eye dominant, and the rest of them were left eye dominant. Further, 13 participants used corrective lenses/glasses, and 11 participants had normal vision. 13 participants were recruited through the SONA PRP system for class credit, and others were graduate students who were compensated at the rate of $12 per hour. All participants were recruited and tested under local IRB rules. Each participant participated in both monocular and stereo condition in two different days within the interval of 2-3 days from the first experimental condition (monocular/stereo). However, as mentioned earlier, each participant has completed 180 tasks and 900 sub-tasks. Therefore, we could have a total up to $180 \times 24)$ = 4,320 tasks and (900×24) = 21,600 sub-tasks from 24 participants to evaluate the whole study.

CHAPTER 4

EXPERIMENT

The goal of this thesis is to empirically measure the impact of context switching and focal distance switching on human performance in AR system. To achieve this goal we divided our experiments into three parts(mentioned in section 1.3): Replication (Monocular Condition), Extension (Stereo Condition) and Comparison (Monocular vs. Stereo). In all the experimental parts, we have used the apparatus described in section 3.1, setup the experiment based on the description of section 3.2 by considering the experimental variables and design described in section 3.3.

4.1 Part 1: Replication (Monocular Condition)

The goal of this part is to observe whether our experiment successfully replicate the previous experiment conducted by Gabbard et al. [19] in completely different experimental environment with a different AR display.

4.1.1 Procedure

Participants were given a short brief about the experiment after coming to the experimental area. Then they filled out a consent form where the requirements of Mississippi State University's institutional review board (IRB) are written in detail. After that, participants completed a pre-experimental questionnaire that asked information about the participant's age, gender, ability to mentally visualize and manipulate shapes or objects, experience about virtual reality, augmented reality, or stereo glasses, uncorrected vision, color blindness, and the ability of depth perception. After that, each participant's interpupillary distance was measured at optical infinity with a pupilometer (see figure 4.1a). The AR haploscope parameters were set according to the participant's interpupillary distance. In monocular condition, the Porta test was administered, which determines the participant's dominant eye [58] (see figure 4.1b). After that, a brief description of the experimental task was given to the participants. During this instruction, participants performed several test trials using information from a single sheet of real-world paper, to familiarize themselves with the task and reduce the significance of learning effects on the results. Then, participants were asked to adjust the chair and place the keyboard in a convenient position so that participants remain comfortable during the whole experiment. After that, participants were instructed to cover the non-dominant eye by an eye patch (see figure 4.1c), and one side of the haploscope was blanked out. Participants' views through the AR haploscope during the experiment shown in figure 4.2.

4.1.2 Results

We considered repeated-measures ANOVA at the 5% significance level to analyze the result of this experiment. In the first part, we report results for the monocular condition, which matches the analysis of Gabbard et al. [19].

(a) (b) (c)

Figure 4.1: (a) Experimenter is measuring the IPD of a participant using a digital pupilometer. (b) Participant performing the porta test to determine her dominant eye before the experiment. (c) Participant covered his non-dominant eye with an eye-patch before participating in monocular condition of the experiment.

Context Switching The results indicate that there is a significant impact of context switching on subtask completion at far distance (4.0m): $F_{1,23} = 7.33, p < 0.05$. In addition, there is a significant impact of context switching on subtask accuracy at far distance (4.0m): $F_{1,23} = 6.56, p < 0.05$. This result implies that participants completed a greater number of subtasks with higher accuracy when left and right textual information were presented in the real world environment rather than one in the real world and another one in the augmented world (See figure 4.3). Therefore, context switching has negative effects on human performance in the monocular condition. This output replicates the effects found by Gabbard et al. [19]. Summary of the task completion and accuracy due to context switching at the monocular condition is shown in table 4.1.

Focal Distance Switching Both 'on' and 'off' conditions of context switching are integrated in focal distance switching. The results show that there are a significant effects of

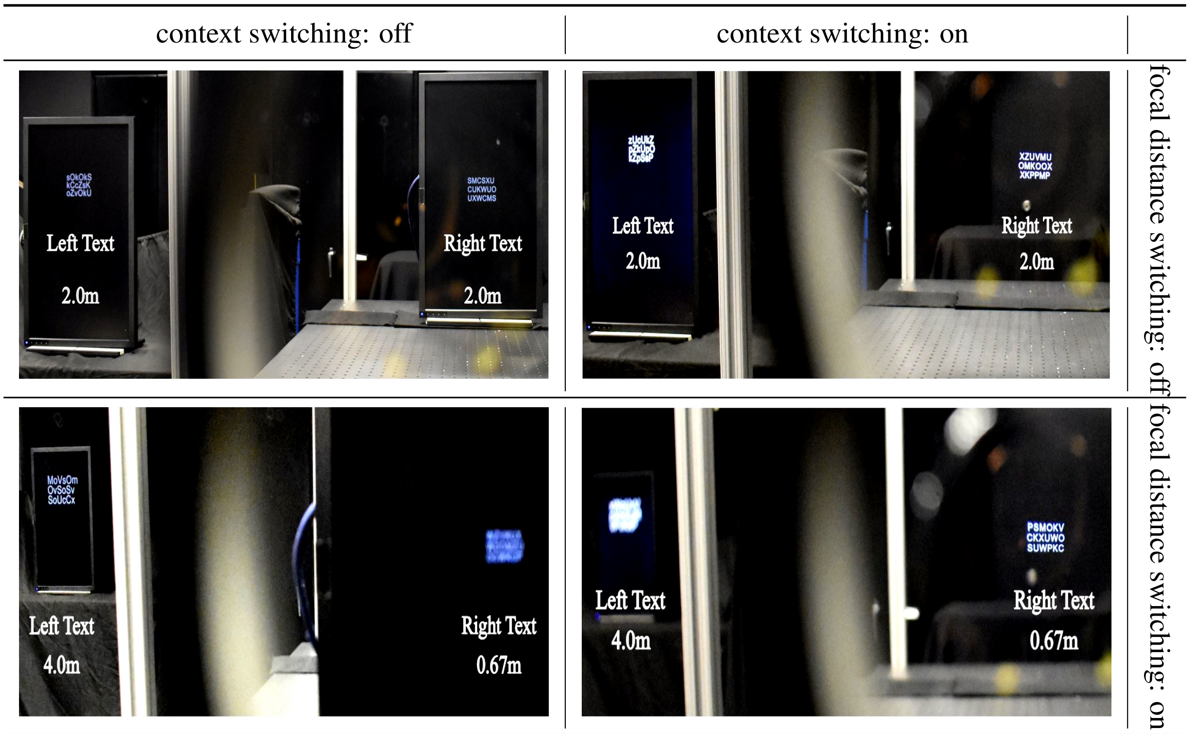


Figure 4.2: Participants view during the experiment. Participants observed both left and right textual information on physical monitors (left column) during context switching:off condition. In context switching:on condition, left textual information presented on the left physical monitor and right textual information displayed through the AR haploscope (right column). Both left and right textual information were presented at the same distance from the participant's eye position in the focal distance switching:off condition (top row). In the focal distance switching:on condition, left and right textual information were placed at different distance levels from participant's eye position. For this reason, when participants focused on the left text, right text became blurry and vice versa (bottom row).

focal distance switching on subtask completion at all three distances (0.67 meters: $F_{1,23} = 5.17, p < 0.05$; 2.0 meters: $F_{1,23} = 10.16, p < 0.05$; 4.0 meters: $F1, 23 = 5.47, p < 0.05$)). Besides, focal distance switching has significant effect on subtask accuracy at all three distances (0.67 meters: $F_{1,23} = 8.77, p < 0.05$; 2.0 meters: $F_{1,23} = 5.87, p < 0.05$; 4.0 meters: $F1, 23 = 4.99, p < 0.05$)) (see figure 4.4).

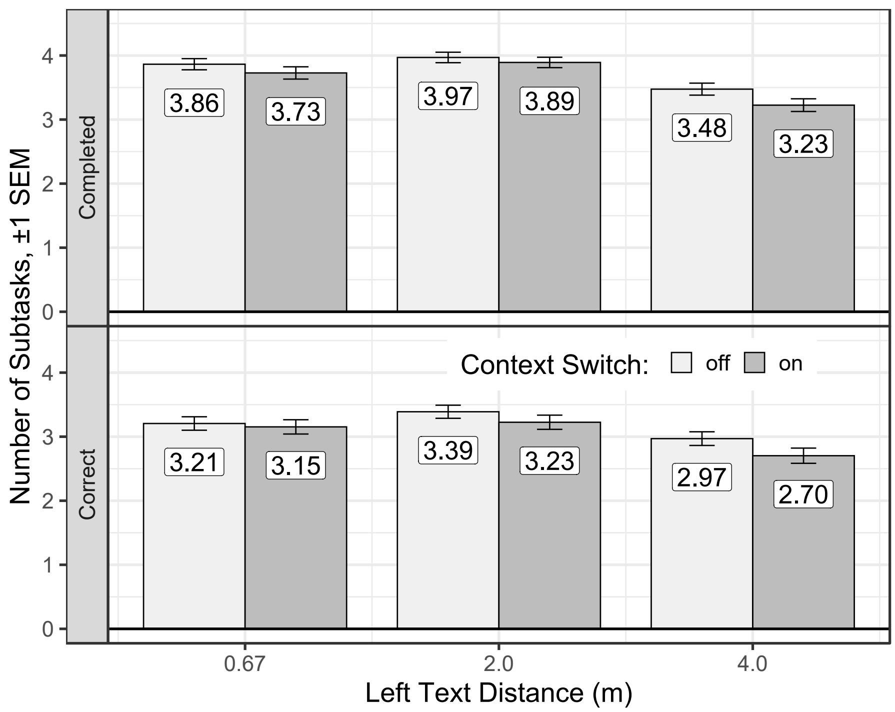


Figure 4.3: This graph shows the impact of context switching on user performance in the monocular condition. The X-axis shows three different left text distance levels: near (0.67m), medium (2.0m), and far (4.0m). The upper grid of the Y-axis shows the number of completed subtasks by the participants, and the lower grid denotes the accuracy of the participants. From the graph, it is visible that participants have better completeness and accuracy in the context switching:off condition compared to the context switching:on condition. At far(4.0m) distance, participants' performance degrades significantly.

Figure 4.5b shows a significant interaction effect of focal distance switching and whether there was a target letter in the first line of text ($F_{1,23} = 24.27, p < 0.05$), as well as related main effects of focal distance switching ($F_{1,23} = 47.17, p < 0.05$) and target letter in first line of text ($F_{1,23} = 54.89, p < 0.05$). The interaction in Figure 4.5b is consistent and support the Gabbard et al. [19] findings.

Eye fatigue Figure 4.6b shows that there is a significant effect of context switching: integrating information between AR and the real world resulted in significantly higher

Table 4.1: Mean and standard deviation of task completion and task accuracy for context switching at monocular condition

Left text distance levels	Task completion			
	Context switching: on		Context switching:off	
	Mean	SD	Mean	SD
Near (0.67m)	3.73	0.67	3.86	0.62
Medium (2.0m)	3.89	0.61	3.97	0.60
Far (4.0m)	3.23	0.74	3.48	0.71
	Task accuracy			
	Context switching: on		Context switching:off	
	Mean	SD	Mean	SD
Near (0.67m)	3.15	0.81	3.21	0.72
Medium (2.0m)	3.23	0.78	3.39	0.67
Far (4.0m)	2.70	0.84	2.97	0.77

levels of eye fatigue at all distances (0.67 meters: $F_{1,23} = 7.58, p < 0.05$; 2.0 meters: $F_{1,23} = 4.87, p < 0.05$; 4.0 meters: $F_{1,23} = 8.63, p < 0.05$)). This replicates the effects found by Gabbard et al. [19] (Figure 4.6a).

Figure 4.7 presents that participants observed less fatigue when there is no or small focal distance switching required. As the requirement of focal distance switching increased, participants' eye-fatigue increased significantly. According to statistical analysis, focal distance switching resulted significant eye fatigue at all the distances (0.67 meters: $F_{1,23} = 24.12, p < 0.05$; 2.0 meters: $F_{1,23} = 38.57, p < 0.05$; 4.0 meters: $F1, 23 = 15.28, p < 0.05$)). Mean and standard deviation of participants' eye fatigue at monocular condition presented in the table 4.2.

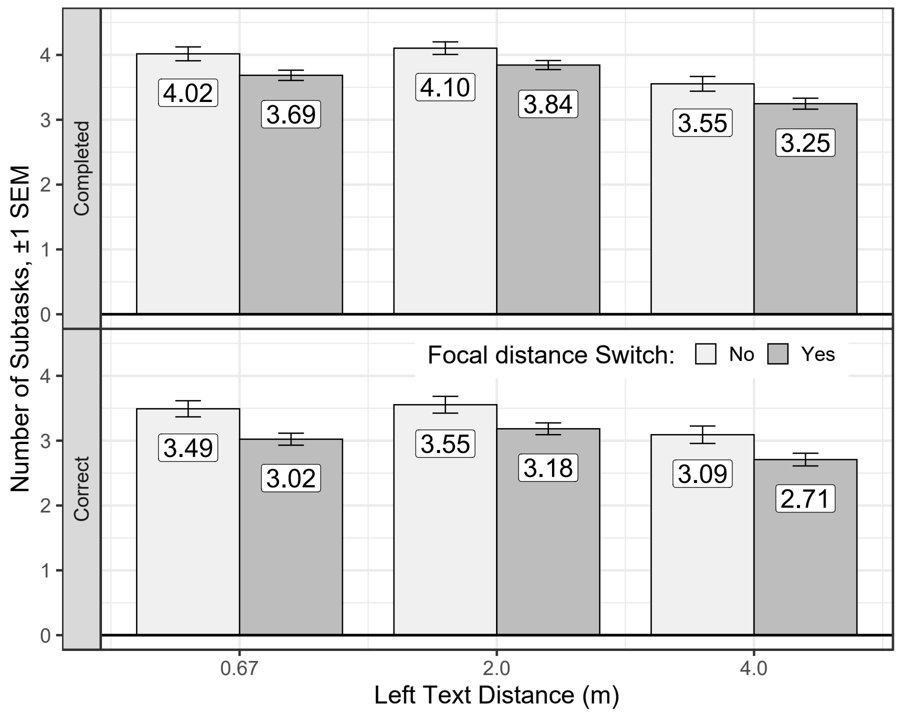
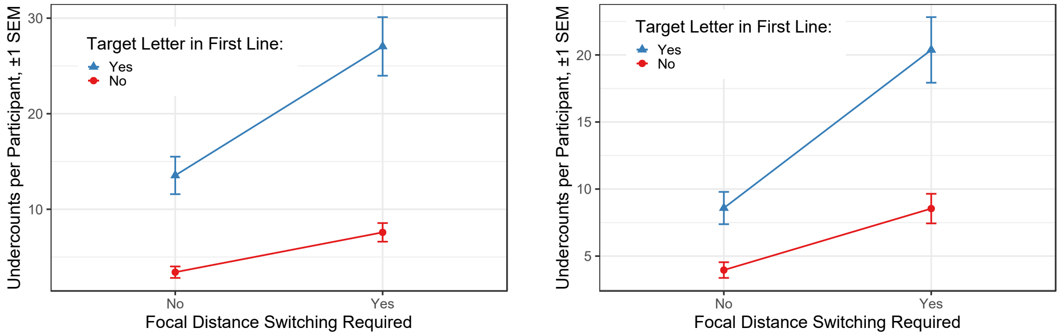


Figure 4.4: This graph shows the effect of focal distance switching on participants' task performance in the monocular condition. The X-axis of the graph shows three different distance levels: near (0.67m), medium (2.0m), far (4.0m) and the upper grid of the Y-axis denotes the number of subtasks completed(out of 5), and the lower grid of the Y-axis indicates the number of correctness by participants. Participants completed a fewer number of subtasks with less accuracy at focal distance switching:on condition compared to the focal distance switching:off condition at all three distance levels.

$$Amount\ of\ focal\ distance\ switching = |Left\ text\ distance - Right\ text\ distance| \tag{4.1}$$

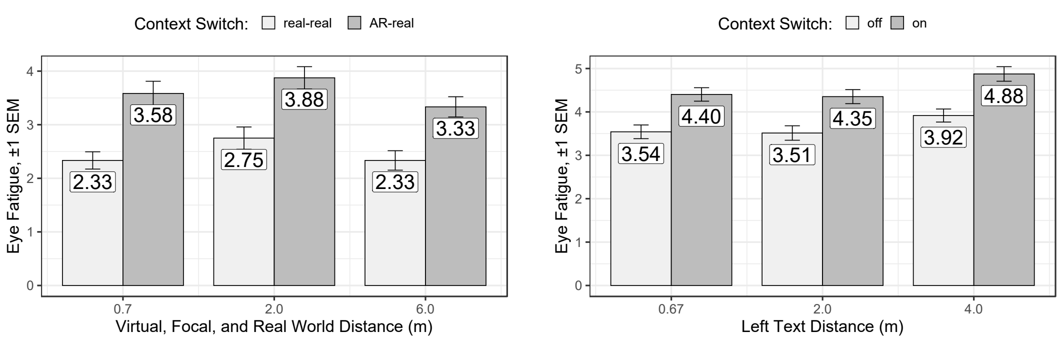
4.1.3 Discussion

As mentioned earlier that our first goal is to successfully replicate the previous experiment conducted by Gabbard et al. [19] to establish a general AR phenomena on user



(a) Data from Gabbard et al. [19]. (b) Data collected from AR haploscope.

Figure 4.5: Participants undercounted more letters when a target letter appeared in the first line of text, and when focal distance switching was required. This indicates that when participants had to switch focal distances, they began scanning the first line for a target letter before their eyes had finished accommodating. This made the text blurry, and therefore they were more likely to miss the target letter (b). This replicates the effect found by Gabbard et al. [19] (a).



(a) Data from Gabbard et al. [19]. (b) Data collected from AR haploscope.

Figure 4.6: Context switching between AR and real-world visual information resulted in significantly higher levels of reported eye fatigue at all distances, from both the AR haploscope (b) and Gabbard et al. [19] (a). Context switching "off" identical to context switching "real-real" and context switching "on" identical to context switching "AR-real".

interface design considering context switching and focal distance switching. However, Gabbard et al. [19] only considered three(3) real-real distance combination in their experiment (R1R1, R2R2 and R3R3) wheres as we considered all the nine(9) distance com-

Table 4.2: Mean and standard deviation of eye fatigue at monocular condition

Left text distance levels	Eye fatigue			
	Context switching: on		Context switching:off	
	Mean	SD	Mean	SD
Near (0.67m)	4.40	1.06	3.54	1.15
Medium (2.0m)	4.36	1.10	3.51	1.26
Far (4.0m)	4.88	1.20	3.92	1.06
	focal distance switching: no		focal distance switching: yes	
	Mean	SD	Mean	SD
Near (0.67m)	3.38	1.01	4.27	0.84
Medium (2.0m)	3.21	0.99	4.27	0.74
Far (4.0m)	3.83	1.12	4.68	0.84

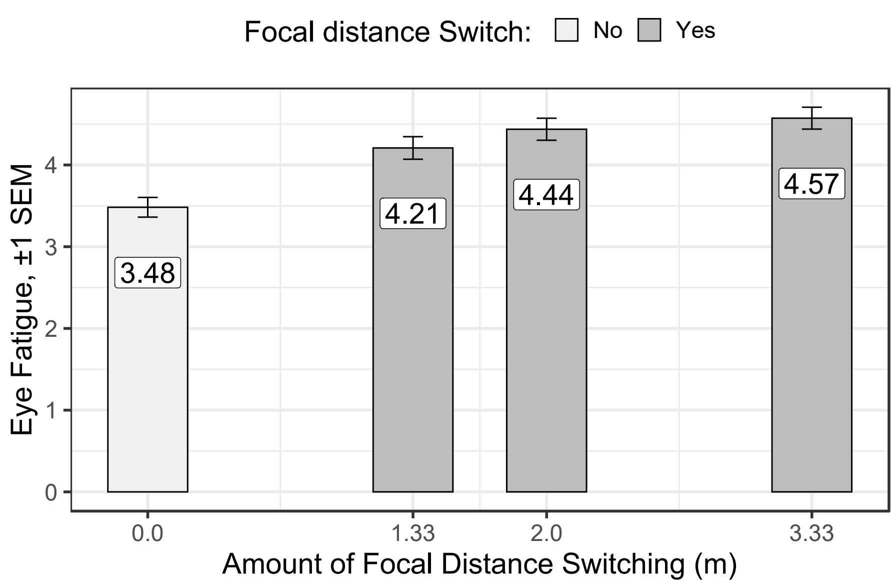


Figure 4.7: The X-axis of the graph shows the amount of focal distance switching and the Y-axis denotes participants' eye fatigue. From the graph, it is visually understandable that participants have less eye fatigue within small focal distance switching distance, but when the switching distance increased, participants' eye-fatigue increased notably.

bination in our experiment (table 3.2). With more empirical data, our results show that context switching and focal distance switching together resulted in significantly reduce performance (figure 4.3 and 4.4), further supporting the observations made by Gabbard et al. [19].

Both our findings and Gabbard et al. [19] findings found that in context switching condition there is no difference in participants' task performance at near (0.67m) and medium (2.0m) distances. However, participants have relatively poor performance in the far(4.0m) distance though the text legibility standard of 22 arc minutes was consistent. Like Gabbard et al. [19], most of the participants also mentioned about blurriness at far(4.0m) distance because of the eye-fatigue and tiredness, which eventually degraded their performance. Furthermore, participants also rated the AR-real condition more fatiguing than the real-real condition (Figure 4.6b).

In focal distance switching condition, participants complete fewer tasks and are also less accurate at all three distances (Figure 4.4). In addition, particiapants' eye fatigue is higher when the focal distance switching amount is significant. According to Tufano et al. [67], eyes accommodation and vergence resting point position is said to be around the arm's length distance (0.67m). When the position of the textual information is closer to the eye's resting point, participants require minimal focus adjustments to integrate the information. For this reason, as the amount of focal distance switching increased, participants' eye fatigue increased gradually (Figure 4.7).

Additionally, participants undercounted more letters when a target letter appeared in the first line of text, and when focal distance switching was required (figure 4.5b). The reason is that when the participant had to switch focal distances and the target letter was in the first line of text, they tried to search the line while their eyes were still accommodating to the new distance. Changing accommodating is relatively slow, taking anywhere from ∼360 to ∼425+ milliseconds [6]. In addition, the task put the participant under time pressure. This is the most likely explanation for the interaction effect in Figure 4.5b. This also replicates the same effect found by Gabbard et al. [19] (Figure 4.5a). Further, participants found integrating information from different focal distances more fatiguing than information presented at the same focal distance.

4.2 Part 2: Extension (Stereo Condition)

Experiment 2 is the extended version of experiment 1. In this experiment, we are interested in determining if there are any effects of context switching and focal distance switching exist in stereo AR display or not. If the effect exists, we are interested to empirically measure the effects of context switching and focal distance switching in AR display. In stereo condition, participants' also observed the same view through the AR haploscope during the experiment(see figure 4.2).

4.2.1 Procedure

The procedure of Experiment 2 was similar to that of Experiment 1, except that the participants did not have to perform the porta test to find their dominant eye and did not use eye-patch to cover their non-dominant eye. Participants experimented with two eyes, and appropriate vergence and accommodation demand provided through the AR haploscope. .

4.2.2 Results

Similar to the experiment 1, we considered repeated-measures ANOVA at the 5% significance level to analyze the experimental results. In this part, we report and discuss the results of stereo condition.

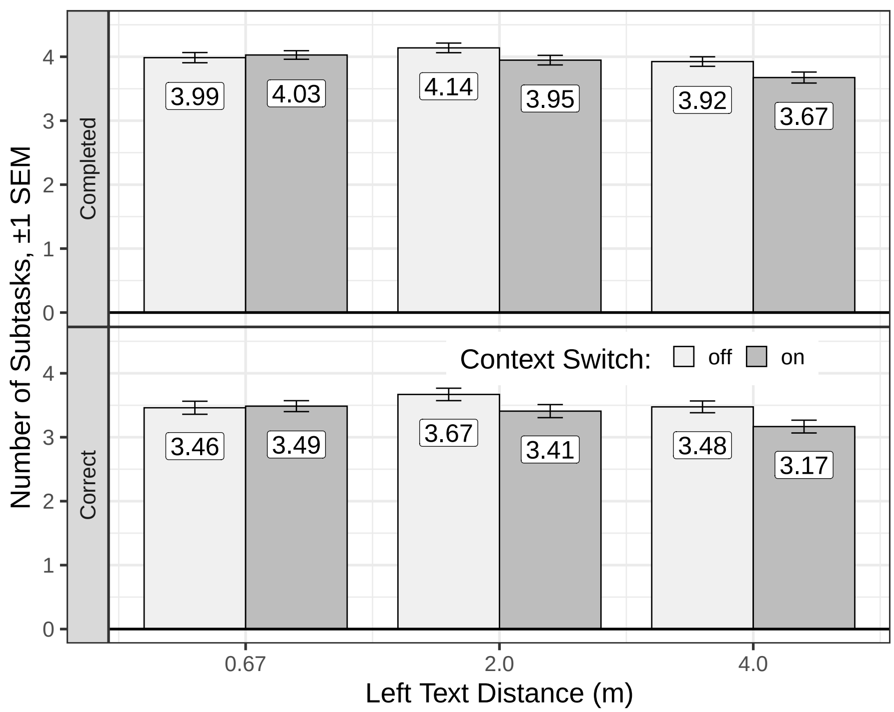


Figure 4.8: The graph shows the effect of context switching on participants' performance in stereo condition. Participants' subtask completeness and accuracy at 0.67m and 2.0m did not differ much between context switching "on" and "off" condition. At 4.0m, participants had relatively poor performance compared to 0.67m and 2.0m distances.

Context Switching In stereo condition, there are significant effects of context switching on both subtask completion ($F_{1,23} = 5.511, p < 0.05$) and subtask accuracy ($F_{1,23} = 4.85, p < 0.05$) at far distance (4.0m). Therefore, participants had better performance when both information were presented in the real world environment rather than on the combination of real and AR world (see figure 4.8). These findings are similar to the finding of the monocular condition and previous work of Gabbard et al. [19]. Table 4.3 reports the mean and standard deviation of task completion and task accuracy for context switching at all three left text distance levels.

Table 4.3: Mean and standard deviation of task completion and task accuracy for context switching at stereo condition

Left text distance levels	Task completion			
	Context switching: on		Context switching:off	
	Mean	SD	Mean	SD
Near (0.67m)	4.03	0.54	3.99	0.64
Medium (2.0m)	3.95	0.57	4.14	0.57
Far (4.0m)	3.68	0.61	3.93	0.58
	Task accuracy			
	Context switching: on		Context switching:off	
	Mean	SD	Mean	SD
Near (0.67m)	3.49	0.69	3.46	0.80
Medium (2.0m)	3.41	0.79	3.67	0.78
Far (4.0m)	3.17	0.76	3.48	0.71

Focal Distance Switching Like monocular condition, both context switching 'on' and 'off' conditions are integrated in the focal distance switching analysis. In stereo condition, there are significant impact of focal distance switching on the subtask completion at medium (2.0m) distance:$F_{1,23} = 8.17, p < 0.05$ and far (4.0m) distance: $F_{1,23} = 7.51, p < 0.05$. In addition, figure 4.9 show that focal distance switching results significant impact on the subtask accuracy at medium (2.0m):$F_{1,23} = 7.32, p < 0.05$, and far distance (4.0m):

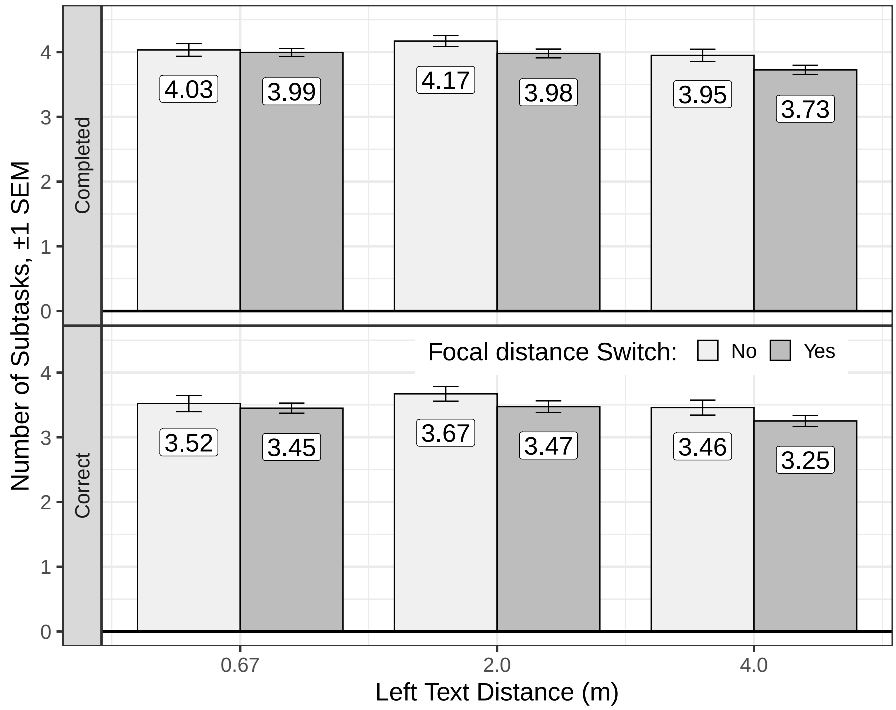


Figure 4.9: This graph shows the effect of focal distance switching on participants' task performance in the stereo condition. Participants' performance in the near (0.67m) distance is not affected much by focal distance switching. Nevertheless, at medium (2.0m) and far (4.0m) distances, participants completed a fewer number of subtasks with less accuracy at focal distance switching:on condition compared to the focal distance switching:off condition.

$F_{1,23} = 4.72, p < 0.05$. There is no impact of focal distance switching on participants' task performance and eye fatigue at near (0.67m) distance.

Similar to the monocular condition and Gabbard et al. [19] findings, figure 4.10 shows that there is also a significant interaction between focal distance switching and whether there was a target letter in the first line of text ($F_{1,23} = 24.92 < 0.05$), as well as related main effects of focal distance switching ($F_{1,23} = 38.67, p < 0.05$) and target letter in first line of text ($F_{1,23} = 32.62, p < 0.05$) in the stereo condition.

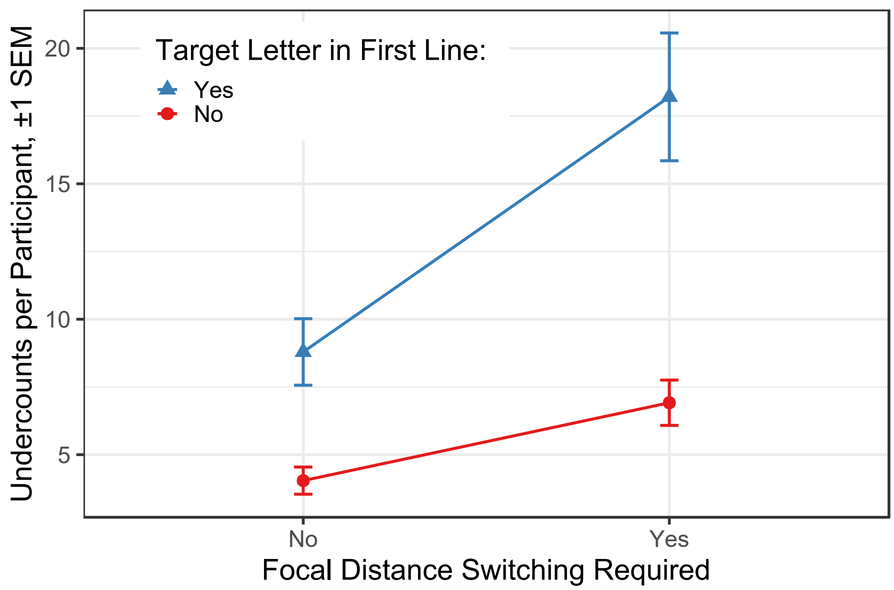
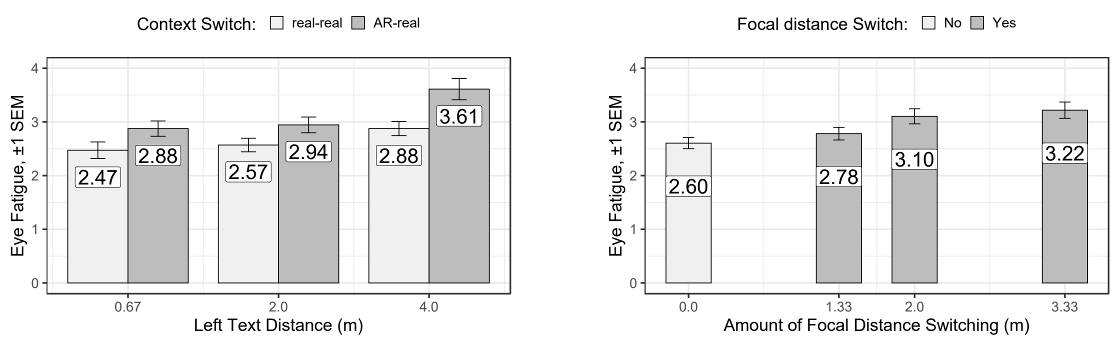


Figure 4.10: This figure shows that participants undercounted more letters when a target letter appeared in the first line of text, and when focal distance switching was required. This outcome supports the effect found by Gabbard et al. [19] and findings of monocular condition (figure 4.5).



(a) Context switching effects on eye-fatigue.

(b) Eye fatigue at various amount of switching distances.

Figure 4.11: (a) Context switching between AR and real-world visual information resulted in significantly higher levels of reported eye fatigue only at the far distance at stereo condition. (b) As the amount of focal distance switching distance increased, participants' eye-fatigue increased significantly in the stereo condition.

Eye fatigue Unlike the monocular condition and findings of Gabbard et al. [19], context switching resulted in significant effects of eye fatigue only at the far distance (4.0m): $F_{1,23} = 7.52, p < 0.05$. This result implies that participants found stereo condition less fatiguing at near and medium distances (Figure 4.11a). Again, we found significant effect of focal distance switching on participants eye fatigue at medium (2.0m):$F_{1,23} = 11.16, p < 0.05$, and far distance (4.0m): $F_{1,23} = 13.53, p < 0.05$ in stereo condition (Figure 4.11b), not at the near (0.67m) distance. Summary of the eye fatigue results is provided in table 4.4.

Table 4.4: Mean and standard deviation of eye fatigue at stereo condition

Left text distance levels	Eye fatigue			
	Context switching: on		Context switching:off	
	Mean	SD	Mean	SD
Near (0.67m)	2.88	1.09	2.47	1.27
Medium (2.0m)	2.94	1.15	2.57	0.93
Far (4.0m)	3.61	1.40	2.88	0.95
	Focal distance switching: no		Focal distance switching: yes	
	Mean	SD	Mean	SD
Near (0.67m)	2.54	1.22	2.74	0.95
Medium (2.0m)	2.56	0.68	2.85	0.86
Far (4.0m)	2.71	1.02	3.51	1.17

4.2.3 Discussion

In this part, our goal is to explore the effects of context switching and focal distance switching on human performance by extending the experiment to stereo condition. We hypothesized that participants would experience less fatigue, and performance would be better as the experiment was done stereoscopically rather than with the dominant eye.

Similar to the monocular condition findings and Gabbard et al. [19] results, there is no difference in participants' task completeness and accuracy at near (0.67m) and medium (2.0m) distances. However, participants have completed a fewer number of subtasks with lower accuracy at far (4.0m) distance. This result implies that positioning textual information at far distance degrades user performance in stereo condition too. Though there is a significant impact of context switching on participants' task performance at far (4.0m) distance, unlike the monocular condition and Gabbard et al. [19] findings, there is no significant impact of context switching on participants' eye fatigue at near (0.67m) and medium (2.0m) distances. According to Tufano et al. [67], resting point of eyes accommodation and vergence is said to be around the arm length distance (0.67m). Near (0.67m) and medium (2.0m) distances are correspondingly equal and closer to the resting point distance of the eye compared to the far (4.0m) distance. Besides, in this condition, participants did not feel eye strain only on one eye. For this reason, participants felt less fatiguing at near and medium distance during context switching in the stereo condition.

There is a significant effect of focal distance switching on participants' performance at medium(2.0) and far(4.0) distance. Participants completed fewer number of subtasks with lower accuracy and rated high eye-fatigue value at medium (2.0) and far(4.0) distances. However, there is no significant effect of focal distance switching on participants performance at near(0.67m) distance. Further, focal distance switching has no significant effect on participants eye-fatigue in the near(0.67m) distance. Participants eye fatigue is nearly equivalent when there is no focal distance switching and amount of focal distance switching is very small (see figure 4.11b). This finding is not surprising. As mentioned above, resting point of eyes is said to be around the arm length distance (0.67m) [67]. So, it is acceptable that there is no impact of focal distance switching on participants' performance and eye-fatigue at near (0.67m) distance. Therefore, in stereo condition, participants performance and eye fatigue will be same when left text (real text) is fixed at near(0.67m) and right text(AR text) is presented at any of the three distances (near(0.67m), medium(2.0m), far(4.0m)).

Supporting the findings of the monocular condition and Gabbard et al. [19] results, there is a significant interaction between focal distance switching and whether there was a target letter in the first line of text. There are also main effects of focal distance switching, and the target letter in the first line of text in the stereo condition exists. It implies the hypothesis: participants were more likely to miss target letters in the first line of the right text in focal distance switching required condition, as opposed to target letters in the second or third line, which is also consistent in the stereo condition.

4.3 Part 3: Comparison (Monocular vs. Stereo)

After determining the effects of context switching and focal distance switching both on monocular as well as stereo condition, finally, we are interested in comparing participants' performance and eye fatigue between monocular and stereo condition. participants participated in both monocular and stereo condition as we considered within-subject design for our experiments. In the monocular condition, each participant observed: $2(\textit{Context Switching}) \times 3(\textit{Left Text Distance}) \times 3(\textit{Right Text Distance}) \times 5(\textit{repetitions}) = 90$ tasks and $5 \times 90 = 450$ sub-tasks. Each participant also performed a similar number of tasks (90) and sub-tasks (450) in stereo condition too. Therefore, 24 participants performed 90×24 = 2,160 tasks and 90×450 = 10,800 sub-tasks both in monocular as well as stereo condition. We considered above mentioned number of tasks and sub-tasks while comparing monocular condition with the stereo condition. We examined repeated-measures ANOVA at 5% significance level to compare participants' performance and eye fatigue between monocular and stereo condition.

4.3.1 Results

Both context switching and focal distance switching conditions are integrated in stereopsis condition analysis. Figure 4.12 interprets that participants completed more subtasks in stereo condition than monocular condition (also see table 4.5), and there are significant main effects of stereopsis on subtask completion at near and far distances (0.67 meters: $F_{1,23} = 5.35, p < 0.05$; 4.0 meters: $F1, 23 = 20.92, p < 0.05$)). In addition, stereopsis has significant effect on the subtask accuracy at near distance (0.67 meters: $F_{1,23} = 9.56, p < 0.05$) and far distance (4.0 meters: $F1, 23 = 17.45, p < 0.05$)). Participants were more accurate in the stereo condition compared to the monocular condition (figure 4.12 and table 4.5). There exists no significant interaction effects in this analysis. Table 4.5 depicts the mean and standard deviation of task completion and task accuracy at stereopsis condition.

We hypothesized that stereo condition would be less fatiguing than monocular condition. Results support our hypothesis and participants experienced less eye fatigue in the stereo condition compared to monocular condition in all distances (figure 4.13 and

Table 4.5: Mean and standard deviation of task completion and task accuracy at stereopsis condition

Left text distance levels	Task completion			
	Stereopsis: mono		Stereopsis: stereo	
	Mean	SD	Mean	SD
Near (0.67m)	3.80	0.56	4.00	0.53
Medium (2.0m)	3.93	0.51	4.04	0.51
Far (4.0m)	3.35	0.69	3.80	0.53
	Task accuracy			
	Stereopsis: mono		Stereopsis: stereo	
	Mean	SD	Mean	SD
Near (0.67m)	3.18	0.73	3.47	0.70
Medium (2.0m)	3.31	0.66	3.54	0.72
Far (4.0m)	2.84	0.77	3.32	0.65

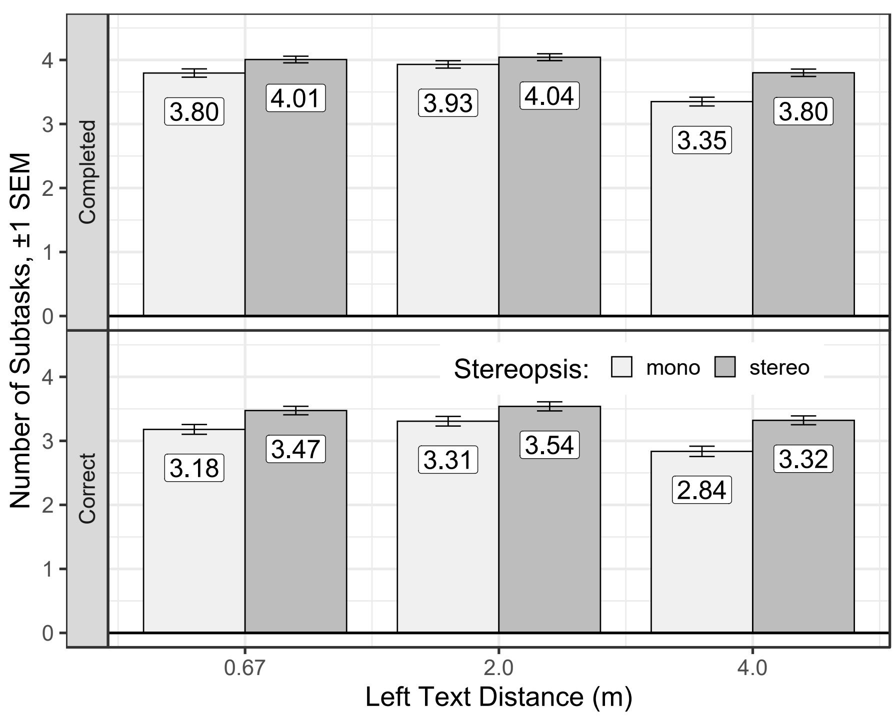


Figure 4.12: Participants completed a fewer number of subtasks significantly with less accuracy in the monocular condition at all the distances compared to the stereo condition.

table 4.6). Statistical analysis also shows that there is a significant effect of stereopsis on eye fatigue at all three distances (0.67 meters: $F_{1,23} = 22.49, p < 0.05$; 2.0 meters: $F_{1,23} = 28.90, p < 0.05$; 4.0 meters: $F1, 23 = 18.68, p < 0.05$)). Summary of the participants' eye fatigue at stereopsis condition is shown in table 4.6.

Table 4.6: Mean and standard deviation of eye fatigue at stereopsis condition

Left text distance levels	Eye fatigue			
	Stereopsis: mono		Stereopsis: stereo	
	Mean	SD	Mean	SD
Near (0.67m)	3.97	0.80	2.67	0.96
Medium (2.0m)	3.93	0.73	2.76	0.78
Far (4.0m)	4.40	0.80	3.24	1.00

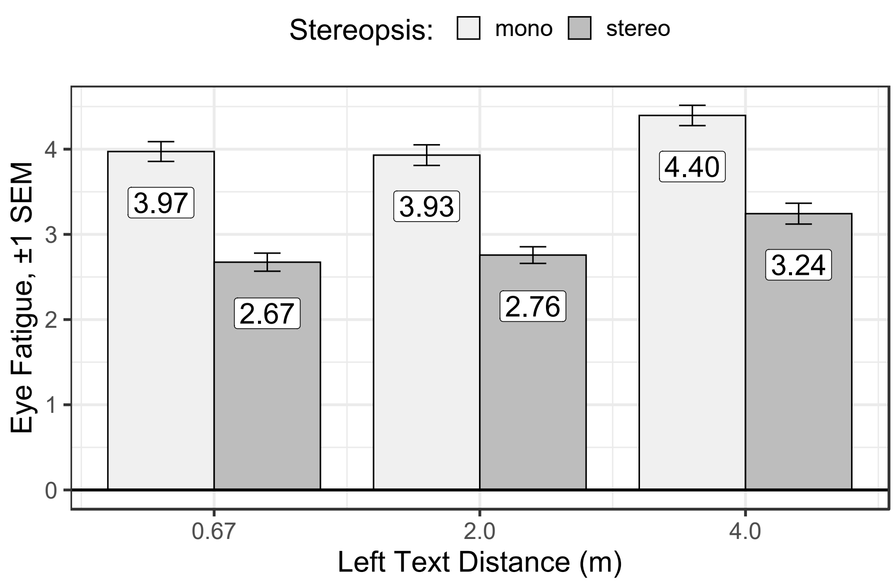


Figure 4.13: The monocular condition resulted in significantly higher levels of eye fatigue at all three distances compared to the stereo condition. This graph illustrates that participants were more comfortable in the stereo condition over the monocular condition.

4.3.2 Discussion

In the final part, we compare participants' performance and eye-fatigue between monocular and stereo condition by considering an extra variable named "Stereopsis". This comparison would be the first empirical comparison between monocular and stereo conditions considering the effects of context switching and focal distance switching in AR. Participants completed a fewer number of subtasks with lower accuracy in monocular condition compared to the stereo condition (figure 4.12 and table 4.5). This result is not surprising, as Laramee et al. [41] stated that user performance would be slower while completing a visual scanning task wearing a monocular display compared to the stereo display. There is a significant impact of stereopsis on participants' subtask completeness and accuracy at near (0.67m) and far (4.0m) distances, not in the medium (2.0m) distance. Moreover, feedback was also gathered from the participants after the experiments in an informal interview. About 80% of the participants preferred left text distance at 2.0m during the experiment, but nobody preferred the left text position at 4.0m in both monocular and stereo conditions. Based on the participants' feedback and statistical analysis, we can say that in the medium (2.0m) left text distance participants performance did not differ between monocular and stereo conditions. Therefore, participants performance are the same when the left text (real text) is fixed at near (2.0m) and right text(AR text) is presented at any of the three distances (near (0.67m), medium (2.0m), far (4.0m)) both in monocular and stereo conditions.

In addition, there is a significant effect of stereopsis on eye-fatigue at all three distances. Participants rated higher fatigue value in the monocular condition compared to the stereo condition (figure 4.13). One of the reasons could be that in monocular condition participants required to complete the task with the dominant eye, which gradually increased eye fatigue because it created additional eye strain only on the one eye. Whereas, in the stereo condition, eye pressure for the task was distributed between two eyes, which eventually resulted in less fatigue among the participants during the experiment. Further, in the post-interview session, all the participants mentioned that they were more comfortable and experienced less fatigue in stereo condition than monocular condition. Therefore, both subjectively as well as objectively, participants preferred stereo condition rather than monocular condition in this experiment.

CHAPTER 5

CONCLUSIONS

In an AR display system, two types of OST HMDs can be found: monocular (e.g., Google glass) display and stereo (e.g., Microsoft HoloLens) display. In both monocular and stereo displays, users need to integrate information both from the real world and virtual contents. As most of the current AR displays have a fixed focal plane, users need to switch from one particular distance to another for gathering information. Therefore, both context switching and focal distance switching are important issues in the current AR display interface design. To fully utilize the power of the AR system, it is essential to understand the effects of AR display context switching and focal distance switching on human performance. Although there has been dramatic progress of research done in AR system, very little empirical data currently exists on the effects of context switching and focal distance switching. The primary goal of this thesis is to replicate and extend the previous study of Gabbard et al. [19] by including the variable *stereopsis* (stereo, mono) and fully crossing the variables of context switching and focal distance switching. To achieve our goal we have divided our experiments and findings into three parts.

In the first part, our purpose is to replicate Gabbard et al. [19]'s task and experiment. The effects of context switching and focal distance switching indeed replicate in this monocular condition. Given the many differences between the Microvision Nomad display and the AR haploscope, this is consistent with the hypothesis that these findings broadly generalize to OST AR user interfaces [2]. These findings also lend further support to the primary finding of Gabbard et al. [19], that context switching and focal distance switching are important AR user interface design issues.

In the second part, we extend the previous study of Gabbard et al. [19] by considering a within-subject design for the stereo condition. Unlike the monocular condition, participants completed the experiment with two eyes. Except for this, experimental task, setup, variables, design, and procedure all are similar in both monocular and stereo condition. The findings of this stereo condition support the same negative effects of context switching and focal distance switching on human performance in the AR system. These findings further support that context switching and focal distance switching are important AR user interface design issues both in monocular and stereo displays. To the best of our knowledge, there is no research to date that has been conducted to empirically measure the effects of context switching and focal distance switching in stereo condition. Therefore, the empirical findings of this experimental part augment empirical data on the effects of focal disparity and context switching in AR.

In the final part, we are interested in comparing the effects of AR display context switching and focal distance switching on human performance between monocular and stereo conditions. For this, we considered the experimental data from the first part (monocular condition) and the second part (stereo condition). Results from the final part demonstrate that participants observed less fatigue in the stereo condition compared to the monocular condition at near (0.67m), medium (2.0m), and far (4.0m) distances. In addition, participants have better task performance in terms of completeness and accuracy in stereo condition rather than monocular condition. However, at medium (2.0m) distance, there is no difference in participants' task performance between monocular and stereo condition. To the best of our knowledge, this study is the first known empirical study that demonstrates the comparison of monocular and stereo AR systems in terms of the effects of context switching and focal distance switching on human performance.

Limitations Although our research successfully replicate and extend the previous study of Gabbard et al. [19], we have some certain limitations in our study. The limitations of our research are given below:

- AR haploscope is designed to overlay the graphical content in perpendicular direction only. For this reason, in our experiments, all the virtual textual information projected perpendicularly rather than any other directions. Generally, AR displays can overlay graphical contents in various directions in the real-world, which identifies as one of the limitations of our experiment.
- During the experiment, participants were directed to fix their heads by placing their chin in the chin-rest. Participants were only allowed to move their gaze during the experiment. However, in current AR displays, participants are allowed to move their heads as well as gaze.
- Participants with bi-focal corrective glasses/lenses found difficulties in performing the experiment. They needed to take off their glasses for certain distances, which distracts them and brings uncomfortable during the experiment.
- Sometimes participants found difficulties placing the numeric keypad in their convenient position, which eventually degraded their performance during the experiment.
- In the experiment, we provided one set of real-world textural information and one set of virtual textual information. However, real-world environment contains more than one information. In that case, the findings of our research are not applicable.

Future work Our research is one of the first steps of many human centered AR research studies in near future. In future, this research can be conducted and extended in various experimental conditions and directions. Some are listed below:

- One of the future works could be including the eye tracker to record the participants' gaze reaction time from the real world text to virtual text. We hypothesize that participants will have a slow reaction time when the eyes become tired and fatigued as the experiment progresses. In addition, as discussed in the experimental discussion section, participants began scanning the first line of text before their eyes had finished accommodating to the new focal distance. Future research is needed to verify this hypothesis with a binocular eye-tracker, which will indicate when observers shift gaze from the left to the right text, and therefore when they begin accommodating the new focal distance.

- One of the potential future research work could include the age effect in the analysis. Our results are based on a sample of young participants. As the accommodation ability decreases with increasing age [14], we can hypothesize that younger participants will have better performance with less eye fatigue than older participants considering the effects of AR display context switching and focal distance switching. Future studies need to be conducted to test this hypothesis.

- In our experiment, we present white textual information on a static black background. However, in reality, the real world consists of different colors, objects, shapes, and lighting conditions. In addition, real-world environment is complex and dynamic. Measuring participants' performance and eye-fatigue by replicating our experiment with various complex and dynamic backgrounds can be interesting future research.

- In our experiment, we have considered the subjective measurement method for evaluating the participant's eye fatigue. Lambooij et al. [40] recommended combining both objective as well as subjective measurement techniques to measure the degree of visual fatigue accurately, reliably, and validly. For this reason, in the future, objective measurement methods could be included in the experiment to measure the visual fatigue.

REFERENCES

[1] Apple, "Project Glass One day Google: Yep, we're testing augmented-reality glasses.,", url =https://www.youtube.com/watch?v=Vb2uojqKvFM, 2012, last checked: 2019-12-04.

[2] M. S. Arefin, N. Phillips, A. Plopski, J. L. Gabbard, and J. E. Swan II, "Impact of AR Display Context Switching and Focal Distance Switching on Human Performance: Replication on an AR Haploscope," *Abstracts and Workshops Proceedings, IEEE Conference on Virtual Reality and 3D User Interfaces (IEEE VR 2020)*. March 2020, IEEE Computer Society, in press.

[3] R. T. Azuma, "A Survey of Augmented Reality," *Presence: Teleoper. Virtual Environ.*, vol. 6, no. 4, Aug. 1997, pp. 355–385.

[4] M. L. Bernard, B. S. Chaparro, M. M. Mills, and C. G. Halcomb, "Comparing the Effects of Text Size and Format on the Readibility of Computer-Displayed Times New Roman and Arial Text," *Int. J. Hum.-Comput. Stud.*, vol. 59, no. 6, Dec. 2003, p. 823–835.

[5] M. Billinghurst and A. Duenser, "Augmented Reality in the Classroom," *Computer*, vol. 45, no. 7, July 2012, pp. 56–63.

[6] F. W. Campbell and G. Westheimer, "Dynamics of accommodation responses of the human eye," *The Journal of Physiology*, vol. 151, no. 2, 1960, pp. 285–295.

[7] T. P. Caudell and D. W. Mizell, "Augmented reality: an application of heads-up display technology to manual manufacturing processes," *Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences*, Jan 1992, vol. ii, pp. 659–669 vol.2.

[8] H. Cho, J. Wang, and S. Lee, "Text Image Deblurring Using Text-Specific Properties," *Computer Vision – ECCV 2012*, A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid, eds., Berlin, Heidelberg, 2012, pp. 524–537, Springer Berlin Heidelberg.

[9] S. Cuendet, Q. Bonnard, S. Do-Lenh, and P. Dillenbourg, "Designing augmented reality for the classroom," *Computers & Education*, vol. 68, 2013, pp. 557 – 569.

[10] J. E. Cutting and P. M. Vishton, "Chapter 3 - Perceiving Layout and Knowing Distances: The Integration, Relative Potency, and Contextual Use of Different Information about Depth," *Perception of Space and Motion*, W. Epstein and S. Rogers, eds., Handbook of Perception and Cognition, Academic Press, San Diego, 1995, pp. 69 – 117.

[11] S. Debernardis, M. Fiorentino, M. Gattullo, G. Monno, and A. E. Uva, "Text Readability in Head-Worn Displays: Color and Style Optimization in Video versus Optical See-Through Devices," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 1, Jan 2014, pp. 125–139.

[12] A. Dey, M. Billinghurst, R. W. Lindeman, and J. E. Swan II, "A systematic review of 10 years of augmented reality usability studies: 2005 to 2014.," *Front. Robotics and AI*, vol. 5, Jan. 2018, p. 161.

[13] A. Doshi, R. Smith, B. Thomas, and C. Bouras, "Use of projector based augmented reality to improve manual spot-welding precision and accuracy for automotive manufacturing," *The International Journal of Advanced Manufacturing Technology*, vol. 89, 03 2017, p. 1279–1293.

[14] A. Duane, "Normal values of the accommodation at all ages," *J. of the American Medical Association*, vol. LIX, 1912, pp. 1010–1013.

[15] F. Echtler, F. Sturm, K. Kindermann, and G. Klinker, "The Intelligent Welding Gun: Augmented Reality for Experimental Vehicle Construction," 01 2003.

[16] G. K. Edgar, J. C. D. Pope, and I. R. Craig, "Visual accommodation problems with head-up and helmet-mounted displays?," vol. 15, no. 2, 1994, pp. 68–75.

[17] A. Eiberger, P. O. Kristensson, S. Mayr, M. Kranz, and J. Grubert, "Effects of Depth Layer Switching between an Optical See-Through Head-Mounted Display and a Body-Proximate Display," *Symposium on Spatial User Interaction*, New York, NY, USA, 2019, SUI '19, Association for Computing Machinery.

[18] Federal-Aviation-Administration, "Character and Symbol Size, FAA Human Factors,", url =https://www.hf.faa.gov/webtraining/visualdisplays/text/size1a.htm, last checked: 2019-09-12.

[19] J. L. Gabbard, D. G. Mehra, and J. E. Swan II, "Effects of AR display context switching and focal distance switching on human performance," *IEEE transactions on visualization and computer graphics*, vol. 25, 2018, pp. 2228 – 2241.

[20] J. L. Gabbard, J. E. Swan II, D. Hix, S.-J. Kim, and G. Fitch, "Active Text Drawing Styles for Outdoor Augmented Reality: A User-Based Study and Design Implications," *Technical Papers, Proceedings of IEEE Virtual Reality 2007*, March 2007, pp. 35–42.

[21] J. L. Gabbard, J. E. Swan II, D. Hix, R. S. Schulman, J. Lucas, and D. Gupta, "An empirical user-based study of text drawing styles and outdoor background textures for augmented reality," *IEEE Virtual Reality (VR 2005)*, March 2005, pp. 11–18.

[22] H. Gross, F. Blechinger, and B. Achtner, *Human Eye*, John Wiley & Sons, Ltd, 2015.

[23] D. Gupta, *An Empirical Study of the Effects of Context-Switch, Object Distance, and Focus Depth on Human Performance in Augmented Reality*, master's thesis, Virginia Polytechnic Institute and State University, Blackburg, Virginia, USA, 2004.

[24] G. HERON, W. CHARMAN, and C. Schor, "Age Changes in the Interactions between the Accommodation and Vergence Systems," *Optometry and vision science : official publication of the American Academy of Optometry*, vol. 78, 11 2001, pp. 754–62.

[25] G. Heron, W. N. Charman, and L. S. Gray, "Accommodation Responses and Ageing," *Investigative Ophthalmology & Visual Science*, vol. 40, no. 12, 11 1999, pp. 2872–2883.

[26] M. Hirota, H. Kanda, T. Endo, T. Miyoshi, S. Miyagawa, Y. Hirohara, T. Yamaguchi, M. Saika, T. Morimoto, and T. Fujikado, "Comparison of Visual Fatigue caused by Head-Mounted Display for Virtual Reality and Two-Dimensional Display using Objective and Subjective Evaluation," *Ergonomics*, vol. 62, 02 2019, pp. 1–33.

[27] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, "Vergence-accommodation conflicts hinder visual performance and cause visual fatigue," *Journal of Vision*, vol. 8, no. 3, Jan. 2008, pp. 1–30.

[28] A. Huckauf, M. H. Urbina, J. Grubert, I. Böckelmann, F. Doil, L. Schega, J. Tümler, and R. Mecke, "Perceptual issues in optical-see-through displays," *Proceedings of the ACM Symposium on Applied Perception in Graphics and Visualization*. ACM, 2010, pp. 41–48.

[29] J. H. Iavecchia, H. P. Iavecchia, and I. Stanley N Roscoe, "Eye accommodation to head-up virtual images," *Human Factors*, vol. 30, no. 6, 1988-01-01, pp. 703–712.

[30] Interaction-Design-Foundation, "Augmented Reality – The Past, The Present and The Future,", url = https://www.interaction-design.org/literature/article/augmented-reality-the-past-the-present-and-the-future, last checked: 2020-02-09.

[31] J. Jankowski, K. Samp, I. Irzynska, M. Jozwowicz, and S. Decker, "Integrating Text with Video and 3D Graphics: The Effects of Text Drawing Styles on Text Readability," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2010, CHI '10, p. 1321–1330, Association for Computing Machinery.

[32] W. Jaschinski, M. Bonacker, and E. Alshuth, "Accommodation, convergence, pupil diameter and eye blinks at a CRT display flickering near fusion limit," *Ergonomics*, vol. 39, 02 1996, pp. 152–64.

[33] S. Kasthurirangan and A. Glasser, "Age related changes in accommodative dynamics in humans," *Vision Research*, vol. 46, 2006, pp. 1507–1519.

[34] H. Kato and M. Billinghurst, "Marker tracking and HMD calibration for a video-based augmented reality conferencing system," *Proceedings 2nd IEEE and ACM International Workshop on Augmented Reality (IWAR'99)*, Oct 1999, pp. 85–94.

[35] H. Kato, H. Ito, J. Shima, M. Imaizumi, and H. Shibata, "Development of Hologram Head-Up Display," *SAE Technical Paper*. 02 1992, SAE International.

[36] K. Kim, A. Erickson, A. Lambert, G. Bruder, and G. Welch, "Effects of Dark Mode on Visual Fatigue and Acuity in Optical See-Through Head-Mounted Displays," *Symposium on Spatial User Interaction*, New York, NY, USA, 2019, SUI '19, pp. 9:1–9:9, ACM.

[37] G. Kramida, "Resolving the Vergence-Accommodation Conflict in Head-Mounted Displays," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 7, July 2016, pp. 1912–1931.

[38] B. C. Kress and W. J. Cummings, "Invited paper: Towards the ultimate mixed reality experience: HoloLens display architecture choices," *SID Symposium Digest of Technical Papers*, vol. 48, no. 1, 2017, pp. 127–131.

[39] E. Kruijff, J. E. Swan II, and S. Feiner, "Perceptual issues in augmented reality revisited," *IEEE International Symposium on Mixed and Augmented Reality (ISMAR 2010)*. Oct. 2010, pp. 3–12, IEEE.

[40] M. Lambooij, M. Fortuin, I. Heynderickx, and W. IJsselsteijn, "Visual discomfort and visual fatigue of stereoscopic displays: a review," *Journal of Imaging Science and Technology*, vol. 53, no. 3, Jan. 2009, pp. 30201–1–30201–14.

[41] R. S. Laramee and C. Ware, "Rivalry and Interference with a Head-Mounted Display," *ACM Trans. Comput.-Hum. Interact.*, vol. 9, no. 3, Sept. 2002, p. 238–251.

[42] T. Lockhart and W. Shi, "Effects of Age on Dynamic Accommodation," *Ergonomics*, vol. 53, 07 2010, pp. 892–903.

[43] K. J. MacKenzie, R. A. Dickson, and S. J. Watt, "Vergence and accommodation to multiple-image-plane stereoscopic displays: "real world" responses with practical image-plane separations?," vol. 21, no. 1, 2012, pp. 011002–1.

[44] K. J. MacKenzie, D. M. Hoffman, and S. J. Watt, "Accommodation to multiple-focal-plane displays: Implications for improving stereoscopic displays and for accommodation control," *Journal of Vision*, vol. 10, no. 8, 07 2010, pp. 22–22.

[45] P. Milgram and F. Kishino, "A Taxonomy of Mixed Reality Visual Displays," *IEICE Trans. Information Systems*, vol. E77-D, no. 12, 12 1994, pp. 1321–1329.

[46] R. J. Miller, R. G. Pigion, M. F. Wesner, and J. G. Patterson, "Accommodation fatigue and dark focus: The effects of accommodation-free visual work as assessed by two psychophysical methods," vol. 34, no. 6, 1983, pp. 532–540.

[47] A. Z. Mohamad Ali, R. Wahid, K. Samsudin, and M. Z. Idris, "Reading on the Computer Screen: Does Font Type has Effects on Web Text Readability?," *International Education Studies*, vol. 6, 03 2013, pp. 26–35.

[48] M. Mon-Williams and J. Wann, "Binocular Virtual Reality Displays: When Problems Do and Don't Occur," *Human Factors*, vol. 40, 03 1998, pp. 42–49.

[49] K. R. Moser, M. S. Arefin, and J. Edward Swan, "Impact of Alignment Point Distance Distribution on SPAAM Calibration of Optical See-Through Head-Mounted Displays," *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, March 2018, pp. 641–642.

[50] K. R. Moser, M. S. Arefin, and J. E. Swan, "Impact of Alignment Point Distance and Posture on SPAAM Calibration of Optical See-Through Head-Mounted Displays," *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Oct 2018, pp. 21–30.

[51] W. Narzt, G. Pomberger, A. Ferscha, D. Kolb, R. Müller, J. Wieghardt, H. Hörtner, and C. Lindinger, "Augmented reality navigation systems," *Universal Access in the Information Society*, vol. 4, 03 2006, pp. 177–187.

[52] C. Neveu, T. Blackmon, and L. Stark, "Evaluation of the effects of a head-mounted display on ocular accommodation," *Presence: Teleoperators and Virtual Environments*, vol. 7, no. 3, 1998, pp. 278–289.

[53] S. Okabayashi, M. Sakata, and T. Hatada, "How an automotive head-up display affects a driver's ability to recognize the forward view," *Proceedings of the society for information display*, 1991.

[54] D. A. Owens and K. Wolf-Kelly, "Near work, visual fatigue, and variations of oculomotor tonus.," vol. 28, no. 4, 1987, pp. 743–749.

[55] R. Patterson, M. D. Winterbottom, and B. J. Pierce, "Perceptual Issues in the Use of Head-Mounted Visual Displays," *Human Factors*, vol. 48, no. 3, 2006, pp. 555–573, PMID: 17063969.

[56] N. Phillips, K. Massey, M. S. Arefin, and J. E. Swan II, "Design and calibration of an augmented reality haploscope," *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, Munich, Germany, Oct. 2018, pp. 75–76, IEEE.

[57] N. Phillips, K. Massey, M. S. Arefin, and J. E. Swan II, "Design, assembly, calibration, and measurement of an augmented reality haploscope," *Proceedings of PERCAR: The Fifth IEEE VR Workshop on Perceptual and Cognitive Issues in AR, IEEE Conference on Virtual Reality and 3D User Interfaces (VR 2019)*, Osaka, Japan, Mar. 2019, pp. 1770–1774, IEEE.

[58] G. D. Porta, "De refractione optices parte: libri novem," 1593.

[59] R. Ratcliff, A. Thapar, and G. McKoon, "Aging, Practice, and Perceptual Tasks: A Diffusion Model Analysis," *Psychology and aging*, vol. 21, 07 2006, pp. 353–71.

[60] H. Regenbrecht, G. Baratoff, and W. Wilke, "Augmented reality projects in the automotive and aerospace industries," *IEEE Computer Graphics And Applications*, vol. 25, no. 6, 2005, pp. 48–56.

[61] U. Rehman and S. Cao, "Augmented-Reality-Based Indoor Navigation: A Comparative Analysis of Handheld Devices Versus Google Glass," *IEEE Transactions on Human-Machine Systems*, vol. 47, no. 1, Feb 2017, pp. 140–151.

[62] B. Schwerdtfeger, R. Reif, G. Klinker, D. Hamacher, L. Schega, and F. Doil, "Pick-by-vision: a first stress test," *2009 8th IEEE International Symposium on Mixed and Augmented Reality*. 2009, pp. 115–124, IEEE.

[63] G. Singh, S. R. Ellis, and J. E. Swan II, "The effect of focal distance, age, and brightness on near-field augmented reality depth matching," *IEEE Transactions on Visualization and Computer Graphics*, 2018, pp. 1–14.

[64] M. B. Spitzer, X. Miao, and B. Amirparviz, "Method and apparatus for a near-to-eye display,".

[65] L. Stark, Y. Takahashi, and G. Zames, "Nonlinear Servoanalysis of Human Lens Accommodation," *IEEE Transactions on Systems Science and Cybernetics*, vol. 1, no. 1, 1965, pp. 75–83.

[66] I. E. Sutherland, "A Head-Mounted Three Dimensional Display," *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I*, New York, NY, USA, 1968, AFIPS '68 (Fall, part I), p. 757–764, Association for Computing Machinery.

[67] D. R. Tufano, "Automotive HUDs: The Overlooked Safety Issues," *Human Factors*, vol. 39, no. 2, 1997, pp. 303–311, PMID: 9302890.

[68] K. Ukai and P. A. Howarth, "Visual fatigue caused by viewing stereoscopic motion images: Background, theories, and observations," *Displays*, vol. 29, no. 2, 2008, pp. 106–116.

[69] D. Van Krevelen and R. Poelman, "A Survey of Augmented Reality Technologies, Applications and Limitations," *International Journal of Virtual Reality*, vol. 9, no. 2, 2010, pp. 1–20.

[70] R. W. West and D. A. Goss, *Introduction to the Optics of the Eye*, Butterworth-Heinemann, 2001.