Perceptual Manipulations for Hiding Image Transformations in Virtual Reality

Maryam Keyvanara

A DISSERTATION SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE PROGRAMME
IN ELECTRICAL ENGINEERING AND COMPUTER SCIENCE
YORK UNIVERSITY
TORONTO, CANADA

July 2022

Abstract

Users of a virtual reality make frequent gaze shifts and head movements to explore their surrounding environment. Saccades are rapid, ballistic, conjugate eye movements that reposition our gaze, and in doing so create large-field motion on our retina. Due to the high speed motion on the retina during saccades, the brain suppresses the visual signals from the eye, a perceptual phenomenon known as the saccadic suppression. These moments of visual blindness can help hide the display graphical updates in a virtual reality.

In this dissertation, I investigated how the visibility of various image transformations differed, during combinations of saccade and head rotation conditions. Additionally, I studied how hand and gaze interaction, affected image change discrimination in an inattentional blindness task. I conducted four psychophysical experiments in desktop or head-mounted VR. In the eye tracking studies, users viewed 3D scenes, and were triggered to make a vertical or horizontal saccade. During the saccade an instantaneous translation or rotation was applied to the virtual camera used to render the scene. Participants were required to indicate the direction of these transitions after each trial. The results showed that type and size of the image

transformation affected change detectability. During horizontal or vertical saccades, rotations along the roll axis were the most detectable, while horizontal and vertical translations were least noticed. In a second similar study, I added a constant camera motion to simulate a head rotation, and in a third study, I compared active head rotation with a simulated rotation or a static head. I found less sensitivity to transsaccadic horizontal compared to vertical camera shifts during simulated or real head pan. Conversely, during simulated or real head tilt observers were less sensitive to transsaccadic vertical than horizontal camera shifts. In addition, in my multi-interactive inattentional blindness experiment, I compared sensitivity to sudden image transformations when a participant used their hand and gaze to move and watch an object, to when they only watched it move. The results confirmed that when involved in a primary task that requires focus and attention with two interaction modalities (gaze and hand), a visual stimuli can better be hidden than when only one sense (vision) is involved.

Understanding the effect of continuous head movement and attention on the visibility of a sudden transsaccadic change can help optimize the visual performance of gaze-contingent displays and improve user experience. Perceptually suppressed rotations or translations can be used to introduce imperceptible changes in virtual camera pose in applications such as networked gaming, collaborative virtual reality and redirected walking. This dissertation suggests that such transformations can be more effective and more substantial during active or passive head motion. Moreover, inattentional blindness during an attention-demanding task provides additional opportunities for imperceptible updates to a visual display.

To Ali, Rose, and my mum and dad.

Acknowledgements

Thank you to my PhD supervisor, Professor Rob Allison, for his constant support during my years at his lab. Thank you for trusting me and having me in his lab, and guiding me through becoming a professional. This research would not be possible without his constant support and advise. I would like to thank my committee members Professor Matthew Kyan and Melanie Baljko, for all their great feedback and encouragements throughout the past years. They were always available to answer any questions I had. Thank you Matt for being an amazing Chair for my PhD oral exam. Thanks to my examining committee, Gerd Grau, Jacob Beck, and my external examiner Gabriel Diaz, who all took the time out their busy schedule to attend my oral exam. Their feedback has been very valuable and appreciated. Thank you Professor Minas Spetsakis for trusting me to be your Teaching Assistant for so many courses. I learnt so much from you. Finally, thank you my lab mates who were all always so friendly and supportive.

Thank you my amazing husband, Ali, who I can know I can always rely on. He supported me during along the way with all the ups and downs. Thank you my wonderful parents. They always believed in me, supported me and encouraged me to achieve my best. Thank you my two sisters. Finally, thank you my daughter, Rose, for coming to my world, and being a part of this PhD journey.

Thank you all for your support along this long way. I could not have been here without your help.

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

Human gaze has compelling features for interaction and can reveal a lot about a person's interests, intentions, and actions. We use our eyes to look at objects and to communicate non-verbally. Gaze direction is a reliable indicator of a person's visual attention and interest as people often look at what interests them (overt attention), compared to keeping it in peripheral vision (covert attention) (Navalpakkam and Churchill, 2014). The observation of eye movements and patterns allows researchers to identify the point of regard, blink characteristics, pupil motion and reaction to their stimuli. With this information, developers are able to produce applications that can be used in many domains (Hansen and Ji, 2009; Ruhland et al., 2015; Young and Sheena, 1975). Eye tracking, the technology that enables a device to measure eye position and eye movement, has been employed in many research and industrial fields for many years. Eye tracking provides enhanced means of interaction with VR content, making it more natural and intuitive. This technology can be used in VR environments and immersive systems such as the CAVE, head-mounted displays (HMDs), 3D Games and assistive applications. (Tanriverdi and Jacob, 2000; Strandvall, 2009; Piumsomboon et al., 2017). It is also an effective tool to study different aspects of visual perception and quantify the dynamics of eye movements in visual behavior. This enables understanding eye movements in natural tasks and reveals a user's cognitive and perceptual processes and limitations. Saccades are one of the most common eye movements we make. They are rapid, ballistic, conjugate eye movements that reposition our gaze, and create large-field motion on our retina (Bahill et al., 1975; Liversedge and Findlay, 2000). They are rapid with peak velocity up to 1000°/s, and because of their high speed motion, the brain suppresses the visual signals from the eye. This is a perceptual phenomena known as the saccadic suppression, and has been studied widely (Bridgeman et al., 1975; Ziat et al., 2010). Leveraging this phenomenon, the details of a visual display can be modified during such blind phases of saccadic eye movements. In such moments, the users of a VR will often be unaware of any changes that occur in the display, even if very important details are changed (Bohrn et al., 2010; Born, 2019; Allison et al., 2010). Saccadic suppression has gained a lot of attention in Virtual Reality applications, as it can allow a VR system to make modifications to its graphical display without the user noticing (Sun et al., 2018; Patney, 2017; Bolte and Lappe, 2015; Schumacher et al., 2004).

Virtual reality (VR) is a technology aiming to provide an interactive three-dimensional virtual environment. It relies on high-quality computer graphics system and other external display and control interfaces, to create an immersive experience for the users. Immersive virtual reality content is advancing rapidly, but still need improvements to create more natural and immersive environment for the users. Insufficient graphical performance and high computations is another challenge in VR systems. These computer graphic systems intend to create a high quality user experience that is perceptually realistic. There are several components that can help achieve this goal. These virtual reality contents must use stereoscopic display and be interactive, rendered with high-resolution and realistic three-dimensional images, computationally efficient, have a low latency, and work with the human visual system. Human senses such as vision, hearing and touch determine a person's interpretation of their surrounding environment. Human perception aims at strengthening sensory inputs so that they can be perceived and acted upon quickly and efficiently (Carbon, 2014). Most VR systems today maintain such features and try to create systems that match visual perceptual capabilities. Focus cues such as accommodation and retinal blur have also been considered in the design of some VR systems (Elias et al., 2019; Zhdanov et al., 2019; Konrad et al., 2020). Our visual perception has other interesting features that can be used to improve a VR experience, such as saccadic suppression, that allows interaction with VR systems.

Gaze-based interaction in virtual environments is becoming very popular in many commercial, gaming and educational fields. For example in gaming, the players' eyes can be used to interact with the 3D virtual environment (Bombari et al., 2015). Furthermore, in a VR, human gaze can be used for pointing and selection purposes (Sidenmark and Gellersen, 2019), to help people with disabilities (Karlsson et al., 2018), to improve remote collaborations (Luxenburger et al., 2016), and also to build gaze-contingent displays. Eye gaze can provide better interaction when a user's hands are fully occupied or even disabled, and can replace other input modalities. There are even more applications of eye tracking in VR such as more human-like social avatars and narratives (Nilsson et al., 2009; Da Silva et al., 2007; Lahiri et al., 2011), biometric identifiers and gaze-based authentication systems (Chappell, 2015; Luo et al., 2020) and virtual item selection or inputting (Luxenburger et al., 2016). Eye tracking has been used as a tool for investigating behavior and to better understand eye movements and eye problems. It has also been used as a means to diagnose eye-related clinical problems. Eye movements deficits such as nystagmus and strabismus, eye movement control and tracking difficulties have been studied through eye tracking devices (Kasneci et al., 2017). In addition, eye tracking has been used to detect brain disorders such as autism, specifically when it is used in conjunction with fMRI (Boraston and Blakemore, 2007). Moreover, eye tracking analytics can show insights about user attention that can help developers in having a deep understanding of their application feedback. It shows what parts of an application users are more focused on, and which parts may elicit negative responses from them. These analytics can have important influence on how software or VR content are designed, and hence shape a much better experience for the users.

Gaze-Contingent Displays (GCD) are dynamic displays that manipulate the display content based on the user's real-time eye movements (Reder, 1973; Duchowski et al., 2004). They were initially introduced to stabilize images on the retina (Riggs et al., 1954; Yarbus, 1967). The purpose of many of these displays is to balance the amount of information displayed against the visual information processing capacity of the observer by relying on real-time eye movement sensing. Modern GCDs tend to improve accuracy and visibility of the object at the point of regard and balance the displayed information with the viewer's visual information processing capacity (Duchowski et al., 2004). GCDs can use eye movements features, such as fixation durations, to provide information for tasks such as reading or visual search (McConkie and Rayner, 1975; Loschky and McConkie, 2002).

Moreover, there are many applications for eye tracking in virtual and augmented reality. Incorporating cameras into VR headsets has become more affordable and feasible in the recent years, making eye tracking in VR more accessible for everyone. Eye tracking applications in virtual environments include foveated depth of field, gaze interaction and gaze navigation. By enabling natural interactions through gaze, eye tracking contributes to more immersive and user-friendly experiences in VR. It can also be used for gaze prioritized graphics also known as Foveated Rendering. Having a virtual reality HMD or desktop equipped with an eye tracker enables us to understand where a user is looking. HMDs can direct high-definition graphics processing power to that exact spot in real time. This enables higher definition displays, more efficient devices, longer battery life, or increased mobility. Eye tracking in a VR headset can further be used for calculating the users' interpupillary distance (IPD) and therefore automatically adapt to each user. IPD is an important measure for each individual who wears and uses a headset. This is because IPD is required to move the lenses and displays into the optimal position to provide comfort and visual quality and appropriate binocular disparities. During longer periods of time, this can reduce eye strain by allowing for more natural vergence eye movement. However, this feature needs to have high quality and expensive lenses.

The motivation for this PhD research was to explore some of the visual perception phenomena that can be used in improving virtual and augmented reality technology and user experience. I was specifically interested in the utility of perceptual saccadic suppression in VR eye tracking applications. With wider and more affordable availability of eye tracking in VR, AR and desktop computing the requirement to improve their quality and speed is very important. The interdisciplinary idea of using human visual perception in improving such content is very interesting and a strong motivation in contributing to this field. In a virtual environment users not only make many eye movements, but also a lot of head movements. Our saccadic eye movements may be combined with head movements, or without them. In a natural immersive VR experience, we expect users to make several of these movements. Studying saccadic suppression during such eye and head movements can provide more insight into higher quality design of a VR experience for users. With a restrained head, saccades can be characterized by the relationships between the amplitude, peak velocity and duration. With a freely moving head, the changes in the direction of the line of sight (gaze shifts) often involve saccades that have simultaneous head movements (Freedman and Sparks, 2000; Freedman and Sparks, 1997; Berencsi et al., 2005). Interaction in VR is not limited to eyes and head, and may involve a hand movement, for interaction, object manipulation, or assisting users with disabilities or rehabilitation purposes. This can be achieved through the use of controllers (Fahmi et al., 2020; Khundam et al., 2021), VR gloves (Kim et al., 2020; Connelly et al., 2010) rigid body wrist attachments (Han et al., 2018) or computer vision.

1.2 Research Objectives

The focus of this research is gaze-contingent displays and how saccadic suppression can be leveraged to improve a user's virtual reality experience. This improvement can be in reducing the VR system's latency and computational and graphical load, hence providing effectively higher quality VR content and enabling applications such as redirected walking (Sun et al., 2020; Langbehn et al., 2016; Schumacher et al., 2004). I am interested in working with the natural human eye movements when viewing virtual environments. Duchowski categorizes interactive eye tracking systems into selective and gaze-contingent applications (Duchowski and Çöltekin, 2007; Duchowski, 2007). In this research, we focus on gaze-contingent applications where we enable display updates according to the users' gaze behaviour, and provide an interaction with them according to their eye movements and their point of regard. To this end, I designed and conducted a series of VR experiments, using a precise eye tracker with a desktop display and an HMD with integrated eye trackers to analyse the research hypotheses. In addition, I would like to be able to provide natural viewing and navigation in a VR and increase the sense of presence and immersion for the users, by applying subtle undetectable changes inside the user's field of view and study the variables that affect the sensitivity and tolerance of the users of a VR environment to these changes occurring during saccade. In a VR, we make many head rotations and translations to view our surroundings. To increase the generalizability of my results, I measured the sensitivity to modifications in the user's viewpoint change when saccades were generated while moving the head. In addition, in an interactive VR, we may use other senses than just vision. That is why I studied how a hand movement during gaze can affect this sensitivity to motions applied to the whole image.

1.3 Contribution

The most significant contribution of this dissertation is the analysis of various perceptual methods that can be effective in hiding graphical display updates, namely image transformations, from VR users and modifying their viewpoint without them noticing.

The combinations of different saccade directions, image transformations, and virtual camera angles have not been studied in depth before. Also, in real context, real motion or locomotion is very common, therefore comparing the effect of an added real or simulated head rotation during saccades, using various types of setup is another important contribution. I found the head rotation affects the amount of suppression during saccades during both a user's active head rotation and a simulated rotation of the virtual camera. Furthermore, I found that multimodal interaction in VR during an inattentional blindness task, modulates the detectability of image changes during hand-eye coordination in VR. These are important as users make many eye, head and hand movements in VR to explore it. These findings can help improve design of VR applications such as redirection and reorientation of users during locomotion or navigational tasks. Knowing which image transformations, and saccade or head rotation directions are more likely to mask changes in the viewpoint of the display, can improve and optimize a more natural navigation for VR users in small physical spaces. Knowing that when users are involved in a task that engages more than their vision (e.g. their hand), can help in creating content that can trigger such tasks, or when such tasks occur apply the image changes to the display. Furthermore, these studies build on previous work that used 2D scene changes, and will help inform models and theories of scene perception in real and virtual worlds.

1.4 Related Publications

The following is a list of published research articles related to the research of this dissertation. Both authors contributed in the idea of the research study, design and implementation, data collection and analysis, and writing of the papers. In addition, the supervision and guidance of the project, supply of the lab equipment and the financial funding for the research project were supported by Prof. Allison.

- Keyvanara, Maryam, and Robert S. Allison. "Effect of a Constant Camera Rotation on the Visibility of Transsaccadic Camera Shifts", ACM Symposium on Eye Tracking Research and Applications (ETRA), June 2020, Stuttgart, Germany.

- Keyvanara, Maryam, and Robert S. Allison. "Transsaccadic Awareness of Scene Transformations in a 3D Virtual Environment", ACM Symposium on

Applied Perception (SAP), Sept 2019, Barcelona, Spain.

- Keyvanara, Maryam, and Robert S. Allison. "Viewers' Sensitivity to Camera Motion during Saccades in a Virtual Environment", European Conference on Eye Movement (ECEM), Aug 2019, Alicante, Spain. (Abstract)

- Keyvanara, Maryam, and Robert S. Allison. "Saccadic Suppression of Natural Image Transformations.", European Conference on Visual Perception (ECVP), Aug 2018, Trieste, Italy. (Abstract)

- Keyvanara, Maryam, and Robert S. Allison. "Sensitivity to Natural 3D Image Transformations During Eye Movements.", Proceedings of the 2018 ACM Symposium on Eye Tracking Research and Applications, June 2018, Warsaw, Poland.

1.5 Outline

This dissertation is organized in five chapters. The current chapter was an introduction to the topics covered in the dissertation. It discussed the motivation for this research, the objectives and also my contribution to this topic area. In Chapter 2, I review fundamental background science, and recent related research papers about human vision and eye movements, gaze-contingent displays, and related topics. Chapter 3 describes two experiments I ran on a desktop-based eye tracker. These were non-immersive VR experiments that were inspired by the saccadic suppression of image displacement. Chapter 4 provides the details of two other experiments I conducted using the HTC Vive Pro Eye headset. During these studies, participants had full immersion, as we used a head-mounted display and collected gaze data in 3d space. These experiments were more interactive and included controlled head rotations and hand movements. Finally, Chapter 5 is a conclusion of the topics discussed in this dissertation, plus the future related works.

Chapter 2

Human Vision and Eye Movements

In this chapter I discuss relevant background including human visual perception, saccadic suppression, gaze-contingent displays, inattentional blindness and other related topics in virtual reality.

2.1 Human Vision and Eye Movements

Our vision is one of our main senses for collecting information from our environment. Human vision is not only an optical process, but also an active cognitive dynamic process, in which a viewer searches for visual information to support cognitive and behavioral activities (Gibson, 1950). Through visual perception, we see the world around us and acquire knowledge about environmental objects and events. Our vision is also important in guiding our motor control and provides information for perception of self-motion in the environment (Abbasov, 2019; Findlay et al., 2003; Lappe et al., 1999).

The densely packed cone cells in and around our fovea allow sharp vision with high-resolution to discern finer details and are sensitive to colors. Visual acuity declines quickly and continuously with retinal eccentricity towards the periphery (Henderson, 2007). Our peripheral vision spreads away from the center and can help the eye focus its attention on important parts of the scene. It is mainly composed of the rod cells which are not reliant on bright-light situations and can notice movement better, enabling us to have a high ability to detect motion in our peripheral vision. Foveal and peripheral vision are known to have different functions and information processing. Studies have shown that humans trust the information from their central vision more than the information inferred from the periphery of the visual field (Gloriani and Schütz, 2019). Also, reaching to the peripheral visual field requires a more extensive cortical network engagement than reaching to the central visual field (Prado et al., 2005). Other studies also confirm the properties of the visual system, specifically the limited spatial acuity and high motion sensitivity of peripheral vision. They also found that participants used peripheral vision to monitor the changes in targets and, additionally, to perceive changes at both near and far eccentricities (Vater et al., 2017).

2.1.1 Types of Eye Movements

We make a variety of different types of eye movements, that play a big role in our visual perception, and help us acquire visual information (Rucci and Victor, 2015). Some eye movements can be gaze shifting in which they redirect the fovea to a new point of interest or track moving targets, while others serve as a gaze stabilizing mechanisms and help to keep images steady on the retina (Leigh and Zee, 2015). Normally we can move our eyes with three degrees of freedom: horizontally, vertically and torsionally. The torsional movements are rotations of the eye around the line of sight. In the movements in this direction, the gaze direction does not change (Duchowski and Duchowski, 2017). In this section, I will briefly discuss different types of eye movements.

Fixations

Fixations are the short periods of eye fixating and acquiring visual information. During fixations the image on the retina moves very little. Fixational eye movements are small eye movements, that occur during fixations, and can fixate on an object of interest (Rucci and Victor, 2015). Fixational eye movements take most of our viewing time. They are an indication of a person's attention, and are important in perceptual and cognitive activities. Fixation patterns of individuals vary depending on factors such as age, cognitive functioning, task, and genetics. Analysis of these patterns has been useful in many applications and research studies. For example fixation patterns can be useful in determining the physiological or psychological state of a person (Martinez-Conde, 2006).

In an eye-tracker enabled VR setting, eye fixations can provide fundamental data for evaluation of eye behaviour and a user's performance and attention. Fixations can be measured spatially by their central tendency and dispersion, indicating where and how stably the eye fixates (in pixels or degrees of visual angle). They can also be measured temporally by their duration. The average duration of a fixation is 300 ms, with a range of 100 to 600 ms. Longer duration of fixations are related to more engagement with the target object/scene or an increased cognitive process (Negi and

Mitra, 2020; Just and Carpenter, 1976; Chien et al., 2015). In addition, the visual saliency information obtained from eye fixations can be used in identifying points of interest. Fixational eye movements are particularly useful in extracting saliency maps and identifying points of interest and users' attention. In another study, fixation prediction networks were developed that predicted the viewer's future fixations. They used sensor related features including HMD orientations, and content related features which are the saliency maps of images and motion maps (Nguyen et al., 2018b). Computational models of visual attention have been employed in a method to predict eye fixation locations as saliency maps (instead of directly extracting features from fixation data) (Volokitin et al., 2016; Eizenman et al., 1984).

Researchers have mentioned that the eye is never perfectly still, and agree that during visual fixation we have three main types of eye movement occurance, including microsaccades, drifts and tremors (Martinez-Conde et al., 2004; Helmholtz, 1985). The small fixational rapid eye movements called micro-saccades, occur when we are fixating our eyes on a target. They are spatially random with amplitudes of over 1 to 2 minutes of arc but less than (Hubel, 1995). During these fixational eye movements, the eye has low amplitude and low velocity motion. Drifts, which are slow random motion of the eye away from a fixation point, are also in the category of fixational eye movements. Drifts occur only within the foveal dead zone and have very low velocities of a few minutes of arc per second (Young and Sheena, 1975). Tremor is an 'aperiodic, wave-like motion' of the eyes. They have a very small amplitude (similar to system noise) and are hard to record accurately (Martinez-Conde et al., 2004).

Saccades

When we visually explore our visual environment, both of our eyes jump simultaneously from one fixation to another about three to five times per second. These very rapid conjugate eye movements are called saccades, and redirect the foveal region to a new area of interest. As discussed, visual information is picked up in the brief fixation periods between saccades and creates a loop of perception and action (Duchowski, 2007; Crevecoeur and Kording, 2017; Rucci and Poletti, 2015). For a saccade to take place, the eyes take off from one position, rapidly accelerate and reach a peak velocity early in the saccade. They decelerate and reach the visual target (Bittencourt et al., 1981). Saccadic eye movements provide us with successive sampling of the environment around us and are important part of our visual behaviour. Saccades are potentially very disruptive as they introduce rapid motions of the image on the retinas that should be ignored or cancelled in a normal situation. The cancellation of these events is a perceptual phenomena known as saccadic suppression, and will be discussed in a later section. Saccades can occur voluntary and reflexively. Voluntary saccades are self-paced eye movements which are made consciously to gather information about the environment and explore the visual scene. They are directed at our will, and are triggered by objects seen, heard,remembered, imagined or from memory or even as part of a strategy to scan the visual scene (Yarbus, 1967; Walker et al., 2000). The voluntary saccade network is believed to include pathways from frontal cortex to superior colliculus and the brainstem (Munoz, 2002). On the other hand, reactive saccades are triggered to move the gaze as fast as possible to a new target of interest which may be potentially threatening or interesting. In research studies, these saccades have gained more attention because they can more easily be provoked in controlled conditions and elicited with reliable timing. The neural network of reactive saccades includes parietal pathways to the superior colliculus and the brainstem saccade generator (Grossberg et al., 1997; Gancarz and Grossberg, 1999; Walker et al., 2000).

During typical eye movement studies, saccades are measured and their parameters are monitored. Some of the important parameters that are measured include latency, amplitude, velocity, acceleration and duration. The event details of a saccade such as the temporal metrics, when the saccade occurred (time stamps in ms), and spatial metrics, where it occurred (pixel or angular position x and y coordinates), can be obtained through an eye tracking device.

Saccade latency is the amount of time the human central nervous system takes to respond with a reactive saccade. Young and Sheena (1975) reported latency of 100 to 300 ms, while Yang et al. (2002). found a latency of 200-250 ms for a normal adult human. During the latency period, several events such as shift of visual attention to the new target, disengagement of ocular fixation, and computation of the metrics of the movements occur in the brain and cause the delay (Young and Sheena, 1975; Yang et al., 2002). Saccade latency seems to decrease with age during childhood. Latencies of visually guided saccades and vergence eye movements were longer in children than in adults (Yang et al., 2002). There is also a minimum delay, also known as the refractory period, of 100-200 ms between consecutive saccadic eye movements (Young and Sheena, 1975).

The amplitude (also called the size) of a saccade is the angular distance the eye travels during the movement. The larger the amplitude of a saccade, the smaller is the overlap between the information processed during the two fixations before and after the saccade (Unema et al., 2005). Often when the gaze shift exceeds 30°, head motion is involved. The shape of the velocity profiles of a saccade varies as a function of the amplitudes of saccades. For amplitudes of up to 15° or 20° the velocity of a saccade depends linearly on the amplitude. In addition, duration of saccades ranges between 10 ms to 120 ms, and it also varies with saccade magnitude: larger saccades have a larger duration (Duchowski and Çöltekin, 2007; Duchowski and Duchowski, 2017). Research has shown that parameters such as horizontal saccade frequency, latency and peak velocities are unaffected by age up to the sixth decade, and after that age the frequency and peak velocities of saccades have a slight increase, while the latency remains the same (Anson et al., 2016; Janky et al., 2018). Furthermore, the amplitude of the saccades highly depends on the task. Smaller saccades may last only 30 ms and are harder to detect (Bahill et al., 1975; Gibaldi and Banks, 2019).

Other Types of Eye Movements

In addition to the eye movements discussed, there are other types including smooth pursuits, optokinetic, vestibulation, vergence and accommodation eye movements. In real life situations, when we observe a dynamic scene, we automatically follow and track moving objects to hold the visual stimuli steady on the fovea. In smooth pursuit eye movements, the eyes move smoothly, instead of in jumps, to maintain a continuously stable view of the target object. These movements will allow us to resolve details of the moving objects. Vergence are movements of both eyes in the opposite direction (disjunctive), often with the purpose of fusing the images of near or far objects. The binocular eye convergence (inward rotation towards each other) and divergence (outward rotation away from each other), ensures that both right and left images of an object can be projected on similar locations on the two retinas. These movements are executed more slowly and smoothly than conjugate eye movements and have a shorter latency (approximately 160-180 ms) in normal adults, compared to saccades (Yang et al., 2002). Lastly, accommodation is an oculomotor process, in which the eye changes its focal power to ensure that the fixated object is focused clearly on the retina. As objects come near the eye, the lens of the eye focuses on the near object and the background becomes progressively less focused and more blurry.

2.1.2 Eye Movements During Head Motion

A change in gaze from one point of interest to another is accompanied by coordinated eye and head movements, specifically during daily activities such as walking, running and exploring our surroundings (Grossman et al., 1988). Such activities also generate unplanned head movements. Additionally, there is a constant interaction between the vestibular system and the human visual system. The vestibular system in the inner ears can detect head motion, and provide a perception of head motion. Our brain uses these vestibular signals to produce eye movements that will keep the gaze stabilized. The function of the vestibulo-ocular reflex (VOR) is to rapidly stabilize the image on the retina during rotations of the head.

The VOR helps in maintaining clear vision and compensates for perturbations of the head. The VOR counter-rotates the eyes in order to keep the line of sight stable during head movements. In cases where a large change in gaze requires that eye and the head to move in the same direction, the VOR may be counterproductive. During active head movements, two mechanisms of cancellation and suppression have been postulated to prevent counter-productive VOR movements. The cancellation mechanism generates a command opposite to the head command and hence negates the action of the VOR, as during smooth pursuit movements (Lanman et al., 1978). VOR eye movements are produced at lower latencies than visually mediated ones and compensate for head movements at a latency of less than 16 ms. Without VOR, our eyes would see a smeared image every time we moved our head rapidly. Previous studies show interaction methods of detecting head gestures based on natural VOR eye movements (Mardanbegi et al., 2012; Špakov and Majaranta, 2012), or based on nod and roll through VOR (Piumsomboon et al., 2017). In Daye et al. (2015), gaze trajectories with and without whole-body rotations during saccades were compared. The saccades were made when both head and body were fixed with respect to the seating position and the seat moved. Their results showed that gaze remained accurate despite the head perturbations, although the VOR was suppressed. During a large saccade, the suppression mechanism causes decrease in the VOR gain at the onset of a head-unrestrained saccade and then increases it before the end of the saccade (Daye et al., 2015; Tomlinson and Bahra, 1986; Lefèvre et al., 1992).

In VR applications, the virtual images should be stable in space, despite a user's head movements. Because of the system latencies in the head tracking, the virtual objects may float around instead of being stable. This causes an unnatural experience, disorientation, cyber-sickness, and mismatching of virtual and real objects in VR and AR (Wu et al., 2019; Jerald et al., 2008). The latencies in the visual updates are a big challenge in the design of VR and AR applications and have been focus on much engineering research. These system latencies affect objective performance measures as well as user's sense of presence in VR (Ellis et al., 1999). In a Head-mounted Display, head motion can be measured by a 6 degree of freedom (6DOF) inertial measurement unit (IMU). A study (Pfeil et al., 2018) compared how our eye-head natural movements in responses to stimuli change in Virtual Reality (VR) as compared to Physical Reality. They ran experiments on groups of subjects to identify differences in virtual and physical environment, while observing their eye and head movements through tracking. Their results indicated that participants moved their heads more often when viewing target stimuli, including text, in VR than in real environments.

2.1.3 Saccadic Suppression

Human visual sensitivity is greatly reduced around the times of saccadic eye movements. Visual sensitivity can be examined through perceptual detectability of very brief visual stimuli. During saccadic eye movements, our perception of motion, scene displacements and visual stimuli is attenuated, and we have very limited capability to obtain visual information (Matin, 1974; Ibbotson and Krekelberg, 2011). This suppression of visual information during a saccade is known as the saccadic suppression effect, and may be related to the perceptual compression of space and time during saccades (Burr and Morrone, 2010). Active suppressive signals which are derived directly from eye movement commands have been postulated to explain this perceptual phenomena. Nevertheless, despite being studied extensively (Matin, 1974; Ibbotson and Krekelberg, 2011; Zuber and Stark, 1966), the neural mechanisms of saccadic suppression still remains controversial, after decades of research (Zuber and Stark, 1966; Idrees, 2021; Bremmer et al., 2009). For example, different views relate the neural origins of saccadic suppression to active suppression triggered by the signals related to eye movements, or to the visual consequences of saccades causing suppression (Diamond et al., 2000; Zuber and Stark, 1966).

The suppression of our visual sensitivity begins around 50 ms prior to the onset of the saccade (Matin, 1974), and research (Ibbotson and Krekelberg, 2011) shows that saccadic suppression can last for at least 100 ms after the saccade, even with a 50 ms saccade. In particular, small displacements to objects that occur during saccades are not noticeable by humans (Allison et al., 2010). This is known as saccadic suppression of displacement (SSD). When a blank is introduced between presaccadic and postsaccadic stimulus, the effect of SSD is reduced (Bridgeman et al., 1975; Deubel et al., 1996), but is strongest when the displacement of the stimulus happens in an elliptical area along the axis of saccade direction, and scales with saccade amplitude (Wexler and Collins, 2014). This means that larger object displacements are better hidden when saccade amplitudes are larger (Bridgeman et al., 1975; Li and Matin, 1990). In addition, saccadic suppression, appears to be stronger in central than peripheral vision (Bridgeman and Fisher, 1990).

Saccadic suppression is not apparent as the brain combines information from successive eye fixations to create a subjective impression of continuous view of the visual field. Some stimuli are actively suppressed by saccades, including stimuli with low spatial frequencies, and are very hard to detect if flashed just prior or during a saccade, while stimuli with high spatial frequencies are more visible (Volkmann et al., 1978). Moreover, stimuli with varying colors but equal luminance are very minimally suppressed during a saccade (Binda and Morrone, 2018).

The saccadic suppression effect can be used in time sensitive applications in 3D visualizations. In interactive computer graphics, there may be the need to perform different modifications to the images on the display to keep the system up to date. Immediate and abrupt changes in the displayed images could cause disturbing effects in the visibility of the 3D scenes for the viewers. The saccadic suppression effect has been suggested as a technique for masking extensive graphic updates in a 3D virtual environment. This allows creation of displays that change properties during a users' saccadic eye movements. Disturbing changes in 3D scenes including viewpoint correction can be masked if the update occurs during the saccade (Schumacher et al., 2004; Franke et al., 2014). It can be used to achieve various manipulations with a computer display without viewer awareness. It can further be useful in studies of visual representation and memory, change blindness and virtual environments (Triesch et al., 2002). There have also been advances in augmented reality displays using gaze tracking for dynamically adaptive resolution and focal depth. Saccadic suppression was used to hide the graphical changes that occurred between transitions of foveal and peripheral regions (Kim et al., 2019).

Research has been conducted to compare the perceptability between transsaccadic (during a saccade) and intersaccadic (between two consecutive saccades) scene changes. Schumacher et al. (Schumacher et al., 2004) predicted saccade amplitude in the early stages of a saccade to determine whether imperceptible updates could be made. They used saccadic suppression of image displacement to mask updates in a 180° hemispherical display with an interactive VR CAVE-like setup. Small horizontal translations were not typically noticeable when they occurred during saccades of at least 58 ms in duration and in cases where they were noticeable, they were not very disturbing for the viewers. Likewise, Herpers et al. (Herpers et al., 2004) confirmed that horizontal translations occurring during a saccade while the user was viewing natural 2D scene images, were perceived as much smaller than similar translations in the absence of a saccadic eye movement. A follow-on study (Allison et al., 2010) found that apparent motion of images was significantly less detectable during saccadic eye movements compared to when the eye was fixating. Similarly, in a 3D virtual reality, users had less sensitivity to scene rotations during saccades compared to the changes occurred during fixations (Bolte and Lappe, 2015).

Furthermore, saccadic suppression has been studied in different age groups, and the research shows that the amount of suppression at saccadic onset depends on user's age. Saccadic suppression of contrast sensitivity has been shown to be stronger for children of 12-14 years of age than in adults (Bruno et al., 2006). There are mixed results for how transsaccadic stimulus displacements are suppressed. Even though children can integrate basic visual feature information from an early age, some studies failed to show an ability that complex sensory information in children can be integrated (Gori et al., 2008; Nardini et al., 2008). Stewart et al (2020), found that children aged 7 to 12 years old had stronger transsaccadic suppression of displacement than the adults aged 19 to 23 years, which they suggest is due to higher intrinsic uncertainty in target localization or saccade execution (Stewart et al., 2020).

2.2 Inattentional Blindness

Inattentional blindness is a person's inability to detect a visible object in plain sight, because of lack of attention (Mack and Rock, 1998). Probably the most famous example of inattentional blindness is presented in (Simons and Chabris, 1999), which asked participants to watch a video and count the number of times a basketball was passed between six people in the video wearing a white shirt. At some point in the video, a person dressed as gorilla walks from right to left of screen, beating its chest. At the end, most participants stated they did not notice the gorilla. This was later studied in a 3D VR setting (Schöne et al., 2021), and the results showed that the participants were more likely to notice the gorilla in the VR, which may be due to the sense of presence and spatial proximity in VR that enhanced the presence of the gorilla. Research has shown that the amount of workload affects the strength of inattentional blindness, and that the perceptual load of a task has an important role in determining conscious perception (Cartwright-Finch and Lavie, 2007). The detection of a stimulus depends on how much it can attract attention and the difficulty of the primary task. With more difficult tasks, there is more attentional resources required (Simons and Chabris, 1999; Jensen et al., 2011). Wang et al.

(2022) studied inattentional blindness on an AR heads up display (AR HUD), and found that a higher workload induced a stronger inattentional blindness than a low workload (Wang et al., 2021). Despite this, the ability of the users in achieving those tasks, i.e. individual differences, did not affect the amount of inattentional blindness (Simons and Jensen, 2009).

Inattentional blindness has been used in different applications of VR, including redirecting the users while navigating in VR (Joshi and Poullis, 2020b), and also through auditory distractions. One study found that an auditory effect can degrade the detection of visual stimuli by VR users (Malpica et al., 2020). They recorded their users' eye movements and found that their gaze behaviour did not change during this degradation. They conclude that this degradation occurs even in the absence of saccades towards the sound source.

2.3 Gaze-Contingent Displays

Gaze-Contingent Displays (GCDs) are dynamic displays that use real-time eye movement analysis acquired from an eye-tracking device, and manipulate the corresponding point (pixel), peripheral regions, or an object in the virtual space accordingly. They tend to improve accuracy and visibility, or rendering quality, of the object at the point of regard and balance the displayed information with the viewer's visual information processing capacity. Model-based GCDs try to reduce the resolution by directly manipulating graphical model geometry (3D graphical models) prior to rendering. These gaze-contingent model manipulations aim to reduce rendering requirements and are specifically important for increasing the speed or improving the latency in immersive displays such as Virtual Reality headsets or other complex graphical rendering environments (Duchowski et al., 2004). With the recent demand for creating life-like virtual environments and high-quality computer graphics, the demands for higher pixel densities, higher processing powers and more bandwidth are increasing too. One way to overcome these limitations in current displays is to tailor resource allocation to the perceptual attributes of our visual system. With that in mind, it is not necessary to render all display pixels at the highest visual quality.

In 1973, Reder (Reder, 1973) introduced several paradigms that centered around a stimuli being presented when the participant's gaze was fixated upon a predefined location. Then the reaction to the newly presented stimuli could be measured and compared to other types of presented stimuli such as reaction time to other target images or to other participants' responses. In his experiments, he blocked a person's peripheral vision to only show the content in their foveal vision, or blocked their foveal vision to pay attention to their periphery, or magnified content in their periphery. Also in 1990, another gaze contingent approach was introduced which rendered volume datasets according to the user's view direction (Levoy and Whitaker, 1990). Since these early developments, GCDs have been advancing very fast and are being used in many fields and applications. Gaze-contingent displays can use real-time eye tracking to modify their content and provide better depth perception by additional depth cues. In these methods, the optical effects of accommodations can be simulated. So only objects at the distance of the fixated object are seen clearly (and blurred otherwise) (Vinnikov and Allison, 2014). Figure 2.1 shows a sample stimuli used in (Mauderer et al., 2014) to study a depth of field GCD.

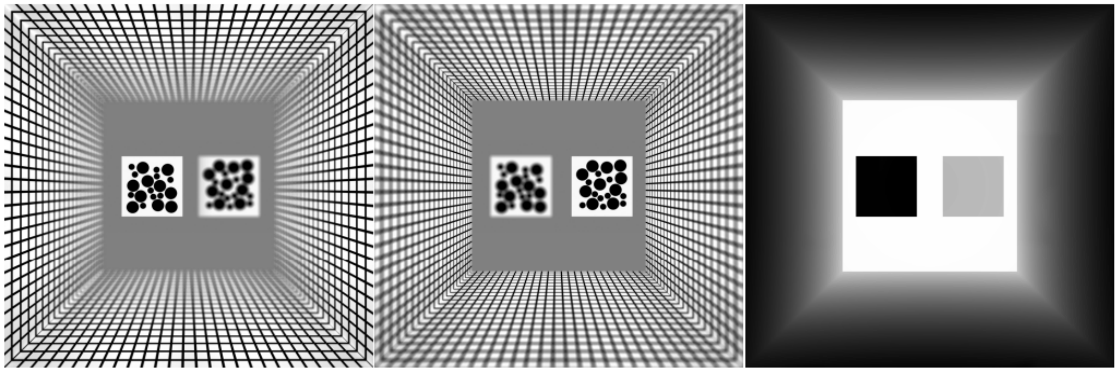


Figure 2.1: Stimuli used in a gaze contingent depth of field study in (Mauderer et al., 2014), showing focus on the left and right panel and their depth map.

Screen-based GCDs modify the contents at a pixel level using image processing, and intend to match the resolution of the human retina, and aim to increase the bandwidth. In some implementations, the functions needed for image blending depend on the alpha channel of the images, and blend the foveal and peripheral image sections. This is a disadvantage of the screen-based methods because it is expensive to translate the alpha channels in real-time to match the foveal region (Duchowski et al., 2004). However, with current graphics card technology, image processing operations can be performed quickly through hardware accelerated convolution. In the spatial domain, a single frame is processed at a time and therefore can have real-time performance.

2.3.1 Applications

Over the years GCDs have been used in many applications to improve quality of rendering and immersiveness in VR. In one study (Chakravarthula et al., 2021), the features of human foveal and peripheral vision were used to reduce speckle noise of computer generated holographic displays. Their perceptual model, leveraged the anatomical and statistical retinal receptor distribution to optimize their computational hologram, and reduced the perceived foveal speckle noise. Another GCD-based study (Albert et al., 2019), looked at changes in the reading speed with various peripheral degradation levels. They examined different text sizes, foveal regions and subsampling kernels. Though some of their results were not significant, they found that for various types of peripheral degradation, faster readers were more negatively impacted, which may be due to them making more use of their peripheral and foveal vision.

Gaze-contingent renderings can overcome certain limitations of displays such as low dynamic range or lack of accommodation cues (Duchowski et al., 2014; Vinnikov and Allison, 2014). Gaze-driven stereo disparity modifications can enhance the impression and perception of depth (Kellnhofer et al., 2016). In this study, depth adjustments were gradually applied at eye fixation, and a model was used to modify local stereoscopic content. These display changes are based on eye tracking information, and try to optimize visual comfort based on depth reproduction. While GCDs are mainly based on visual processing, they can be extended to update auditory content too (Vinnikov et al., 2017). GCDs can also be recognized for manipulating the perceived color. Jacobs et al., (2015), simulated visual events such as loss of acuity and aftereffects with a GCD and found that brightness perception can be altered. Also their results show that in tone mapping and luminance adaptation, GCDs can allocate the dynamic range effectively by reducing the image contrast with eccentricity (E. Jacobs et al., 2015). Mauderer et al. (Mauderer et al., 2016) conducted two studies that manipulated peripheral background and object colors to influence the viewer's color perception. Their findings showed that gaze-contingent simultaneous contrast can be used to modify color appearance. They also suggest that the existing color appearance models may not fully predict perceived colors with GC presentations, and show that gaze-contingent adjustments can be used to enhance color discrimination. Gaze-contingent color can expand the perceived color gamut of existing display technologies and enable users to discriminate color with better precision. Also, in (Chen et al., 2022), the perceived flickering of an image is reduced. This is achieved through leveraging saccadic suppression, and reducing the flickering after the gaze lands. Their method does not completely remove temporal artifacts, but a local adaptive rendering is suggested to be helpful.

GCDs and saccadic-based manipulations have been used to redirect the users of a VR to new positions. Gaze-contingent research and applications depend a lot on the different saccade amplitudes and the detection algorithms of eye movement events (Stein et al., 2021). During moments of effective blindness (saccades or blinks), the viewpoint of the users can be subtly changed (Bolte and Lappe, 2015; Langbehn et al., 2018; Schumacher et al., 2004). For a user in a virtual reality, room-scale VR provides an empty space that lets the user move freely. The VR system uses sensors to track the user in all directions and translates them into the virtual world space. Room-scale VR creates more comfort for users than stationary VR where the user must use an input device (such as a joystick) to navigate in the VR space. It reduces the visual-vestibular inconsistency and allows users to walk freely in a physical space. However, the room space is typically small compared to the virtual environment and a direct mapping from VR space to physical room space is not practical in many cases. The goal in this situation is to embed a large VR space within a finite physical space with minimum interruptions for the user (Sun et al., 2018). Treadmills have been proposed to allow infinite walking, but they are expensive and not very accessible. Redirected walking is one of the methods proposed to solve this mapping issue (Razzaque, 2005; Steinicke et al., 2009). Re-orientation and re-positioning of the users ensures the safety of users in a VE as the system moves them away from obstacles such as boundaries of the tracking space. By applying subtle rigid-body and nonlinear transformations to the virtual world, these methods create a distorted mapping of the VE to the real world, which magnifies the effective physical space. Unlike traditional methods, these techniques can in principle respond to the real-time environmental changes such as displacement of objects or obstacles in the room.

Saccadic suppression of image displacement, used in GCDs, has been suggested as an effective tool for masking redirected walking manipulations in virtual environments. By detecting and tracking the type of eye movement the user makes in a virtual environment, it is possible to subliminally reposition them. This can happen during a blink or during a saccade (Langbehn et al., 2016). Bolte and Lappe (2015) studied saccadic suppression of image displacement in an immersive virtual environment by rotating or translating the camera during saccadic eye movements. Their results showed that participants were less sensitive to transitions during saccades than during fixations (Bolte and Lappe, 2015). However, they only looked at this problem for two degrees of freedom and analysed movements about the yaw axis for rotation and forward and backward translations. Sun et al., investigated infinite walking in VR using saccadic redirection techniques (Sun et al., 2018). Their method imperceptibly rotates the virtual scene when the user performs saccades and head rotations. They implemented a real-time dynamic path planning method which avoids the user hitting the moving obstacles. To do that they used saccade detection thresholds and the physical space around the user to dynamically determine the best virtual camera orientation for the redirection. To trigger more saccades from the user they use Subtle Gaze Direction, which uses image-space modulation to direct a viewer's gaze to a specific target. Pinson et al.(2020) used saccadic redirection on a HTC Vive Pro Eye to allow infinite walking in VR (Pinson et al., 2020). They only rotated the player's viewpoint on one axis, and allowed only one rotation in a specific period of time, to avoid multiple consecutive scene rotations in response to saccades with short fixation times in between.

Studies have also used blinks as a means of hiding user re-orientations in a VR. During eye blinks small changes can go undetected because visual signals and retinal responses are suppressed by the visual system and inhibitory mechanisms, or vision is obscured by the lids (Volkmann et al., 1980; Riggs et al., 1981). Thus, target displacements during eye blinks can also go unnoticed by the users (Maus et al., 2017). Langbehn et al. (Langbehn et al., 2016; Langbehn et al., 2018), looked at both visually-simulated translations and rotations of the users in any direction during eye blinks. The virtual scene that participants viewed was rotated or translated when they were asked to blink and subjects then indicated the direction in which they were virtually moved. The changes to the virtual camera were suppressed during an eye blink, suggesting that imperceptible position movements are possible during blinks. Furthermore, change blindness was studied during eye blinks while the user wore an HMD, and results showed that detectability of scene changes was not only dependent on the angle of rotation, but also on the layout of the objects in the scenes (Bruder and Langbehn, 2017). Similarly, Nguyen and Kunz (2018), found much higher detection thresholds for scene rotations during eye blinks compared to when the eyes were open. Their results showed that detection thresholds had a negative correlation with walking speed, and that gender of the user and their prior gaming experience did not have any effects on these detections (Nguyen and Kunz, 2018). But in a follow-up study (Nguyen et al., 2018a), they found that women had a higher tolerance of curvanture redirection, but did not report the extent of these differences. Although eye blinks seem to be a good mechanism for redirecting users in a virtual environment, and compared to saccades are easier to detect, they do not occur as frequently as saccadic eye movements in a real application. Therefore saccadic eye movements are preferred over eye blinks for natural immersion and locomotion in VR. In (Joshi and Poullis, 2020b), inattentional blindness was used for real-time redirected walking in small physical spaces. The user's view was divided into zones based on their fixations, and using foveated rendering, the zones receive spatially-varying rotations based on their importance. This approach used both saccades and blinks (that occurred naturally) to update the framebuffer. Also, incongruent multimodal visual-auditory cues in VR, such as alteration of the objects' visual location through auditory cues during redirection, were found to have a better application for redirected walking, than congruent cues (Gao et al., 2020).

Figure 2.2: Saccadic redirection implemented as a VR chess board game (Sun, 2021)

An example of such redirections is shown in Figure 2.2, where the red curve is for virtual direction and green curve is used for physical path. The game allows participants wearing an eye-tracked HMD to freely walk around to place chess pieces on a virtual board, and resulted in none of the participants noticing that they were being redirected (Sun, 2021).

2.3.2 Challenges

An important issue in GCDs is the issue of latency. Real-time synchronization of eye movements and the image displayed is crucial in such applications. The system latency is the delay between a change in gaze position and the related update in the display, which causes a mismatch between the rendered view and the actual gaze location. This lag time depends on gaze sampling, and also the time to update and refresh the GCD. This is the reason some studies are now working on some sort of predictive capabilities (Loschky and McConkie, 2000; Duchowski, 2002).The experiment designer should bear in mind the challenges of choosing thresholds for delays that yield imperceptible display modifications to the users. Arabadzhiyska et al.

(Arabadzhiyska et al., 2017) proposed a model to predict the landing position of the saccade and update the display image for the new fixation location. Because of the skewness of the velocity profiles of large saccades, a saccade velocity look-up table was created for each user. Hence the display rendering was not performed according to the current gaze position but to the predicted gaze location, and a correct image for the new fixation was provided before the fixation was established. Due to saccadic suppression, the quality mismatch during the saccade was not noticeable for the observer. Albert et al. (Albert et al., 2017), looked at how latency affects foveated rendering in VR. They looked at three different visual foveation techniques including Subsampling, Gaussian Blur and the perceptual fCPS (foveated Coarse-Pixel Shading). They used a staircase method with varying peripheral blur and also maximum blur to find the perceptual threshold latency. They found that overall a system latency of 50-70 ms can be tolerated in a VR application with eye tracking. On the other hand, there are content-based approaches to gaze prediction (without using real-time eye data) which use saliency maps or machine learning approaches to classify important scene objects or locations (Koulieris et al., 2016). An eye tracker is used, and a classifier is trained to find the correlations between a player's actions and the current state of a game, and is used at runtime to predict gaze (by predicting object category in the game). However, the classifier requires specific training for every content, has a low speed, and was not participant's best choice for comfortable viewing. In recent studies, artificial neural networks have also performed well in predicting gaze. In (Morales et al., 2018), recurrent neural networks are used to predict the saccade landing position. In (Zhang et al., 2018), the authors predict gaze on future frames using adversarial networks. Although these approaches are accurate and perform well and could give priors for real-time predictors, they may not be very suitable with gaze-contingent displays which require fast and robust online gaze detection and graphical updates.

2.4 Eye, head and hand movement

Our movements in daily life depend on visual information for feedforward and feedback control. When we interact with the world, we make many gaze shifts through a combination of eye, head and body movements. We often make two to three gaze shifts per second, to obtain visual information from our environment. When we move our eyes from one point, for example a point in front of our eyes, to another point on our side, we actually move our eyes in coordination with the rotations of our head. Often neither our head nor our eyes can cover the full distance. During saccades, we move our eyes in the head (relative to the head), and move our head relative to the torso, and move our torso relative to the world (Land, 2004). Normal human self-motion is accompanied by various eye movements. With self-motion, our eye movements are also important in our visual perception. Self motion induces image motion on the retina and so the world image on the retina of our eyes is also in motion. VOR eye movements counteract this induced visual motion and stabilize the fixated image. When we are walking on foot, our locomotion is guided by various visuomotor factors. For example, the placement of the foot in the step cycle is an important parameter that needs to be controlled (Lappe et al., 2000). Gaze in walking is found to be mostly directed towards future landing positions of the subject's feet (Zietz and Hollands, 2009). In an outdoor environment, while walking, subjects gaze was mostly directed to objects close to them (Wagner et al., 1981). Overall, research shows that the pattern and distribution of eye movements and gaze depends on a person's task (Lappe et al., 2000). In some actions, such as looking up at a display above our head, we move our eyes, head as well as shifting our torso towards the new target. This coordination of the three systems movement is seamless and provides efficient acquisition of gaze targets.

There are also slow eye movements that occur between gaze shifts. These VOR eye movements use vestibular, proprioceptive and visual signals to keep the retinal image on the eye stable (Leigh and Zee, 2015). Head rotation and translations in space induce corresponding signals that are directly used to move the eyes opposite to the movement of the head. These eye movements are known as the rotational and translational vestibular ocular reflexes. Gaze stabilization requirements differ for rotational and translational self motion. When the head or body rotates, the whole visual scene rotates with a single angular velocity. The rotational VOR eye movements compensate for head rotations by rotating the eyes in the opposite direction, and its speed is very close to the speed of the head movement (Lappi, 2016; Angelaki and Hess, 2005). Hence, good image stabilization is achieved. With lateral or up-down translations of the head, eyes are again rotated against head movement. But for accurate image stabilization, the required speed of the eye movement cannot be determined from the head movement alone. So the geometry of the visual scene should also be considered. There are more complications with a forward motion, as it induces patterns of optic flow in the eyes, and different points in the visual field move in different directions. Therefore, only the stabilization of part of the visual image on the retina is possible, not the entire image (Berthoz and Droulez, 1982; Lappe et al., 2000).

Interaction with 3D immersive VR systems is becoming more accessible, and easier. VR environments can provide better control for their users by allowing them various interaction means, such as eye tracking, gaze selection methods, and using haptics. Goal-oriented hand movements are often accompanied by a saccade. These saccades direct a person's gaze precisely to an observed target location, or to a predicted target location when the target was moving before the hand starts to move (Abrams et al., 1990; Frens and Erkelens, 1991; Lünenburger et al., 2000). The hand movement follows the gaze shift (Lünenburger et al., 2000). These saccades provide visual information about the observed or predicted target location to guide a hand movement. Research suggests three aspects for having an effective visual guidance of the hand. One is timing (gaze arriving at the target before the hand), accuracy (gaze reaching the goal as precisely as possible and stability (gaze remaining there until the hand reaches the goal) (Prablanc et al., 1979; Vercher et al., 1994; Lünenburger et al., 2000). While performing a line drawing hand movement, a series of small saccades followed the pencil's trajectory, which contributed to feedback control (Tchalenko, 2007). Therefore, studies suggest that two types of saccades may be associated with hand movement. One type includes a saccade that directs an individuals gaze to a target position, as in during a reaching movement. Another type directs a person's gaze to a hand position, and is seen when drawing a simple line. In manipulation tasks, it is important to study how eye movements are directed to objects of interest. In (Johansson et al., 2001), the authors showed that the planning and control of manipulatory actions are supported by gaze, and are achieved by marking key positions that the hands or grasped objects are directed to. The gaze targets were mainly determined by the demands of the sensorimotor task.

Saccadic suppression seems to be stronger when a hand movement accompanies a saccadic eye movement (Blouin et al., 1995). Results of one study showed that when the subjects combined simultaneous eye and arm movements towards a target stimuli that is displaced during a saccade, the perceptual thresholds of target displacements and was of saccadic suppression were increased.

2.5 Summary

In chapter 2, I reviewed some of the important theoretical topics of human visual perception and eye tracking which were related to this PhD project. In addition, I reviewed some of the most that studied eye tracking in virtual reality and how human perceptual features can be related in such studies. We described different eye movements and showed how relevant of visual perception can contribute in the design of virtual reality systems. The visual characteristics of the human peripheral and central vision have widely been leveraged for improving depth perception, distance judgment, foveated rendering and also to reduce cyber-sickness. Perceptual phenomena of saccadic suppression, change blindness and blink suppression can be used for hiding graphical updates in a virtual environment, and these updates can support more complicated tasks such as redirected walking in VR and to subtly reposition the users. Although real-time modelling of a saccadic eye movement remains a challenge due to its speed, there are some features of it such as duration, amplitude and velocity profiles that have been used to predict the landing position (Schumacher et al., 2004; Arabadzhiyska et al., 2017; Morales et al., 2018). While gaze-contingent techniques have been around for a long time, eye tracking technology, saliency estimation methods and graphical hardware have recently become fast enough and affordable for all users. In this chapter we discussed the main features of Gaze-Contingent Displays and some of their applications.

Chapter 3

Gaze Contingent Experiments: Saccadic Suppression Sensitivity

3.1 Introduction

Eye movements involve cognitive, motor and perceptual processes and are required to fixate objects and for clear stable vision. When exploring or interacting with an environment, whether real or virtual, we typically successively fixate areas and objects of interest. These fixations place the images of objects we look in the high-resolution center of our retinas, the fovea. To change fixation, both of our eyes jump simultaneously from one location to another about three to five times per second. These rapid conjugate eye movements, known as saccades, are used to redirect the foveas to a new area of interest (Bridgeman et al., 1994; Rayner and Pollatsek, 1992). We acquire visual information through consecutive fixations, although we are normally unaware of this process and have the impression of a coherent and stable visual scene. During saccades the eye senses information as well, but no sharp images can be obtained because the image of the stationary world smears across the retina. However, we do not normally see these smeared images. Our visual acuity and sensitivity to motion is reduced during saccades. This reduced transsaccadic visual sensitivity is known as the saccadic suppression effect and is a well documented phenomenon (Hopp and Fuchs, 2004; Bridgeman, 2012). In virtual reality, this perceptual phenomena has been studied as means of interaction with the virtual environment. Some of these applications include hiding graphical updates and increasing the processing speed (Wei et al., 2020; Schumacher et al., 2004), navigation and infinite walking in VR (Sun et al., 2018; Pinson et al., 2020), imperceptible redirecting and repositioning the viewpoint of the user (Bolte and Lappe, 2015; Langbehn et al., 2018), prediction of saccade landing position (Arabadzhiyska et al., 2017) and foveated rendering techniques (Kruchinina et al., 2020; Joshi and Poullis, 2020b; Joshi and Poullis, 2020a; Sun, 2021). This diversity and breadth of an application was a motivation for the experiments conducted in this dissertation.

In this chapter, I discuss the details and results of two gaze-contingent experiments. These experiments tested saccadic suppression sensitivity of users in two different experiments. The saccadic suppression effect, in which visual sensitivity is reduced significantly during saccades, has been suggested as a mechanism for masking graphic updates in a 3D virtual environment (Schumacher et al., 2004; Patney, 2017). In the first experiment I investigated whether the degree of saccadic suppression depends on the type of image change, particularly between different natural 3D scene transformations. During the second experiment I followed a similar procedure and added a constant camera motion in the scenes to simulate a head motion, and investigated whether a constant camera rotation in a virtual scene modulates saccadic suppression.

3.2 Experiment I: Saccadic Suppression of Image Transformations

In gaze-contingent displays, the viewer's eye movement data are processed in real-time to adjust the graphical content. Such updates can be used to introduce imperceptible changes in virtual camera pose in applications such as networked gaming, collaborative virtual reality and redirected walking. To provide a high-quality user experience, these graphical updates must occur with minimum delay. For such applications, perceptual saccadic suppression can help to hide the graphical artifacts. In this experiment, I investigated whether the visibility of these updates depends on the type of image transformation.

3.2.1 Hypothesis and Objective

In this study, I leveraged the fact that our visual sensitivity and perception of motion are significantly reduced during saccadic eye movements. The objective of this study was to discover how various combinations of image transformation in different rotational and translational directions, with various sizes, can be detected by users of VR when they make saccadic eye movements. This objective was important as this sensitivity is not known and it provides substantial background knowledge in VR content creation, games, and other applications. I looked at the patterns of detectability of changes in the camera angle and camera position in the scene while the user made a saccadic eye movement. I hypothesised that there would be more suppression for transitions that are parallel to the saccade direction. For instance, I expected lower rates of correct detections for cases such as right or left scene translations during horizontal saccades, compared to vertical transformations during these saccades. Intuitively and confirmed by previous research (Schumacher et al., 2004; Allison et al., 2010), I was also expecting less sensitivity to image transitions that are smaller in size than the larger ones.

3.2.2 Stimuli and Apparatus

The stimuli were presented in the form of two virtual environments simulating an urban intersection with an open square, buildings and other features. The complexity of the virtual environment was kept modest to ensure it could be rendered consistently in real time. Screenshots of the displays that the users viewed are shown in Figure 3.1. For horizontal saccades, users fixated a car initially on the left side of a street-level view of the scene and made a 15° horizontal saccade as the car was displaced to the right, as can be seen in Figure 3.1 (a) and (b). For the vertical saccade condition, participants were asked to look at the car in Figure 3.1 (c) and make a 15° vertical saccade as the car was displaced vertically. This is shown in Figure 3.1 (d).

I was interested to see users' sensitivity to rotational and translational changes along the following directions, during these horizontal or vertical saccades. Some of these image transformations are shown in Figure 3.3.

- Rotational changes:
  - Roll axis : clockwise and counter-clockwise camera rotations
  - Pitch axis : upward and downward camera rotations

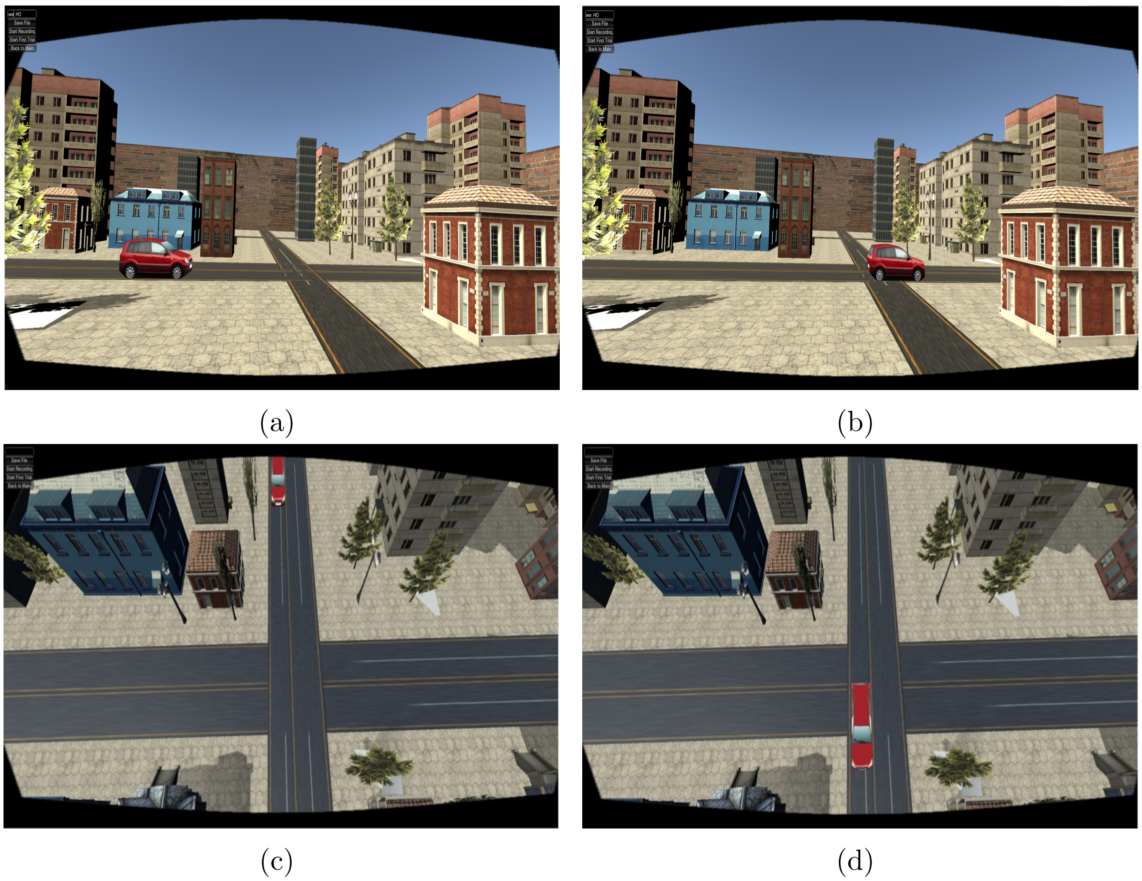


Figure 3.1: Images (a) and (c) show the location of the target stimuli (the car) before the saccade and images (b) and (d) show the final location. Subjects were to fixate their eyes on the target object and when the car was displaced to the final location, execute a horizontal or vertical saccade in response to its jump. The irregular border reduces frame cues to the transformation. (There is no scene transformation in the above images.)

  - Yaw axis: right and left camera rotations

- Translational changes:

  - Horizontal axis : right and left camera translations
  - Vertical axis : upward and downward camera translations
  - In depth : forward and backward camera translations

For the six translational directions (forward, backward, left, right, up and down) I presented each of two sizes of translation, 0.5 m and 1.5 m. Also, for each of 6 rotational transitions (yaw left, yaw right, pitch up, pitch down, roll clockwise and roll counter-clockwise) I presented each of two sizes of rotation, 2° and 7°. Thus, there were overall 24 different conditions (6 directions x 2 types (rotation/translation) x 2 sizes), that were ran in several experimental blocks. I chose these values based on previous research studies, and also based on running pilot testing to find values that their shift amounts on the scene would be.

All 3D scenes were created and rendered in real-time on a Windows 7 desktop computer with AMD FirePro W9000 FireGL, Intel Core CPU 3.5GHz and 3.50 GB RAM. The visual environments were designed in Unity3D and C# scripts, and were presented on a 27-inch 3D Samsung LCD monitor, with a resolution of 1920H * 1080V pixels, and a refresh rate of 120 Hz. The users' eye movements were recorded with a video-based EyeLink 1000 system [SR Research Ltd, Oakville, ON, Canada]. The setup of the eye tracker was in the tower-mount mode with a chin-rest for stabilizing the participant's head and ensuring a fixed viewing distance of 55 cm from the display, as shown in Figure 3.3. I set the sampling rate to 1000Hz.

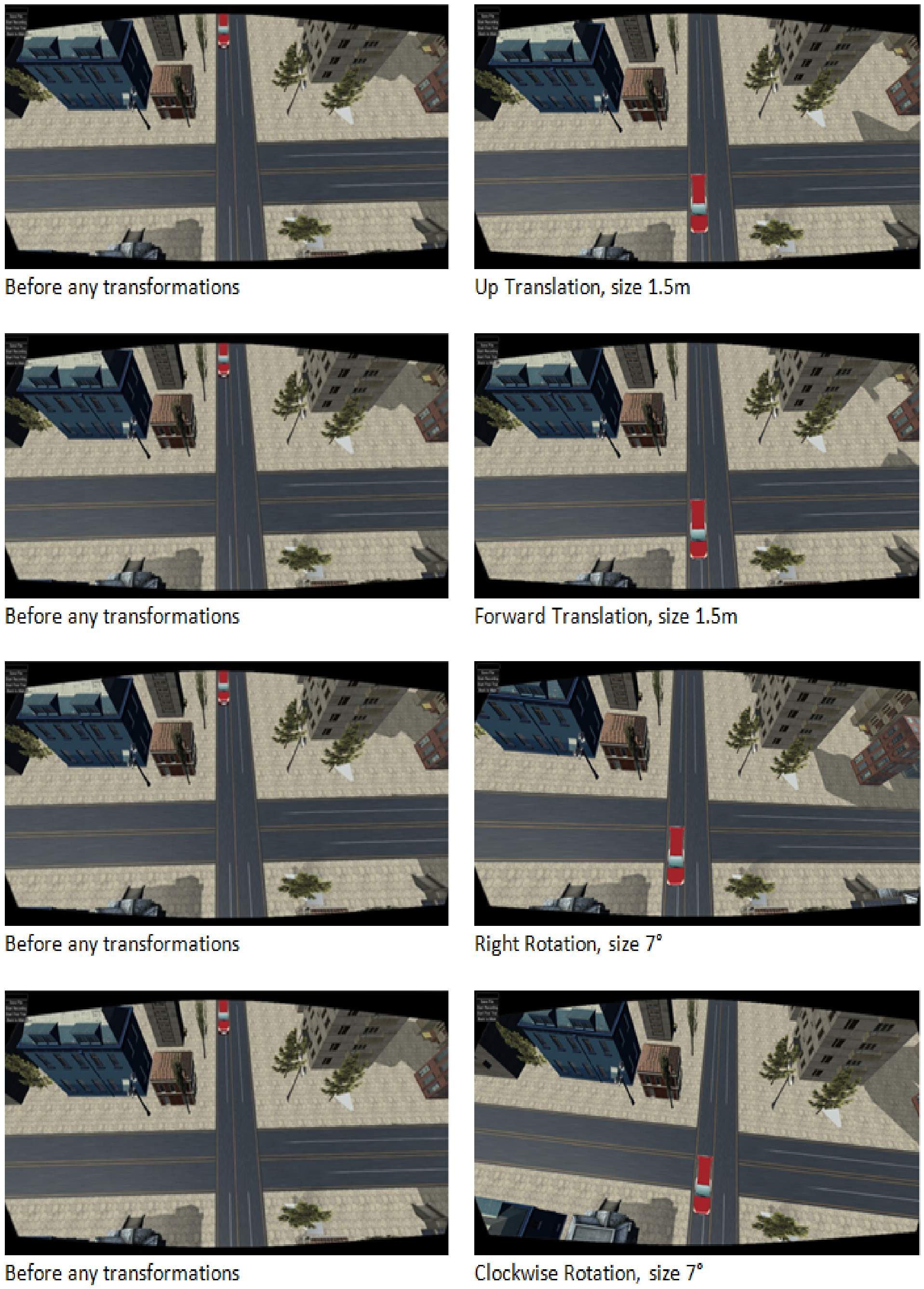


Figure 3.2: Sample image transformations applied during the saccade.

Figure 3.3: Eyelink eye tracker in tower-mount setup mode

3.2.3 Experiment Design and Procedure

Participants sat at the desktop-mounted eye tracker which was placed in front of the monitor. Before beginning their main block, they were given instructions in a short training, about the goals of the experiment and what they were expected to do. The scene used in the training was the same scene used in the experiment. They were shown how each camera movement looked (not during a saccade) and what questions they would be asked after each trial. When I ensured that they were clear about their task, they started the first block. Each block started with a calibration and validation of the eye tracker, followed by 72 trials in random order. In every block, each condition was repeated three times. Every participant attended two sessions of three blocks (three block with horizontal saccades and three with vertical) on two separate days, to avoid eye strain and fatigue. To counter balance order effects, I presented the blocks in alternating order, starting with either horizontal or vertical saccade block, and randomly assigned them to each participant. They fixated their eyes on the target object in the 3D scene and follow it with their eyes as it jumped from one point to another. The duration of each trial was 2.0 seconds and the object was displaced after 1.0 second. As the participant looked at the object being displaced, they performed a saccade of 15 degrees of visual angle. At the same time as this saccadic eye movement and upon detection of a start of a saccade, a translation or rotation in a specific direction was applied to the scene. This transformation was applied through a stepwise translation or rotation to the virtual camera that showed the scene. Participants were asked to indicate in which direction they detected the camera change. They had to choose one of the eight directions in a forced-choice question and indicate their confidence level for their answer, on a spectrum of *Not Confident* to *Very Confident* . They then proceeded to the next trial. In trials where they did not notice any changes, they were still required to guess the direction of camera change.

There were catch trials in each experimental block. In such trials, no display updates occurred but the users were still asked to indicate the direction of camera change during that trial. In cases where a saccade was not detected in a trial, the current trial was repeated until a saccade was detected. However, the users did not notice this and answered these trials in the same manner. These trials were counted as additional catch trials.

Saccade Detection

Human eyes do not move in a deterministic manner, and detecting saccades automatically is challenging. To detect a saccade in a gaze-contingent display, we need eye position estimates obtained through an eye tracking device as well as a robust real-time saccade detection algorithm. There have been different approaches proposed, such as dispersion-based, velocity-based, acceleration-based and area-based algorithms for detection of saccades (Salvucci and Goldberg, 2000; Duchowski, 2007). However, some of the methods are unsuitable for online saccade detection. Although they may be very accurate, they require that the entire saccadic profiles be recorded before performing the classification (Andersson et al., 2017; Nyström and Holmqvist, 2010).

There have also been studies that classify eye movements using deep neural networks. In (Bellet et al., 2019), convolutional neural networks were used to detect saccades. Four different datasets were used to train the network. Although their model produced high accuracy classification of the eye samples, the datasets used did not vary on different sampling frequencies, and may not generalize well to larger eye movements. Another study used a Bayesian approach for the detection of microsaccades based on a generative model (Mihali et al., 2017). Bayesian methods can provide estimates of uncertainty for the detected eye movements. Choosing an eye event detection algorithm is challenging and depends on a variety of parameters, including the eye tracking hardware. There are equipment-dependent hyper-parameters that need setting, including detection thresholds that are required in most algorithms, sampling frequency, or eye tracker measurement noise. In addition, the parameters of the eye movements also need consideration. For example, the saccade amplitude (large saccades vs microsaccades) or velocity can make a difference in the detections. Also, lack of correctly labeled data, in cases of training models for automatic saccade or fixation detections is another important challenge (Schweitzer and Rolfs, 2020; Nyström and Holmqvist, 2010)

In velocity-based algorithms, eye samples are classified based on their instantaneous velocity. Using velocity thresholds, eye movements are classified as low-velocities for fixations and high-velocities for saccades. In the saccade detector used here, the time series of eye positions are converted into velocity values by using a FIR (Finite Impulse Response) differentiator of velocities over five eye data samples, as in Equation 3.1. This velocity calculation suppresses the noise in the eye tracking data (Engbert and Kliegl, 2003).

$$v_n = \frac{x_{n+2} + x_{n+1} - x_{n-1} - x_{n-2}}{6\Delta \mathrm{t}} \tag{3.1}$$

In my study, I used a velocity based saccade detection algorithm. I used an EyeLink 1000 [EyeLink Research Ltd, Oakville, Ontario, CA] to sample the user's eye positions with a sampling frequency of 1000Hz. The display update rate was 120 Hz, that is one frame every 8.33 ms. This means that I received one eye tracker sample every millisecond (Δt) and therefore I had 8 samples in every refresh update of the display. Hence I was able to perform a five point FIR differentiator for velocity computation. I computed velocity from horizontal and vertical components of eye sample positions. Then I scaled them by the instantaneous pixel per degree, which I calculated for each current eye sample. I applied elliptic thresholds to the horizontal and vertical velocity components using a median estimator (Engbert and Mergenthaler, 2006). I applied the scene change when more than three consecutive eye samples were detected outside the ellipse determined by the horizontal and vertical thresholds. By carefully adjusting the parameters, I suppressed noise and reduced false positives. I verified the reliability of our algorithm by running a pilot session and further adjusting the parameters.

3.2.4 Subjects

Ten users participated in this experiment, 5 female and 5 male, with average age of 25.4 years range [20-32]. Four of these participants habitually wore glasses for myopia, but participated without glasses, and could see clearly as the distance of the monitor was close to viewer's eyes. All participants were university students, had no prior experience in using an eye tracking device and were naive about the purposes of the experiment. Each participant signed a written informed consent form before starting the experiment, which was undertaken according to a protocol approved by the Human Participants Review Committee of York University. After completing the experiment, they received $10 financial compensation for their participation. Figure 3.4 demonstrates the setup of a user during the experiment.

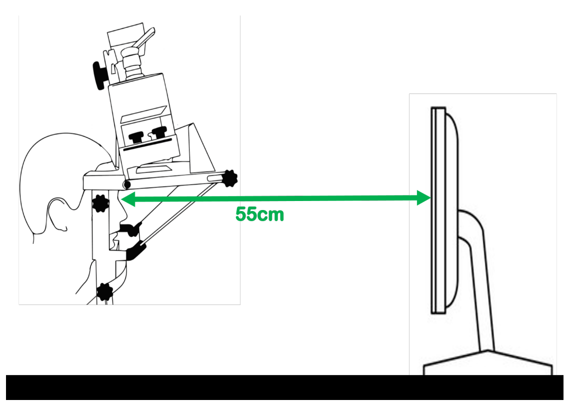


Figure 3.4: Subjects sat at a viewing distance of 55 cm from the display. (Display dimensions are not to scale)

3.2.5 Data Collection

The data collected for this experiment consisted of two separate data files for each participant. One of the files recorded data for the participant's eye movement features, such as eye position coordinates on display and the timestamps for each eye sample. From those features, I calculated other eye movement details like velocity. The other data file was the perceived scene motion responses I collected for each participant. These responses included records such as the trial number, subject's selected response for the task after each trial, and the correct response. These files were then used for data analysis and extracting the results.

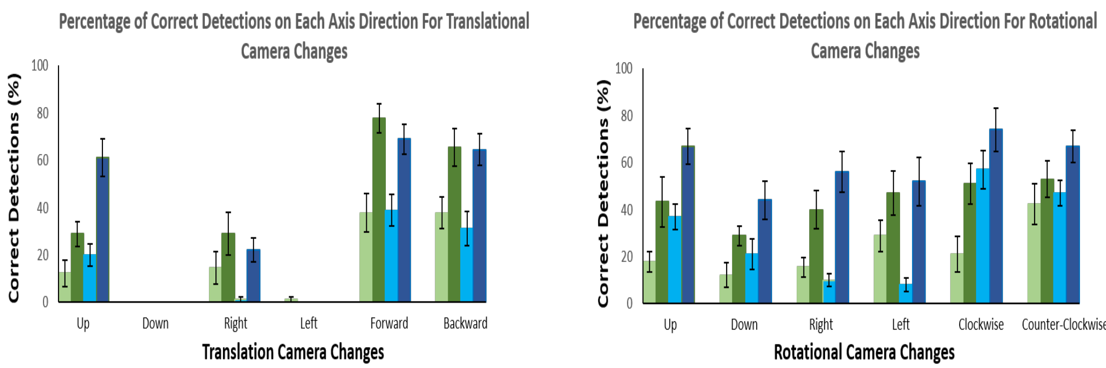
3.2.6 Results

For each participant and each trial, I recorded their reported camera movement with the corresponding confidence level. In my analysis of data, I looked at catch trials (with no camera movement) and scene-changed trials (with camera movement) separately.

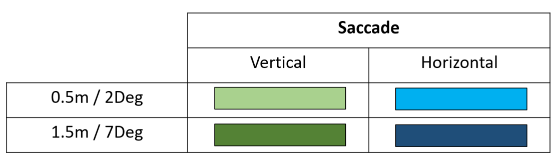
Scene-Changed Trials

Discrimination of Both the Transformation Axis and Direction: In Figure 3.5, I plotted the percentage of correct responses for discriminating the type and sign of transformation, as can be seen in Figure 3.5(a) for translational and Figure 3.5(b) for rotational changes. As there were 8 possible responses (Right, Left, Up, Down, Forward, Backward, Clockwise, Counter-Clockwise), participants should identify the correct direction on 12.5% of trials on average simply by chance. For small translations and rotations it is clear that performance was near or below chance levels.

Performance below chance levels can occur from random variation or when observers have biased responses. As these results show, for both directions of saccades, users found it difficult to differentiate Right and Left camera translations from each other. They also had difficulty discriminating Up and Down translations. However, Forward and Backward movements seemed easier to discriminate. As for rotational changes, the data showed that identifying the specific direction of movement on each axes was not as difficult as with translational changes, particularly for rotation about the roll axis.



(a) (b)



(c)

Figure 3.5: Percentage of correct detections, classified by correct discrimination of both axis and sign of camera movement, as a function of condition. The error bars show the standard error.

Discrimination of the Transformation Axis Regardless of Sign: Observers often reported that they had an impression of along which axis the change occurred, even if they were unsure of the direction. That is, most users reported they found it difficult to distinguish directions on the same axis (e.g. left versus right). Therefore in the main analysis below, I considered a participant's answer as correct if they identified the appropriate axis for the transformation regardless of the sign. Note that this is an easier task as there were only 4 possible outcomes (Left/Right; Up/Down; Forward/Backward; and CW/CCW). Thus chance performance corresponding to guessing predicts 25% correct, on average, for this analysis. In Figure 3.6, the average percentage of correct detections is shown, categorized by the actual directions of the camera movement (average of both sizes) during vertical and horizontal saccades to allow a more clear comparison. It also shows the average subjective confidence level of the correctly detected answers for each transformation. Rotational movements about the roll axis and also translations on the depth axis were easier to detect during both horizontal and vertical saccades than other transformations, and were associated with fairly high confidence levels.

The percentage of the correctly selected camera shifts is presented in Figure 3.7, grouped by directions and sizes of translations and rotations. As Figure 3.7 (b) shows, during both types of saccades, the translation of the camera on the depth axis had a significantly higher detection rate than translations on the horizontal ($\chi^2(3)=17.82$, $p<0.001$) and vertical axes ($\chi^2(3)=8.51$, $p<0.05$), and rotations along the roll axis were significantly easier to detect than rotations along the vertical ($\chi^2(3)=50.89$, $p<0.00001$) and horizontal axes ($\chi^2(3)=43.65$, $p<0.00001$).

The results show that for translation changes, the larger forward and backward updates were more detectable for the users, while the horizontal changes were de-

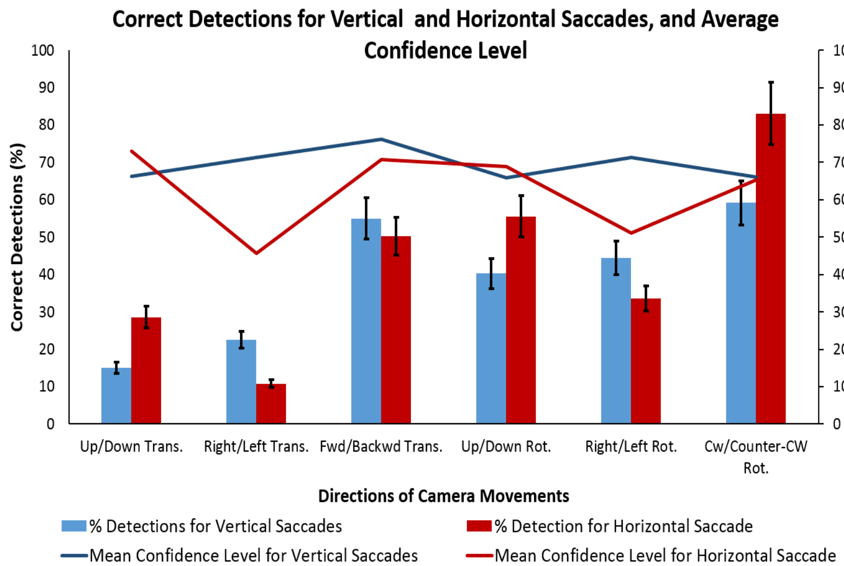
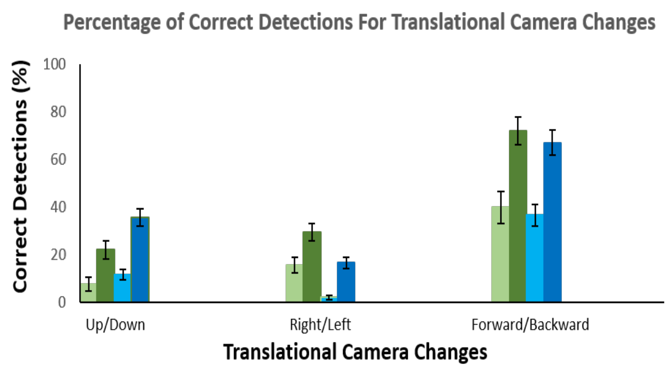
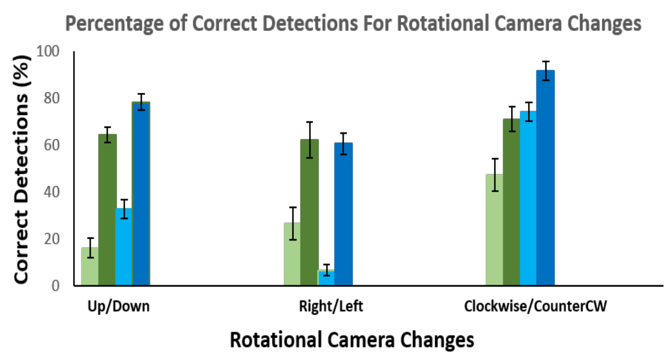


Figure 3.6: Average percentage of correct responses (correct axes regardless of sign) for the six types of translational and rotational eye movements during vertical and horizontal saccades, and the average level of confidence for correct responses for each transformation type.

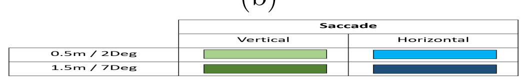
tected the least. For rotational display shifts, changes in the yaw were the hardest to detect. Rotations of 7° were easier for users to detect, than the smaller size ones. The users' levels of confidence increased as the size of translation and rotation increased. A Chi-Square test revealed that there was a significant difference in detection rate in the two different sizes of translational scene changes along each of the three axes, ($\chi^2(2)$=8.68, p=0.013). In addition, there was a significant difference in detection rate between different sizes of rotational scene changes along each axes ($\chi^2(2)$=48.4, p<0.001). Similar to horizontal saccades, during vertical saccades, there were significant differences in the detection rate between the two different sizes along each of the three axes, for translational camera movements ($\chi^2(2)$= 7.51, p= s0.234), and also for rotational camera changes ($\chi^2(2)$=15.5, p<0.001).



(a)



(b)



(c)

Figure 3.7: Percentage of Correct Detections for (a) translations and (b) rotations, grouped by direction of transsaccadic camera shift. Answers were classified as correct if they identified the correct axis, regardless of the sign. The error bars show the standard error.

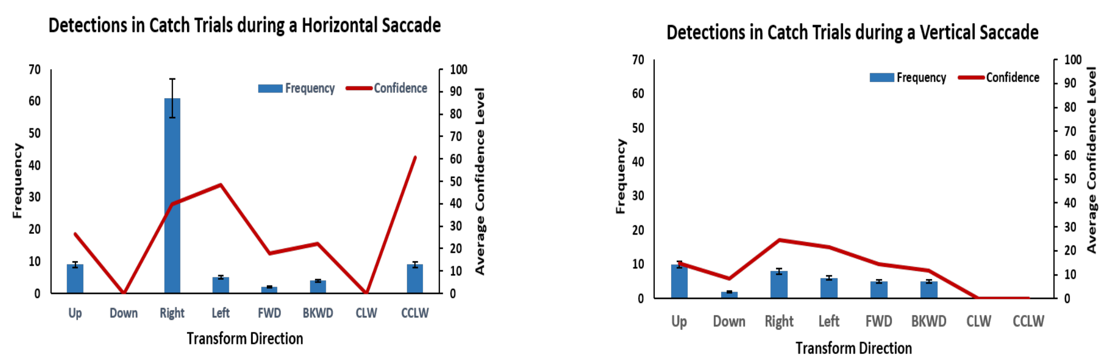
Confidence in Choice A t-test showed that users had a higher confidence level when they made correct responses (M=70.35, SD=19.58) compared to incorrect responses (M=17.57, SD=7.12), $t(9)=9.45$, $p<0.001$, $r=0.43$, $d=3.58$. Participants had an average confidence level of 76.37 for their correct direction responses for translational camera changes and 69.45 for rotational camera movements during vertical saccades. These values were slightly higher than the confidence levels for horizontal saccades (72.42% for correct translational and 67.29% for rotational direction responses).

Congruence of Change with Saccade Direction I found that detection rates for translational horizontal (Right/Left) movements were significantly higher during vertical compared to horizontal saccades ($\chi^2(1)=6.2005$, $p<0.0127$). Similarly, rotational horizontal changes were also correctly identified more often during vertical than horizontal saccades ($\chi^2(1)=16.5466$, $p<0.01$). This means that the detection rate of translational/rotational horizontal camera movements depends on the direction of saccade, as seen in Figure 3.6, and was lower when making a horizontal compared to vertical saccade.

For rotational vertical changes (Up/Down), there was also a significant difference between the two saccade directions ($\chi^2(1)=3.9925$, $p=0.0447$) although this was not the case for translational vertical movements ($\chi^2(1)=0.0261$, $p=0.8716$). This indicates that the correct estimation of Up and Down rotational (but not translational) movements of the camera depends on the direction of saccade, as seen in Figure 3.6, and in this case detection was more difficult when making a vertical saccade. Thus, in both the vertical and horizontal movement cases, correct discrimination was more likely when the saccade was orthogonal to the image transformation.

Catch Trials

Subjects were required to respond on all trials whether they saw a change or not. I defined an effective false alarm on a catch trial (where there was no change) as a response with a confidence level greater than zero. I found the false alarm rate, while low, was not zero. In fact, during horizontal saccades, in 90 (out of 932 across all participants) catch trials, participants answered with confidence levels of larger than zero, with average confidence level of 39.8%. This means that although there was no change in the position or angle of the camera, they still believed there was a change. Also, during vertical saccades, in 36 (out of 856 across all participants) catch trials, participants answered with confidence levels of larger than zero, with average confidence level of 15.8%.



(a) Horizontal (b) Vertical

Figure 3.8: Frequency of directions selected in catch trials, and the average confidence levels (over the non-zero responses) of users for different saccade directions.

As demonstrated in Figure 3.8a, during horizontal saccades, users selected the rightward scene change in 61 trials which is 4.29% of all the catch trials. But they never indicated downward and clockwise transitions. Figure 3.8(b) shows the user confidence levels for their answers in catch trials. Most users were not very confident about their guess of the scene change. However, in more than 35 catch trials, users made a guess with more than 60% confidence about their answers. Moreover, during vertical saccades users selected Up and Right changes most often in catch trials (Figure 3.8). This is consistent with Up and Right camera transformations being predominantly selected in the catch trials of horizontal saccades, as shown in Figure 3.8(a), although the preference for the rightward response was more pronounced in the horizontal saccade case. During vertical saccades, they also responded that the camera made Down, Forward and Backward movements during their saccades. However, they never selected Clockwise or Counter-Clockwise rotations. The confidence levels for selecting answers during these catch trials was comparably higher during horizontal saccades, than it was during vertical saccades.

3.2.7 Discussion and Conclusions

The results of this experiment show that the ability to identify changes in a 3D environment during a saccade depends on the type and magnitude of camera transformation. We confirm that large 3D corrections to the scene viewpoint can be introduced unobtrusively and with low latency during saccades, but the allowable amount of the correction varies with the transformation applied. This is consistent with previous research that showed transsaccadic transitions are suppressed for stimuli that filled a large area of the visual field (Currie et al., 2000; Allison et al., 2010). We showed how the detectability increases with the size of the camera shift. We used a fixed size saccade with a known direction (either horizontal or vertical). We looked at two directions of saccades (right and down), which are common in many tasks such as reading, for our sample population. We believe the results would generalize, but future work could consider other saccade directions or populations with different cultural biases. These controlled experimental conditions are important because depending on the user, task and the content of a VR display, the frequency and amplitude of visual saccades vary. To make sure that the users reliably generated a saccade, we had them follow a target object being displaced in the scene. The length of each trial was short to allow a high number of trials without being too tedious for participants.

As the task was related to saccadic suppression and the scene transitions occurred transsaccadically, in many trials the subjects were not able to see the camera movement and answered based on a guess or based on the edges of the scene after the saccade. Note that unlike an image update incorporated into the background processing of a VR application, our users were primed to expect an image transformation. The users made their 15 degree saccade consciously with a specific direction, and were focused and attentively looking for movements in the camera/scene that would occur at a known time (during the saccade). The probability of detecting a change will almost certainly decrease when the users make natural arbitrary saccades, are concentrated on another main task, and are not aware that a display update is likely. Therefore, we can say that the results presented in this paper are conservative estimates of the likelihood of hiding translational and rotational changes. In a number of large transition trials, some users reported verbally that they judged the camera movement direction by the skewed angle or shift of the scene after the saccade had ended. Although this is an indication that those particular transitions with those sizes were more detectable for the users, we should bear in mind that the users' sole task was to determine shift direction and hence in a more natural viewing VR when users are not aware of any camera shift, they may not notice these shifts after their saccades. Moreover, when wearing an HMD, the viewers have a large field of view of the virtual environment with weaker frame cues, which will make the difference between the before and after images less noticeable, particularly if the scene is changing or they are in motion. The results also showed that the most detectable changes were clockwise and counter-clockwise camera movements during both horizontal and vertical saccades. As discussed, these rotations change the apparent vertical position of the user (simulating a body tilt). Observers are quite sensitive to visual self-tilts (Howard and Childerson, 1994; Allison et al., 1999) and thus these changes are expected to be salient after the saccade, particularly with larger magnitudes, even if the change itself was not seen.

We found that users are less sensitive to certain image transformations during saccadic eye movements. For both horizontal and vertical saccades, users had a higher confidence level in their answers for translational changes as opposed to rotational changes. These results are consistent with previous research (Bolte and Lappe, 2015; Allison et al., 2010). Saccadic suppression of image displacement is larger with bigger saccades and smaller target displacements (Stark et al., 1976). Results of the current experiment show that when viewing a 3D scene, users are sensitive to scene transitions that occur during a saccade; and can more easily recognize direction of changes in which the camera makes large angles of rotation (7 degrees) or sizes of translation (1.5 m) as compared to smaller ones of 2 degrees rotations and 0.5 m translations.

During catch trials with no camera movement, participants still occasionally (on about 4-10% of trials) believed that they saw movements in the camera/scene. In such trials, participants often believed they saw horizontal changes as the object in the scene moved horizontally from left to right. This could be because users expected some amount of shift in the image with their eye movement, or because they were less certain that they did not see small image shifts in this direction on top of the large retinal image shift produced by the eye movement. As the object in the scene and users' eyes moved horizontally to the right, users may have interpreted the motion of the object as image shift. When a viewer and a fixation stimulus experience simultaneous acceleration, the fixated stimuli seem to move in the direction of acceleration despite of no physical movement relative to the observer. So during self-motion, fixation on a stationary environment results in perceived object motion (Whiteside et al., 1963). In fact these may be "illusional movements" which are a part of autokinetic effect caused by a short-term imbalance of the neural systems directly concerned with the visual registration of movement (Gregory and Zangwill, 1963). Overall, since the participants were given a forced-choice question to guess a direction, they seemed to act very cautiously and tried to detect any movements in the scene and make as many correct direction guesses as possible and attributed them to the noisy horizontal direction when uncertain or unseen.

We translated or rotated the user's viewpoint during saccadic eye movements. They did not notice most of the reorientations (rotations) of 2 degrees and translations of 0.5 meters during saccades. However, it was different for clockwise and counter-clockwise rotations of the camera along the roll axis, as these were more obvious, changed the simulated standing angle of the user, and felt very unnatural. These camera shifts also have larger motion in the periphery. It is worth noting that the participants may not have noticed the camera shift during the saccade itself, but guessed the direction of camera change correctly after the saccade had ended. This is because when the camera moves clockwise or counter-clockwise it is much easier (than other directions) to detect its change of position because the image rotated relative to the display. This could also be the case for large rotational changes by noticing parts of the image shifting on or off the display. The users mentioned that even though they did not see the image shift, they could guess the direction of its change correctly according to the image they saw after their saccade had ended. This result might be different with a larger field of view or when wearing an HMD where frame cues are typically weaker, as these before and after image changes should be much less detectable, especially when the users are not aware that there are any display updates.

During vertical saccades, subjects had a weak predisposition to report these transsaccadic movements to be Up translations or rotations. This is the same axis that the target object was displaced (from top of the scene to the bottom) and also in the direction opposite to the saccade. The reason for this may be users relating the movement to the retinal image shift when unsure about direction. Also, they might have interpreted the displacement of the object as a scene movement. This tendency appeared to be much stronger for horizontal saccades where the majority of responses during catch trials were in the Right direction.

The results obtained in this research can be used in designing applications that require imperceptible graphical updates as in gaze contingent displays. In a redirected walking application, rotation gains can reportedly be adjusted between 5 and 30% without being detected (Nilsson et al., 2018). Therefore assuming a 20% gain limit on a 20° virtual rotation, we could introduce about 4° distortion or use a 16° real rotation. If we could hide another 2° yaw change during a saccade, then we could effect a 30% change. On smaller rotations, the effect could be proportionally larger and on larger rotations we might expect multiple saccades. Similarly, suppose we were traveling in a vehicle at 100 km/h (27 m/s) in a collaborative virtual environment (CVE) consisting of multiple simulators, each maintaining local dead reckoning models of vehicle state (for example using IEEE 1278.1-2012 Annex E or IEEE 1516-2010). If we assume that a saccade occurs approximately every 300 ms then we could instantaneously correct for network related inconsistency errors of up to 6% during a saccade while keeping absolute size of the corrections below 0.5 m. Tolerance to change during such fast movements is likely to be much larger than during the viewing of static scenes like the ones used in our experiment. In other applications, saccadic masking of camera shifts could allow for subliminal diversion of a user's attention to a specific location, or even to guide navigation tasks in VR. The present research suggests that the tolerance to such transsaccadic updates depends on their magnitude, direction and transformation type as well as the direction of the eye movement. Previous research shows that it also depends on the size of the saccade (Allison et al., 2010). Gaze-contingent hiding of graphics updates is an application of eye tracking that can be used to improve the design of 3D virtual environments. Accurate eye tracking technology like that employed in the current study is increasingly being integrated into VR displays and provides a framework for producing such displays as well as other interactive virtual reality and gaze-contingent applications. Based on the results of this research, saccadic suppression is an applicable tool for hiding graphics updates when users view a 3D virtual setting through an eye tracker. As seen, there are image transformations in certain directions that are more apparent and recognizable for the viewers, such as rotations along the roll axis.

The results are consistent with the previous research (Bolte and Lappe, 2015; Allison et al., 2010). Saccadic suppression of image displacement is larger with bigger saccades and smaller target displacements (Stark et al., 1976).

The desktop-based findings can be generalized to immersive settings. One important reason for choosing a desktop-based eye-tracker is that it provides more reliable, accurate gaze data with a high sampling rate, as opposed to the HMD integrated eye-trackers with a sampling frequency nearly ten times lower. The fixed desktop display had a smaller FOV than many VR systems (although similar to most AR systems) and we restrained the head. The fixed head mode allows for much improved tracking and was appropriate to focus on saccadic suppression. Our study assessed the extent of changes that can be tolerated when observers are still and most sensitive. Thus it provides valuable baseline performance, obtained under better controlled experimental conditions that is well-suited for application by researchers and developers of VR, not to mention other 3D contexts such as gaming and fish-tank VR. In the application context, our gaze contingency technique may be more effective, because with a large FOV or when wearing an HMD, these before-and-after image differences should be less detectable. In this case it would only matter if users notice the camera change and how disruptive they find it. In most cases the users will not be aware that there were any display updates. Even in an HMD setup there is a peripheral reference from the edge of the field of view, these references are located well away from typical fixation points and the large degree of saccadic suppression found here suggests they will be relatively ineffective.

3.3 Experiment II: Effect of Head Motion on Saccadic Suppression

Often in 3D games and virtual reality, changes in fixation occur during locomotion or other simulated head movements. In the following experiment, we investigated whether a constant camera rotation in a virtual scene modulates saccadic suppression. We studied two directions of camera motion; horizontal (camera tilt) and vertical (camera pan).

3.3.1 Hypothesis and Objective

The aim of the current research was to investigate whether an on-going camera motion in a scene has any effects on how the users perceive transsaccadic image transformations. I designed an experiment to find the threshold values of image transformations (being scene translations and rotations imposed on the camera showing the scene) that were imperceptible, during an ongoing camera rotation. I looked at canonical camera shifts corresponding to 3 translation directions (horizontal, vertical, in-depth) and 2 rotation directions (yaw and pitch) during either horizontal or vertical saccades. These transformations were applied during one of three overall motion conditions:

- Constant velocity camera pan, or
- Constant velocity camera tilt, or
- No camera motion (static).

The main hypothesis here is that head rotations that are along the same rotation axis as the image transformations will be better hidden from the subjects, compared to ones in other directions.

3.3.2 Stimuli and Apparatus

My stimuli were rendered from a 3D scene showing a view of a street with two streetlights and some buildings. We used the exact same scene with the same camera view for both horizontal and vertical saccade trials. Each trial was two seconds in duration. The virtual scenes were designed in Unity3D and the scripts were written in C#. They were presented to the participants on a 27 inch 3D Samsung LCD monitor, with a 120 Hz refresh rate and a 1920H*1080V resolution. Although the monitor is stereoscopic, I did not use this feature. The processor used for rendering the 3D scenes in real-time consisted of an AMD FirePro W9000 FireGL, Intel Core i7 CPU 3.50 GHz, 32 GB RAM and a 64 bit Win 7 OS. To record the users' eye

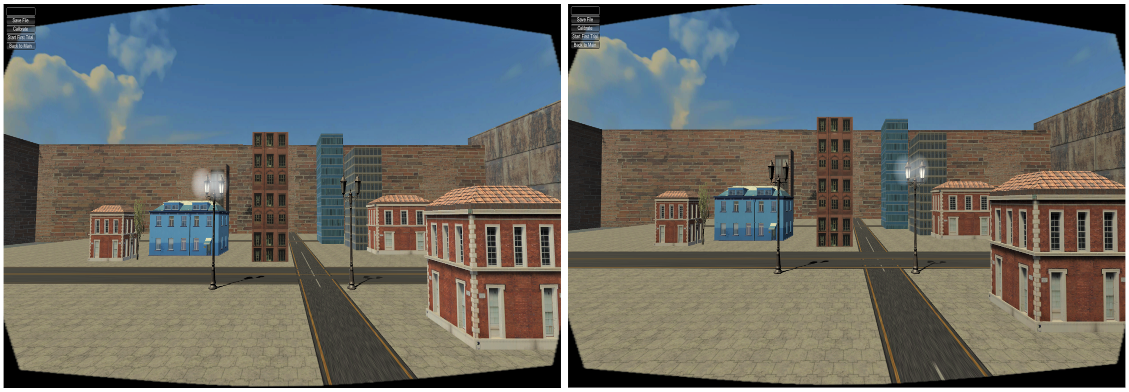


Figure 3.9: The virtual scene used in the experiment. For horizontal saccade trials, the user fixated on the left image first, and made a saccade to the lamp post on the right image when it turned on.

movements in real-time, the video-based Eyelink 1000 [SR Research Ltd, Oakville, ON, Canada] eye tracker was used in its tower-mounted mode, with a chin-rest for ensuring the stabilization of the participants' head. The eye tracker was placed at a fixed viewing distance of 55 cm from the display. We placed mechanical jacks beneath the display to adjust its height to center at eye level for each viewer.

Trials Stimulating a Horizontal Saccade

Upon the start of these trials, the participants fixated their gaze on an illuminated lamp post on the left side of the screen. After one second, the left lamp post turned off and and at the same time the lamp post on the right side turned on, as shown in Fig. 3.9. The participants were instructed to move their eyes to the lamp post on the right side of the scene when the left light turned off and the right one turned on. This resulted in them producing a 12° horizontal saccade.

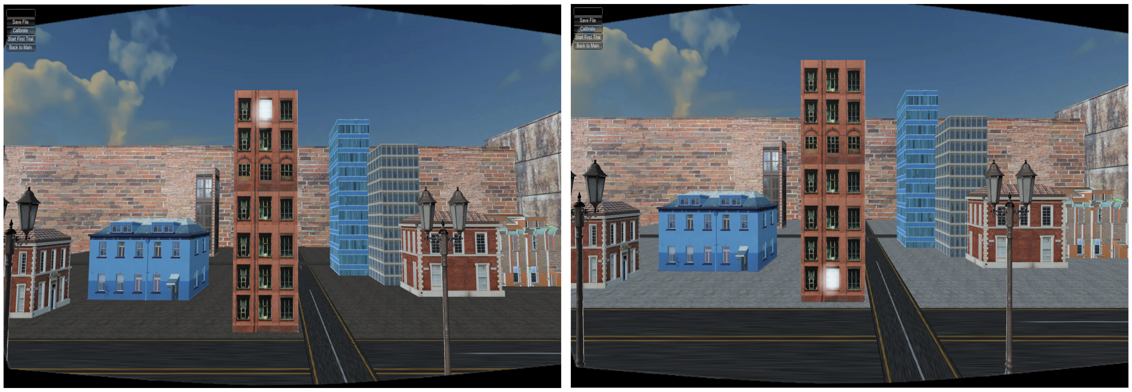


Figure 3.10: The virtual scene used in the experiment. For vertical saccade trials, user fixated on the lights on the upper floor (left image) first, and made a saccade as the bottom floor lights turned on (right image). The snapshots were taken from scenes with the constant vertical camera motion.

Trials Stimulating a Vertical Saccade

In these trials the users were triggered to generate a vertical saccade. A light in a top story of a building was turned on (Fig. 3.10 left) and the participants were asked to fixate their eyes on it upon the start of each trial. Once it turned off and a light in the bottom story turned on (Fig. 3.10 right), the participants moved their fixation to the newly illuminated window, making a vertical saccade. Compared to the horizontal saccade scenes, the camera showing the scene was slightly moved in to guarantee a stimulation of a saccade of 12° in magnitude.

3.3.3 Experiment Design and Procedure

My study consisted of three motion conditions, for which I recruited different groups of participants. In group I of our experiment, the constant camera motion in each scene was a pan (left to right) rotation of 0.04 degrees per frame, which is equivalent to 4.8 deg/sec at our frame rate of 120Hz. In group II, the ongoing camera motion in each scene was a tilt (top to bottom) rotation with the same speed. I chose this value as this made a smooth camera motion in the scene and could simulate a constant head rotation in a VR, while it was not too fast to interfere with the purpose of our experiment. In the last group of our experiment, there was no constant camera motion during the trials (static camera). Figure 3.11 shows a participant setup of this experiment.

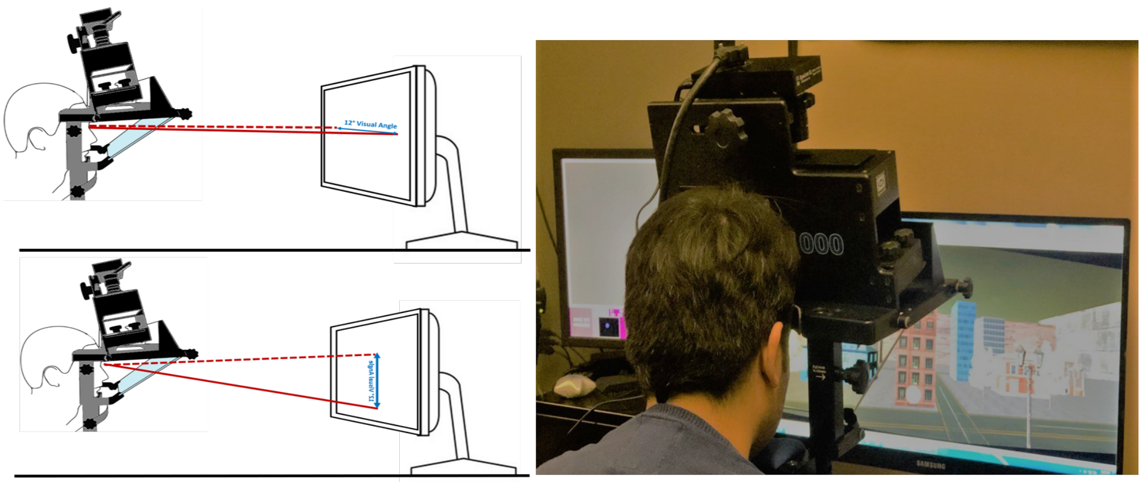


Figure 3.11: Experimental setup. Left: Sketch of user producing a horizontal saccade (Top) and a vertical saccade (Bottom). Both saccades are of size 12°. Right: A participant taking part in the experiment. Room lights were extinguished for the actual experiment.

Adaptive Methods for Threshold Estimation

Adaptive threshold estimation methods are used often in psychophysical experiments to determine the value of a single point or set of points on a psychometric function (Levitt, 1971). While these procedures run, the stimulus level is adaptively increased or decreased during each trial based on the user's response to the stimuli. With the progress of the trials, the stimulus level converges to a single value.

There are different adaptive methods for estimating thresholds. The classical non-adaptive methods for psychophysical measuring such as the method of constant stimuli (Simpson, 1988), present a number of fixed stimuli in a random order many times. For each stimulus, participants perform the same task. These methods are costly as many trials are necessary. Having many trials is not preferred as it would take a long time, yield learning effects and it can affect the participant's concentration on the task, and also the amount of data collected. On the other hand, adaptive methods that use the maximum likelihood procedures include the Parameter Estimation by Sequential Testing (PEST) (Lieberman and Pentland, 1982; Pentland, 1980) and QUEST (Watson and Pelli, 1983), both calculating the optimal threshold value for the next trial while the experiment is being run. PEST assumes a logistic underlying function and is a variant of the weighted step-size method. It takes larger steps at the beginning and changes its step size as it runs process. Quest is a Bayesian version of the PEST and estimates the threshold based on a Log-Weibull function. The down/up staircase (Levitt, 1971), is another adaptive method that uses user's responses in each trial to step the threshold up or down. It stops after a given number of trials or transition points. In Madigan and Williams (1987), the authors found that in a Yes-No and 2AFC task PEST was less efficient than QUEST but that QUEST was more accurate (Madigan and Williams, 1987). Also in Otto and Weinzierl (2009), it was shown that the PEST method requires a high number of measurements to provide a high accuracy estimate (Otto and Weinzierl, 2009).

The quest (Watson and Pelli, 1983) method is an adaptive inference procedure that relies on Bayesian statistics. It uses prior knowledge to estimate Bayesian probability in Yes-No or Forced-Choice tasks. Although it requires a more complicated procedure, it is very efficient and accurate as it considers the user's responses on each trial. It can lead the experimenter to the estimated threshold in a lower number of trials compared to the other methods. In quest, the data collected during each trial is used to fit a psychometric function to them. This fitted function is used for selecting the level of intensity or threshold for the next trial. The Weibull function is good at fitting participants' data (Klein, 2001). However, the log-transform version of it, known as the Gumbel function, is often used with the quest method (Kuss et al., 2005).

To measure the thresholds of imperceptible scene movements during saccades, we implemented and used a quest algorithm (Watson and Pelli, 1983). For quest to be used in Unity, we implemented the toolbox in C#, which we were able to integrate into the main C# scripts of our experiment. It is also available as a separate package and can be used with any other C# script. Our implementation is based upon and follows the same procedure as the quest algorithm in the MATLAB Psychtoolbox. The quest algorithm uses a Weibull function as the default psychometric function (Watson and Pelli, 1983). This function has a number of parameters that we set so that our quest produced the correct threshold values and converged to the desired estimates. To set the parameters of our probability density function (pdf), we selected a value of 3.5 for β, 0.01 for δ, and 0.5 for γ. The β parameter controls the steepness of the slope of the psychometric function, δ is the lapse rate and γ is the guess rate. As a probability threshold we used 0.82. We selected an initial threshold estimate guess (denoted as x in equation 3.2) of 0.5 m for translations and 0.7° for rotations. We used the quantile of the posterior probability density function to recommend the next testing value.

$$f(x) = \delta\gamma + (1-\delta)[1-(1-\gamma)e^{-10^{\beta x}}] \quad (3.2)$$

Procedure

The participants were given guidance on their task in a five minute training. They were shown examples of the types of camera transformations to expect, without being told whether they were rotations or translations. They completed the trials over two block, one with horizontal saccades and one with vertical saccades.

To begin each experimental block, participants performed a 9-point eye tracker calibration procedure. If a failure occurred, we repeated the calibration steps. They then started the first trial of the block. Each block took approximately 25 minutes. Participants could take a break between the two blocks to avoid any eye strain or loss of concentration. As the first trial began, we initialized two interleaved Quests one for each of same-axis translational directions (for example Right/Left). The direction of these camera transitions was selected randomly on each trial so the user could not anticipate what to expect. When an estimated threshold was reached for both of these directions, two new Quests were started for the two new directions. The same procedure was executed for all translational and rotational directions tested. This process was repeated for the second block.

After each trial, a graphical dialog was presented to the users to respond on a 2-alternative forced choice task. We used a 2AFC task to measure the threshold of imperceptible camera movements during saccades. This is a general approach in psychophysical experiments (Langbehn et al., 2018; Langbehn et al., 2016). After each two-second trial, the question of "Which Direction did you notice a shift in the camera?" was displayed on a blank background. They were given two choices, one of which was correct and they had to click the button of the direction they thought the change occurred. Participants were instructed to guess the direction if they did not notice any changes, which would still result, on average, in 50% correct responses. Depending on the correctness of their response, the relevant Quests updated their values automatically and the next trial appeared for the user. The participants did not know whether it was a translational or rotational change. In trials where there was no saccade detected, the quest did not update its values. At the end, each running quest provided a final threshold estimate for a given condition. Different criteria can be used to stop a running Quest, of which a predefined number of trials or reversals are the most common ones. We used the standard deviation of the posterior probability density function (pdf) as our stopping criterion for a running Quest. We set this value to 0.1 m and 0.09 degrees for translational and rotational changes, respectively, to ensure the quest converged appropriately on its estimation of a final threshold. In addition to that, a running quest had to meet another criterion before stopping and that was reaching a set number of reversals, which we set to 5. The combination of both criterion ensured our quest ran for enough trials before stopping. For each trial, a threshold was suggested by the quest adaptive estimation algorithm based on the participant's responses in the previous trials. The user was unaware how the values of camera transformations changed while the experiment ran. On a given trial, the user was presented with a stimulus based on the current threshold estimated by Quest. The estimated threshold value was then updated by the quest according to the participant's response. This way quest converged to the threshold value.

For detecting a saccade I used a velocity based algorithm similar to the one used in previous experiment presented in Section 3.2.4.

3.3.4 Subjects

We recruited 36 participants (three groups of twelve subjects as described below). In all three groups, every participant took part in two experimental blocks, either with horizontal saccades or vertical ones. The difference between groups lied in the continuous motion of the camera. The participants were naive to the scientific purpose of the experiment and were all students at York University. They volunteered through the Undergraduate Research Participant Pool and gained course credit for their participation. Prior to beginning the experiment, they signed an informed consent form which was approved by the Human Participants Review Committee of York University.

Group I: Continuous Camera Pan

Twelve students participated in these experimental blocks of the experiment with a camera pan in all trials. They included 9 females and 3 males. The average age of the participants was 21.58 years old [18-47], nine of them reported that they had myopia, however only five of them wore glasses for the experiment while the rest stated they could see clearly and participated without glasses. None had previous experience with using an eye tracker.

Group II: Continuous Camera Tilt

A different group of twelve individuals (4 males and 8 females) participated in this condition, where there was a camera tilt in all trials. The subjects had an average age of 20.33 years [18-25], and all were naive with using an eye tracking device. One of them wore glasses for the experiment, two wore contacts and the rest participated without glasses.

Group III: Static Camera

Another group of twelve individuals (6 female) participated in this condition, where there was no continuous camera motion in any trial (static). The average age of this group of participants was 21.58 years [18-32]. Only one of the participants had previous experience with eye tracking devices and the rest of them were naive with them. Five participants had normal vision without optometric correction, two of them wore contacts for the experiment and the other participated with glasses.

3.3.5 Data Collection

I collected gaze data for this experiment, as well as user responses, and QUEST estimation values for each condition. The gaze data included the participant's eye movement features, such as eye position coordinates on display and the timestamps for each eye sample. From those features, I was able to calculate other eye movement values like velocity. The other data file was the perceived scene motion responses I collected for each participant. These files included records such as the trial number, subject's selected response for the task after each trial, and the correct response. These files were then used for data analysis and extracting the results.

3.3.6 Results

We analysed the data separately for each transformation type in terms of two factors and their interaction. The first was the between subjects factor of constant camera motion and had three levels: pan, tilt and no motion. The second factor was a within subjects factor of the saccade direction which had two levels of horizontal and vertical. In some conditions, the effects of saccade direction could also be analyzed in terms of whether they were parallel or orthogonal to the direction of camera motion. We used MATLAB to perform our statistical analysis.

Though participants were instructed to respond to the object displacement by making a 12° saccade, they sometimes made other saccades or blinks during the trials as well. In our analysis, we excluded the trials in which subjects did not make 12±0.5° saccades, or were not fixating, or blinked at the time of the required saccade. Figure 3.12 shows a fixation map of a sample trial where more than the required saccade were made. There were small saccades away from fixations before the time of the required saccade. The actual expected vertical 12° saccade was also produced upon change in lights, and this trial was included in our data analysis. On average the participants' saccade size in response to stimuli movement, was 11.96°˙

Initially, we performed a series of two-way mixed ANOVA analyses to look at the interaction and main effects of camera motion direction (with three levels) and saccade direction (with two levels) on each of the possible transsaccadic change types. In none of the analyses was the interaction of camera motion and saccade direction

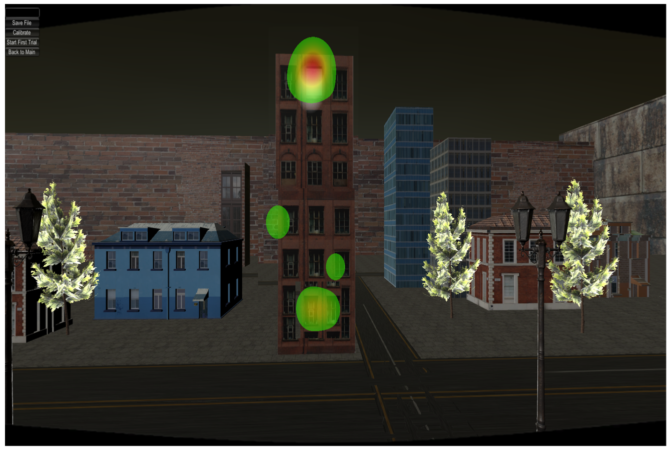


Figure 3.12: A fixation heatmap of a trial, during which the participant made the expected vertical saccade, as well as making other unexpected saccades.

significant. The results revealed that there was a main effect for the direction of the continuous camera motion on the detectability of right ($F(2,66)=10.64$, $p=0.0001$), left ($F(2,66)=3.64$, $p=0.0317$), and down ($F(2,66)=3.65$, $p=0.0313$) translations, but not up, forward or backward translations.

The direction of the constant camera motion during the trial also had a significant effect on the detection thresholds for up ($F(2,66)=5.54$, $p=0.006$), down ($F(2,66)=7.1$, $p=0.0016$), and left ($F(2,66)=9.86$, $p=0.0002$) camera rotations during saccades.

Saccade Direction

In different blocks, participants were triggered to produce either a horizontal or vertical saccade. When we compared the threshold values obtained during scenes with a continual camera pan, there were no significant differences in sensitivity during horizontal compared to vertical saccades. Similarly, the detectability of transient camera shifts were not significantly different between horizontal and vertical saccades in scenes with a steady camera tilt.

Continuous Camera Motion

We were interested in determining how an ongoing camera motion, which simulated a user's head rotation in a VR, affected the detectability of transsaccadic changes. The direction of camera motion had a significant effect on right, left and down transsaccadic translations. The sign of the effect depended on the direction of the transsaccadic change as can be seen in Figure 3.13 and Figure 3.14. Leftward and rightward translations were less detectable (had higher thresholds) when the ongoing motion was also a camera pan than when it was a camera tilt. conversely downward translations were less detectable when the ongoing motion was also a tilt than when it was a pan. For upward translations a similar pattern held but as noted above the effect was not significant. These patterns held during both vertical and horizontal saccades.

As stated above, the direction of ongoing camera rotation also affected detectability of all four types of rotational transsaccadic changes, as can be seen in Figure 3.13 and Figure 3.14. Up or down pitch rotations during the saccade were harder to detect when the ongoing motion was also a tilt compared to when it was a pan ($F(3, 92)=6.41$, $p=0.0005$). Conversely, left and right yaw transsaccadic camera rotations were harder to detect when the ongoing motion was also a pan compared to when it was a tilt ($F(3, 92)=10.13$, $p<0.0001$).

During horizontal saccades, with a constant rightward yaw camera motion in the scene, the right translational camera jumps during saccades were significantly less detectable ($F(1,23)=7.09$, $p=0.0096$). Right rotations were significantly harder to detect (F(1,23)=10.64, p=0.0036) as opposed to the scenes where the ongoing camera direction was downwards, and so were left rotations (F(1,23)=7.1, p=0.0142). This is also evident by comparing the corresponding bars in Figure 3.13.

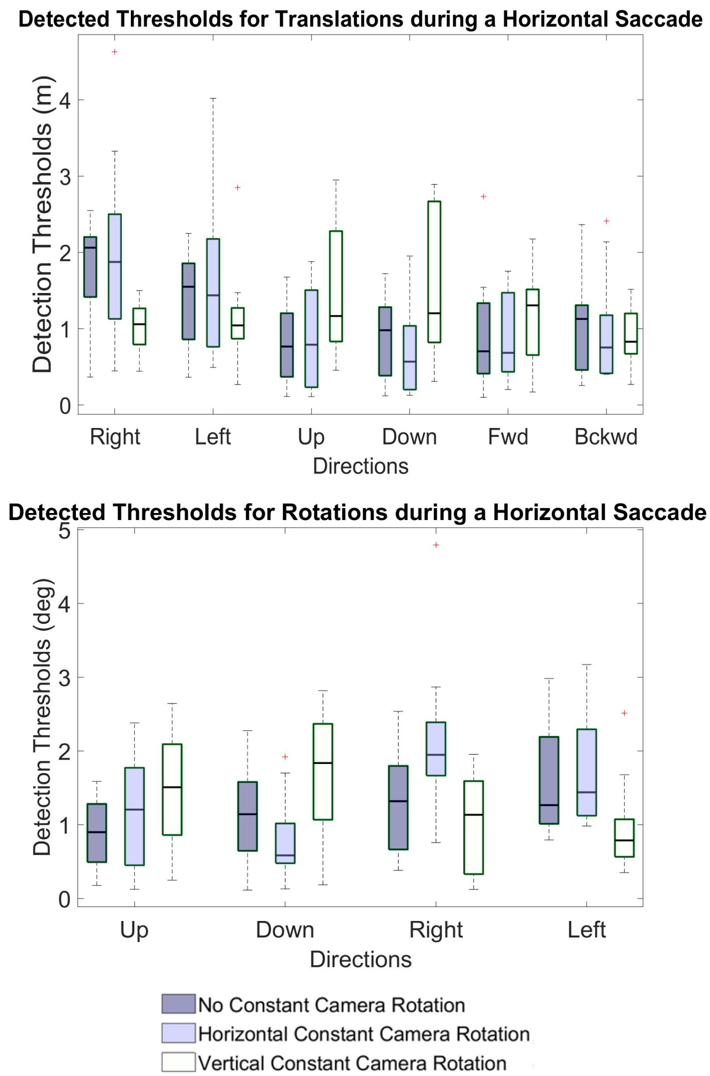


Figure 3.13: Detected thresholds of transsaccadic camera shifts during a *horizontal* saccade for scenes with different directions of a constant camera rotation, Top: Translational Changes, Bottom: Rotational Changes

We also compared differences of the values obtained for image transformations during vertical saccades, between trials with a constant yaw and a constant pitch camera motion. We found that during scenes with a camera pan, right translational image shifts during saccades were less detectable (F(1,23)=10.34, p=0.004) than during camera tilt (although the saccade direction was vertical). This shows additional evidence for the strong effect of a constant camera motion in a scene, simulating a head rotation, compared to saccade direction.

In addition, during scenes with a constant camera pan, up transsaccadic rotations were easier to detect (F(1,23)=7.46, p=0.0122) compared to scenes with a camera tilt. Also, right rotations (F(1,23)=4.38, p=0.0481) were harder to detect when the ongoing camera motion in the scenes was horizontal (pan) compared to when it was vertical (tilt).

Furthermore, for horizontal saccades, a one way between-subjects ANOVA was conducted to compare the camera shift detectability during three different conditions of yaw, pitch and no camera motions. We found the differences in these groups to be significant for right translations (F(2,33)=4.63, p=0.0169), down translations (F(2,33)=4.76, p=0.0153), and down rotations (F(2,33)=4.97, p=0.0130). We performed the same statistical analysis for vertical saccades. The detectability of right transsaccadic translations (F(2,33)=7.83, p=0.0017) and left translations (F(2,33)=4.83, p=0.0145) in all conditions of constant camera motion was higher, as well as left rotational camera shifts (F(2,33)=8.60, p=0.0010).

Scenes with no constant camera rotation versus scenes with a continuous camera rotation

We compared the difference between threshold values obtained during trials with a pan/tilt camera motion with blocks in which the scenes had no camera motion.

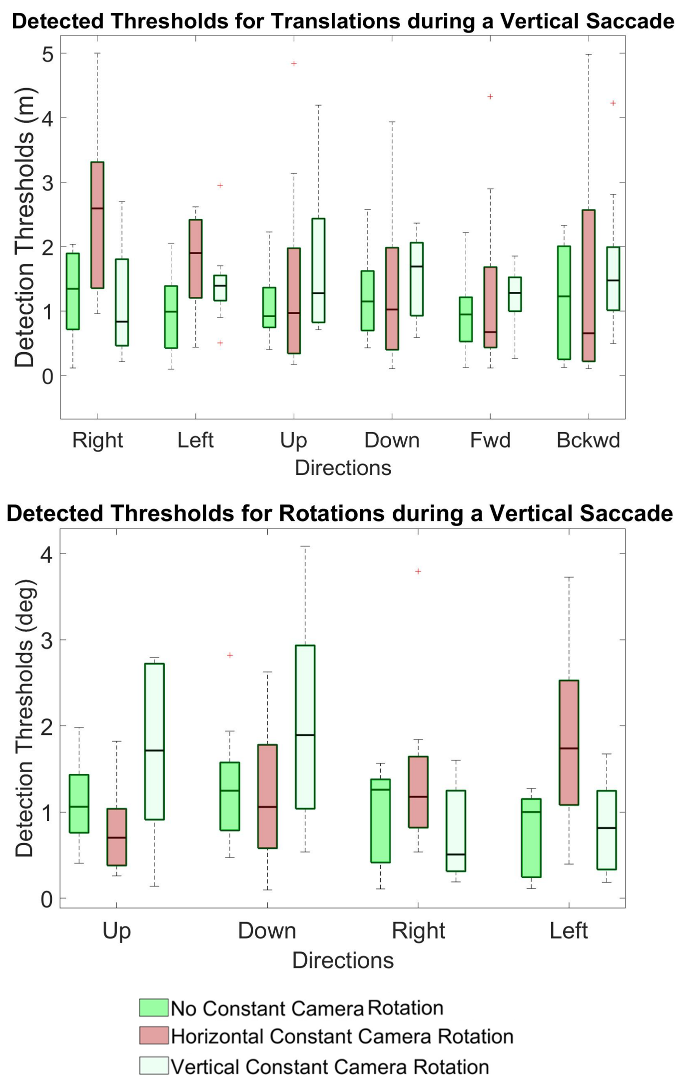


Figure 3.14: Detected thresholds of transsaccadic camera shifts during a *vertical* saccade for scenes with different directions of a constant camera rotation, Top: For Translational Changes, Bottom: For Rotational Changes

Horizontal Saccades When we compared between-subjects threshold values obtained during horizontal saccades, between static scenes and scenes with a constant camera pan, only right yaw image rotational shifts were more easily detectable (lower thresholds) during static scenes ($F(1,22)=5.57$, $p=0.0207$). We also found that up translations ($F(1,22)=4.58$, $p=0.0384$) and down translation ($F(1,22)=5.04$, $p=0.0352$) were harder to detect during trials with a constant downward camera tilt compared to a static camera. Results can be seen in Figure 3.13.

Vertical Saccades In trials with a vertical saccade, right transsaccadic translations ($F(1,22)=9.70$, $p=0.0051$) and left transsaccadic translations ($F(1,22)=8.7$, $p=0.0054$) were significantly harder to detect when there was a constant camera pan in the scene compared to when there wasn't any camera motion. Figure 3.14 demonstrates this.

Angle between constant camera motion and saccade direction

We were interested in finding whether the angle between the direction of the ongoing camera motion in a scene and the saccade direction influenced image displacement detectabilities. We compared angles of zero and 90 degrees.

In the parallel case, when the continuous camera motion (pan) and the saccade direction were both horizontal, we found higher thresholds for transsaccadic horizontal (right/left) translations compared to vertical (up/down) translations ($F(3,43)=5.06$, $p=0.0043$). Further, when both the continuous camera motion (tilt) and the saccade were vertical, transsaccadic pitch rotations were significantly less detectable than yaw rotations (F(3,43)=6.23, p=0.0013).

In the orthogonal case, when the continuous camera rotation in a scene was a pan (horizontal) but the saccade was vertical, horizontal transsaccadic rotations were significantly less detectable than vertical transsaccadic rotations (F(3,43)=3.45, p=0.0243). Thus, in agreement with the main analysis, the direction of camera motion seems to be a more important factor in determining the detectability of the transsaccadic changes than the saccade direction (either in absolute terms or relative to the constant camera motion).

3.3.7 Discussion and Conclusions

In a real scenario in a virtual environment, a user moves freely in different directions, making many combinations of head and eye movements. In our experiment, a steady camera motion simulated head motion in either a right yaw or a down pitch rotation. We also ran static trials without any constant camera rotations.

In our study, the saccades were triggered in response to change in the stimulus (reactive). Such reactive saccades are common to the onset of stimuli and are useful experimentally because they can be elicited consistently with reliable timing (Gremmler and Lappe, 2017).

Our gaze-contingency experiments revealed that the direction of a constant-velocity camera rotation running through the whole trial, affects how users perceive different transformations applied to the image of the environment during a saccade. Our analysis indicates that the direction of the ongoing camera motion in the scene has a stronger effect than the saccade direction on the transsaccadic detectabilities.

We found that when there was a constant rightward pan of the virtual camera in the scene, the users were less sensitive to horizontal scene changes during a saccade, and with a smooth camera tilt, the vertical transformations were harder to notice. This is true for both saccade directions we studied. It is interesting to see that in most cases, scene shifts parallel to the camera motion had a higher estimated threshold compared to the other rotations. Our findings lead us to conclude that when users make horizontal head rotations in a virtual environment, they have a better tolerance of additional horizontal rotational transsaccadic scene changes whether their saccade is horizontal or vertical. Likewise, when making vertical head movements (simulated as a camera tilt), vertical transsaccadic image displacements should be better hidden from the users for both horizontal and vertical saccades. This insensitivity during head motion will allow designers to leverage the additional suppression that occurs when the image displacement is in the same direction as the head moves. It also suggests further investigations to explore how well these findings generalize to other combinations of image displacements, head rotations and eye movement directions.

Saccadic suppression can mask the motion-blurred images produced during a gaze shift from one position to another. Our study revealed how this suppression can differ when more variables (i.e. head rotation) are added to the viewing environment. According to previous studies (Allison et al., 2010), we expected users to have a higher tolerance for horizontal transsaccadic camera shifts during a continuous (camera pan) and horizontal saccades. Also, we were expecting higher threshold values for vertical transsaccadic camera transitions during trials in which both the continuous camera motion and the saccade direction were vertical. Our results confirmed these assumptions for head motions but we found little effect of saccade direction. Our results are consistent with previous studies (Allison et al., 2010) and further show how adding a dynamic variable (the constant camera pan or tilt) can influence the tolerance of VR users to transsaccadic image displacements. In the current study, we restrained the head and simulated head motion through the virtual camera. This allowed for better control of the simulated head motion making interpretation of the data clearer. In an HMD-based experiment there would be less control of the conditions. Furthermore, an important advantage of the eye-tracker used in the present experiments is a high sampling rate of up to 2000 samples per second as opposed to the 90 or 120 samples per second available with HMD-integrated eye trackers. The high sampling rate helps in saccade prediction and reducing gaze-contingent latency. However, it is important to ensure that our results translate to real application and actively generated head motion. Therefore, the next experiment, that I discuss in Chapter 4, was designed and conducted to validate the current findings.

3.4 Summary

In this chapter I reported two experiments that I conducted to investigate the utility of saccadic suppression in hiding graphical updates. I studied how visible different image transformations, including scene translations and rotations, were to participants during their saccadic eye movements. Users viewed 3D scenes in which the displacement of a target object triggered them to generate a vertical or horizontal saccade. During the saccade a translation or rotation was applied to the virtual camera used to render the scene. After each trial, users indicated the direction of the scene change in a forced-choice task. The results of the first experiment revealed that type and size of the image transformation affected change detectability. During horizontal or vertical saccades, rotations along the roll axis were the most detectable, while horizontal and vertical translations were least noticed. We confirm that large 3D adjustments to the scene viewpoint can be introduced unobtrusively and with low latency during saccades, but the allowable extent of the correction varies with the transformation applied.

In the second experiment of this chapter, we studied how a simulated head motion can affect the degree of suppression of image changes during saccades. The head motion simulation was achieved by adding a constant camera motion to each trial. The users viewed 3D scenes from the vantage point of a virtual camera which was either stationary or rotated at a constant rate about a vertical axis (camera pan) or horizontal axis (camera tilt). During this motion, observers fixated an object that was suddenly displaced horizontally/vertically in the scene, triggering them to produce a saccade. During the saccade an additional sudden movement was applied to the virtual camera. We estimated discrimination thresholds for these transsaccadic camera shifts using a Bayesian adaptive procedure. With an ongoing camera pan, our results showed higher thresholds (less noticeability) for additional sudden horizontal camera motion. Likewise, during simulated vertical head movements (i.e. a camera tilt), vertical transsaccadic image displacements were better hidden from the users for both horizontal and vertical saccades. Understanding the effect of continuous movement on the visibility of a sudden transsaccadic change can help optimize the visual performance of gaze-contingent displays and improve user experience.

Both experiments showed interesting results which can be very helpful in designing gaze-contingent displays and employing eye tracking in virtual and augmented reality. These findings can be used to improve the user experience in VR by recognizing that users have less sensitivity to changes occurring during saccades when making certain head and eye movements. These types of changes can be used in different applications such as redirecting users to new locations in a VR, avoiding hitting obstacles and also in foveated rendering applications. With an understanding of user tolerance to these changes as a function of head motion designers can dynamically optimize and adjust the amount of update allowed in a head movement dependent manner. In particular, when a head rotation is in the parallel to a required image update the extent of the allowable transsaccadic shift is increased.

Chapter 4

Hiding Graphical Updates in 3D VR Spaces

4.1 Introduction

In this chapter I describe two experiments that were carried out in a head mounted display. These experiments looked at the extent of saccadic suppression during real and simulated head motion, and also inattentional blindness when a task is involved. In the first experiment I investigated how saccadic suppression affected by real and simulated head motion when using a 3D VR headset. In the second experiment, I analysed how detectable scene changes were when we make eye movements that are accompanied by a hand motion moving an object.

4.2 Experiment I: Saccadic Suppression during Real and Simulated Head Motion

Users of VR move their heads quite frequently for various reasons. Navigation and moving around is important in interaction with the 3D virtual environment (Bowman et al., 1997). When we make a gaze shift, our head movements can contribute towards reaching a target and can enhance our saccadic eye movement.

When the head moves faster, the saccade that is produced is smaller (Sidenmark and Gellersen, 2019; Guitton and Volle, 1987). When wearing an HMD, the user's head is mapped to the virtual camera that shows the virtual environment, and hence the position and orientation of the virtual viewpoint is based on the user's head movements. Users of a VR frequently make head rotations to explore the virtual world they are viewing, and so the VR control through head motion can enhance VR interaction. Other studies showed showed that in VR, the users move their heads more often than they do in a physical space, and more when text is presented in VR (Pfeil et al., 2018; Sidenmark and Gellersen, 2019).

Rotating the virtual camera in a virtual scene accomplishing directional changes in VR, and can be triggered by head movement. Changing the viewpoint of a virtual camera, whether through user's head motion or programmed rotations, can be useful in many applications in VR. In (Rietzler et al., 2020), head rotations alone were used as input for directional changes, without any body movement. Therefore no physical movements of the user's body were needed for keeping the user inside the path in the real world. These directional changes could occur with head rotations during body movement as well such as when walking.

Moreover, there are situations in virtual reality where a simulation of head rotation, i.e. camera rotations, is useful or even required. For example, for people with a limited range of head motion or restricted physical movements, a subtle rotation of the virtual camera can be very effective in providing a more immersive and effective VR experience. Norouzi et al. (2019), proposed a method that would rotate the virtual scene camera (denoting a user's virtual head) to a desired direction, using gaze tracking. Their technique can enable VR users to experience 360-degree rotations through discrete and continuous rotations of their virtual view (Norouzi et al., 2019). Scene camera rotations can also be used in VR-mediated surgery applications. For example in Khakhar et al. (2021), a virtual microscope rotates around the surgeon's point of interest to simulate the user's head redirection. It is used along with gaze tracking and voice commands to be of better assistant to the surgeon. Although these camera rotations may be of small size, they can still approximate the head rotations of VR users (Khakhar et al., 2021).

4.2.1 Hypothesis and Objective

The main idea in this study was to compare different extents of saccadic suppression, while wearing a head mounted display, with and without making head rotations. I leveraged how our visual perception attenuates during saccadic eye movements, and studied how an accompanying head motion affects that. I also tested whether our previous findings from desktop-based and non immersive VR, are generalizable to 3D immersive VR.

I hypothesized that there will be stronger suppression of visual stimuli during real and simulated head motion, compared to static scenes with no head motion. Also based on previous studies (Schumacher et al., 2004) conducted in 2D settings, I hypothesized that transsaccadic image displacements will be less detectable compared to intrasaccadic image displacements.

4.2.2 Stimuli and Apparatus

I designed a simple outdoor setting with a 3D bee object, and a wall behind it, as can be seen in Figure 4.1. I tried to keep the scene contents simple and stick to the goal of the experiment. All 3D scenes were created and rendered in real-time on a 64-bit Windows 10 desktop computer with Intel Core i7 6700 CPU, and 16GB RAM. The visual environments were designed in Unity3D and programmed in C# scripts, and were presented on a Dual OLED 3.5-inch diagonal 3D HTC Vive Pro Eye headset display, with a resolution of 1440H * 1600V pixels per eye (2880 x 1600 pixels combined), and a refresh rate of 90 Hz. The display had a rendered Field of View of 110° diagonal. The headset was connected to the desktop computer through a power adapter, a USB, and a video cable connection. The users' eye movements were recorded with a Tobii eye tracker, which was integrated inside the headset. The maximum sampling frequency of the eye tracker was 120 Hz.

The HTC (High Tech Computer Corporation) Vive Pro Eye, shown in Figure 4.2, was released in the summer of 2019, and includes the Tobii eye tracking system. As specified, the headset has a declared eye tracking accuracy of 0.5° to 1.1° at 120Hz, and a display with a refresh rate of 90 Hz (HTC, 2019). I used it with Unity3D, SRanipal, and SteamVR packages. The eye tracking technique used in this headset is binocular dark pupil tracking, with 10 infrared illuminators per eye. There are two AMOLED screens and each one has a resolution of 1.440 * 1.600 pixels, which gives a pixel density (PPI) of 615 pixels per inch. The total screen resolution is 2.880 * 1.600 pixels (Vive, 2019a).

(a) (b)

(c) (d)

Figure 4.1: The stimuli used in Experiment I with HTC Vive Pro Eye, in which (a) and (c) show the scene before the saccade was made, and (b) and (d) show the scene after the target object was displaced and the saccade was made.

Figure 4.2: HTC Vive Pro Eye Headset, with Controllers and Trackers

4.2.3 Experiment Design and Procedure

I designed this study using Unity3D and C# scripting. I used the packages provided by HTC Vive to connect my Unity3D projects to SteamVR and the eye tracking API provided by Tobii eye tracking (SRanipal). The packages I imported into my project were SteamVR, SteamvRInput, SteamVRResources, as well as TobiiXR and ViveSR. The last two packages were exclusively used for the eye tracking activities. Once I imported those in my Unity project, I established a connection with the eye tracker through C scripts. My main camera in the scene represented the user's head in my experiments. However, I replaced the Unity3D's default MainCamera with SteamVR's CameraRidge. The CameraRidge includes a camera (HMD display), and the two controllers. I used the Vive SRapinal SDK for eye tracking, and to access the eye tracking data.

The nature of my experiments required fast real-time processing of gaze data. For this reason, I preferred to use threads instead of collecting and processing the gaze data in the main thread. I let Unity's Update() function run the Unity-related actions such as environment, objects and camera manipulations, and parallel to that I ran the eye tracking thread which worked in the background and dealt with the gaze data. This ensured that the two tasks executed at the same time and that the main Unity thread did not need to wait for other actions before collecting and processing the gaze data.

This study was designed to compare three main head rotation conditions, and hence I had three main Unity scenes, each acting as one experimental block. In the design of this experiment, I had several independent variables. The main one that I was testing was the type of head motion, whether it was a real head rotation performed by subjects, or if it was a simulated one, with the rotation of the virtual camera in the scene with the head physically still. I also had a 'No Head Motion' block. I used the instantaneous position and rotation of the headset to perform head tracking, and calculated the head rotational velocity in conditions where users had to move their head in all trials. I also recorded the head rotation in the other two blocks to ensure the users' heads were steady and not moved. But I also observed the participants during all blocks to verify their head movements matched the purposes of that block. Other independent variables included:

- *Timing of the Scene Shift.* This was either intersaccadic (between successive saccades) and occurred at 1.05 seconds after the start of the trial when the target object had not yet moved and the subjects were not supposed to make a saccade, or it was transsaccadic (during a saccade) and occurred once the target moved and the onset of a saccade was detected. This would occur after 1.25 seconds into the trial. I let the intersaccadic shift occur at least 200 ms before trigger of a saccade as it can be assumed that no saccades were performed preceding the saccade. This way chances of coincidental saccades occurring near the image rotation were reduced.

- *Direction of Scene Shift.* During all trials on which a saccade was detected or an intersaccadic change scheduled, the whole scene received a sudden shift in one of four directions: Right, Left, Up or Down. We chose two axes for this variable, the horizontal axis and the vertical axis, as our previous experiments, in Chapter 3, showed that image shifts along these directions have a higher chance of being hidden from the user. These instantaneous rotations were added to any ongoing rotation.

- *Size of the Scene Shift.* I used two sizes of rotations to transform the displayed VR image before or during saccades. The sizes were 0.5° and 1°İn my analysis, I checked the eye movement data to ensure these conditions were true.

I used two horizontal saccade directions in my trials. The saccade that users were triggered to make was horizontal in all trials; a rightward for one group of participants, and a leftward for the other group. The size and direction of the saccade was specified by the displacement of the target object. In each trial there was either a transsaccadic scene shift or an intersaccadic one. After the target object jumped, I allowed a 100 ms time window for the participant's saccade to start and be detected. Once the saccade was detected, the transsaccadic scene rotation was applied. If the saccade was not detected within this time interval, then no scene shifts occurred and that trial was repeated until a saccade was detected and then the scene shift was applied. These trials were counted as Catch Trials. During catch trials, the subjects were not aware of this repetition of the trial, nor that there was no scene shifts, but were still asked to indicate the direction of scene change during that trial.

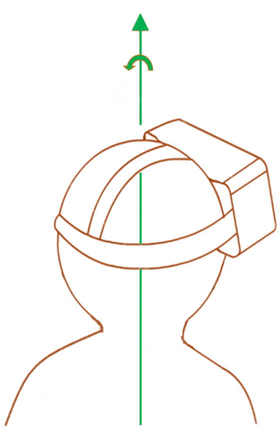
I calculated velocity of eye movement from the normalized 3d gaze direction vectors, and filtered gaze direction data with a median filter of size three. When implementing the gaze data collection, I used multi-threading in Unity. If threading is not used, the eye tracking data polling will be restricted by the main Unity thread running at the HMD screen refresh rate. The VR display often has a different refresh rate than the computer monitor. The eye tracking data is output at 120Hz (120 samples per second, and the sampling interval of 8.33 ms) on average. I used callbacks functions (provided in the Vive SDK for eye tracking), which are not limited by the refresh rate of the Unity's main thread. Using the eye callback function, which runs on a separate thread, the frame rate and sampling frequency of the eye tracker run independent from each other. The Unity FPS varies depending on the VR content (which I kept simple) and the computer specification (which was capable for our purpose). But the sampling frequency of the eye tracker remains consistent at the specified value. Using the callback function, the sampling frequency of gaze data was around 120Hz, and graphic updates were matched to this.

I conducted this study in two groups of participants (four in each group), and three blocks in each group. All three blocks were conducted in one session. To control for a learning effect, we used a random number generator between 1 and 3 (for each saccade direction) to select the order in which each participant completed the blocks. Each participant was given descriptions about the study and instructions. These included setup, general purpose, length, and the general procedure of the experiment.

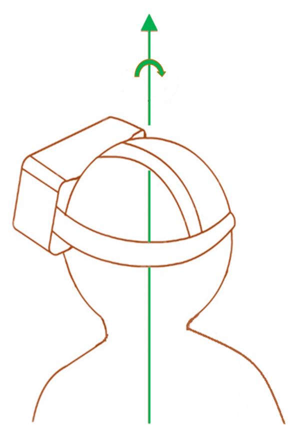
The following steps were then completed for all conditions: Participants signed a written consent form, approved by the Human Participants Review Committee at York University, before beginning the experiment. Because the headset's cushion directly touches the participant's skin, disposable hygiene face covers were offered to the participants. At the start, we ran a couple of trials of the first chosen block as a training for the participant, until we made sure they are aware of the experiment procedure and the task. These trials were excluded from our data analysis. Once they were familiarized with the procedure of the experiment, we started one of the main blocks. For the duration of the experiment, the users were seated on a chair within the Vive's tracking space to keep the environment constant for all participants.

Block with Subject's Head Rotation

During this block, we asked the subjects to rotate their head to follow the movement of the target stimuli. We recorded instantaneous rotation and position of the participant's head which was mapped as the Camera in the scene. We used this data to calculate head velocity from it. Figure 4.3, shows the two types of head rotations that followed the target object in one of the experiment blocks.



(a)



(b)

Figure 4.3: Horizontal head rotations during trials of one of the blocks. Initial head position for trials with a (a) leftward head rotation and saccade, (b) rightward head rotation and saccade.

Block with a Constant Camera Rotation

In this block of the experiment, during all trials there was a constant camera pan, in the same direction as target movement. The speed of the camera motion matched the target and was set to a constant value of 0.18 rotational degrees per frame, in the Unity's Update() function, which is equivalent to 16.2 degrees per second. We chose this value to simulate a natural rotation of the head, track the movement of the target object, and also avoid causing motion sickness for the subjects. This constant camera motion simulated a head motion in VR. The participant's head was kept still during all the trials of this block, as instructed to them. I also observed the participants during these trials, and reminded them to avoid any head rotations during these trials.

Block with No Head Rotation

The main procedure in this experimental block was similar to the previous two blocks, except that there were no head or constant camera rotations in any of the trials. The subjects were asked and monitored to keep their heads still, and only make a saccade when the target stimuli jumped horizontally.

Each block started with a calibration step. We used Tobii's SDK for HTC Vive Pro Eye headset, which included an IPD setting, and a five point calibration procedure. In HTC Vive Pro Eye headset, the calibration is a five point procedure, and starts with a point in the center. The point shrinks to fade when a subjects fixates on it, and appears in another position on the display. After every appearance, it shrinks to disappear, and this continues until it has appeared five times on different positions of the display. The eye tracking routine ends the calibration with a successful or failed calibration message. In this experiment, the calibration was performed successfully for all participants. Then the first trial started. In every trial, the users followed a moving target stimuli with their eyes. During the trial it made a sudden jump of 20° of angle triggering a saccade. To calculate the degrees of visual angle, using trigonometry, we used the formula in Equation 4.1, and determined the angular subtense between the two normalized 3D gaze direction vectors that would hit the target stimuli before and after the saccade. We calculated it in radians and then converted to degrees. This is visualised in Figure 4.4.

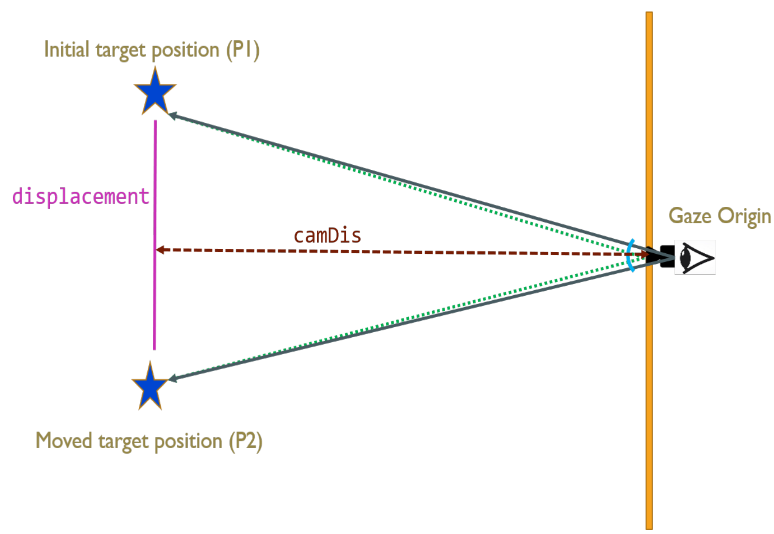


Figure 4.4: The visual angle when making a saccade upon object displacement. The virtual camera and the gaze origin are superimposed, but are shown offset for the purpose of a more clear graphics.

$$\theta = 2 * \arctan\left(\frac{displacement}{2 * camDis}\right) * \frac{180}{\pi} \quad (4.1)$$

When the target object jumped from one point to another, to trigger a saccade, we applied a one-time rotational transition to the whole scene. We rotated the virtual camera showing the scene in one of the four directions: Up, Down, Right or Left. These rotations appeared randomly in two sizes, and at two time points. They were applied either intersaccadically, before the target moved and a saccade started, or transsaccadically, once a saccade was detected and during it. To detect a saccade, we used a velocity based detection algorithm. We used a median filter to smooth the noise of the gaze direction data, and calculating the changes in degrees of visual angle over time. The trials with different variables appeared randomly. The duration of each trial was two seconds. After each trial, the subjects responded to a two-alternative forced choice task about the direction of the scene shift in the scene. Depending on the shift axis, the options were (Right or Left), or (Up or Down), of which one was the correct scene shift. Even if they did not notice a shift, whether it was a catch trial or the shift was too small to be seen, they still had to select an option. They would only proceed to the next trial when they selected a response for the task. Participants used a controller with a virtual laser beam to select a virtual response button. The procedure of the trials and tasks was the same in all three blocks. In one experiment block they followed the moving target object with their head, while in the other two blocks, they kept their head still. In one of the head still blocks, the virtual camera showing the scene moved with target (simulating a head motion) during all trials. While the other was static with no simulated or actual head movement. The bee stimuli target always moved smoothly horizontally, from left to right for one group of four participants, and from right to left for the other group, until it jumped to trigger a saccade.

The trials in each block were comprised of 16 different conditions that were randomly presented to the participants. The conditions included (2 sizes of scene rotations) * (trans- or intersaccadic scene rotations) * (4 directions of scene rotations). Every unique condition was repeated 10 times for each participant, and therefore we had 160 trials in every block for each subject presented in random order. Each block took approximately 15 minutes, and so with the three blocks plus the training and the breaks in between, the session took approximately one hour for every participant, and was completed in one visit of the participant to the lab. We cleaned the headgear and controllers using sanitizing wipes after each participant completed their session.

4.2.4 Subjects

In this experiment, eight users participated including two male and six female. The average age of the participants was 23.1 years old, [range [18 - 30]]. Three of these participants habitually wore glasses for myopia, but participated without glasses, and could see clearly. All participants were university students, and were naive about the purposes of the experiment. All, except two of them, were not regular users of virtual reality headsets. Each participant signed a written informed consent form before starting the experiment, which was undertaken according to a protocol approved by the Human Participants Review Committee of York University. Six of them received course credit for their participation, and two volunteered to take part.

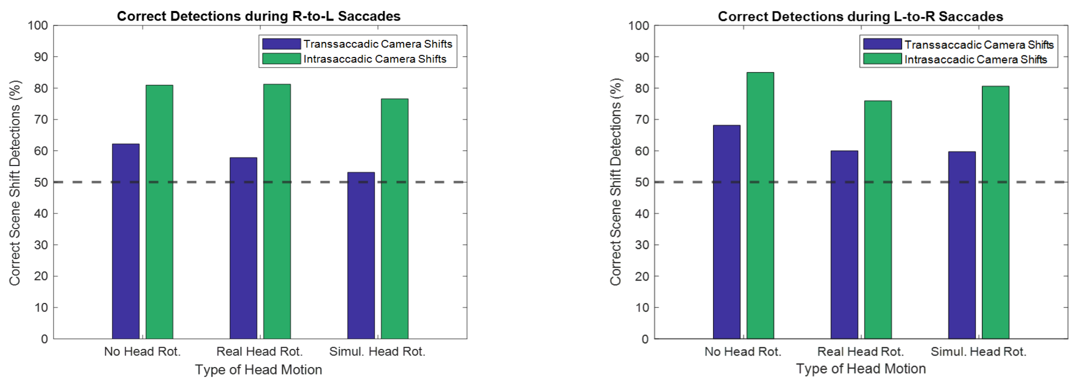
4.2.5 Data Collection

In a 3D Virtual Reality, we need to calculate the 3D gaze vector going from subjects' eye in the look direction in order to find where their gaze is in the scene. The 3D gaze vector and 3D gaze direction variables were provided to us in the eye tracking data of the device. The degree per second is a unit of angular rotational speed, and can be instantaneous or average. The average angular velocity can be obtained by measuring the angle in degrees in which an object rotates in a specified number of seconds, and then dividing the total angle by time.

The data collected for this experiment consisted of three separate data files for each participant. One file included data for the participant's eye movement features, such as gaze origin and gaze direction vectors, and the timestamps for each eye sample. From those features, we calculated other eye movement details like velocity. The second file saved the time stamped head rotation and translation coordinates in every frame and calculated the head motion values from those. The other data file we saved was the responses subjects selected during the 2-AFC task that we collected for each participant. Each line of this file included values such as the trial number, subject's selected response for scene shift direction detection, and the correct response. These files were then used for data analysis.

4.2.6 Results

To analyse the collected data, I used MATLAB's statistical toolbox and Python. I compared the differences in transsaccadic sensitivity during right-to-left and left-to-right saccades. As we expected, there were no significant differences between these conditions for real head rotation ($\chi^2(1)$=0.31, p=0.57), a simulated head rotation ($\chi^2(1)$=2.8, p=0.09), or when there was no head rotation at all ($\chi^2(1)$=2.4, p=0.12). Likewise, intersaccadic sensitivity to scene shifts was also not significantly different between right to left and left to right saccades, for real head rotations ($\chi^2(1)$=2.4, p=0.11), simulated head rotations ($\chi^2(1)$=1.56, p=0.21), and static head trials ($\chi^2(1)$=1.89, p=0.18). Figure 4.5 shows the rate of correct scene shift detections for both saccade directions.



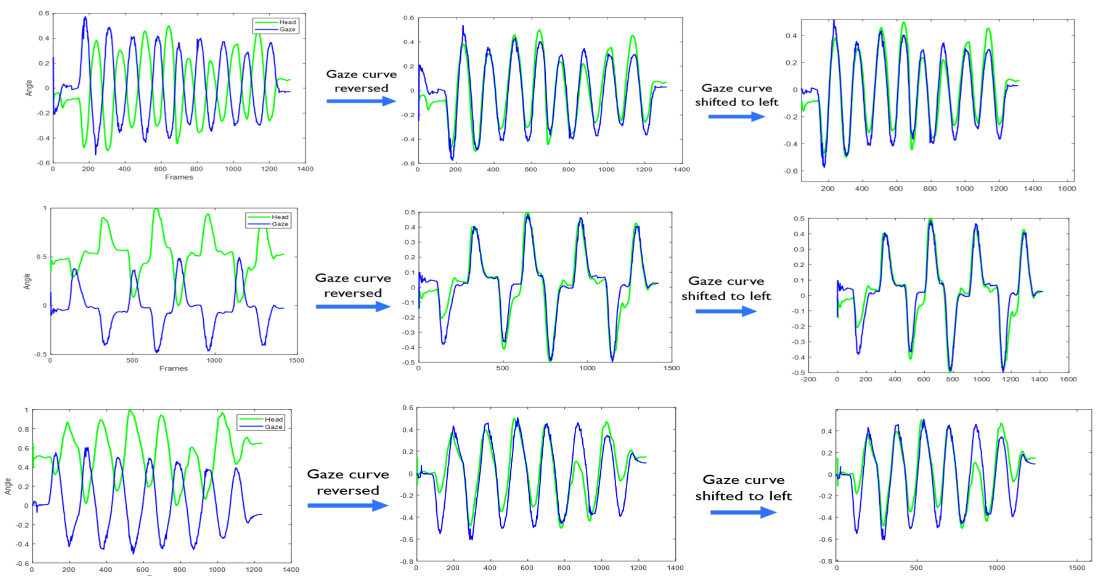
(a) (b)

Figure 4.5: Percentage of Correct Detections for (a) Right to Left Saccade and (b) Left to Right Saccades, grouped by type of camera rotation in the trial.

The accuracy for the integrated Tobii eye tracker in the headset was reported as 0.5° to 1.1° by the manufacturer. However, when we measured the spatial accuracy for each subject as they were viewing the 3D scenes, we found lower accuracy. After calibration, participants gaze was directed to the target (before it moved). On each trial, we knew where they were expected to be looking when the saccade started. This helped us determine the spatial accuracy (and whether it drifted) over an experimental session. For the end point of the saccade, we used the trials where there were no background image transformations. This helped to determine the position of where the saccade landed, without the background image being manipulated. Across all subjects, We found the average deviation to be 1.85 ° visual angle, with a range of [0.71–3.83].

In addition, by performing VOR eye movements, I measured the relative latency between the rotation of the head and the rotation of the eye on the same axis of rotation. The VOR rotates the eyes as a function of vestibular input. I measured these eye movements in a dark room, and stabilized a visual fixation point to rotate head and gaze based on it. The latencies between the two variables were in the range of [22-55] ms. Figure 4.6 shows sample measurements. As reported average VOR latency is approximately 8 ms (Collewijn and Smeets, 2000), this corresponds to a relative latency of [14-47] ms. The head movement is accompanied by the vestibular ocular reflex, but the estimate is relative to the head motion.

Here we rely on motion tracking to determine when the head is moving. Thus we are not measuring the sensed rotation of the eyes relative to the true rotation of the head. But rather we are measuring latency between the sensed rotation of the eye and the sensed rotation of the head. Hence, the assumption we had here is that the sensed rotation of the head is near zero latency. Additionally, the end-to-end latency includes the display update which is not accounted for. The VR device is optimized to minimise head tracking latency and display update. Therefore, this is a useful measurement in this context.



(a)

Figure 4.6: The green curve shows the head rotation on the yaw axis, and the blue curve shows the gaze rotation on the yaw axis. The vertical axis shows the normalized angle of change, and the horizontal axis is the frames. To match them, we first reversed the gaze signal then shifted it for five frames in the top figure, two frames in the middle figure, and four frames in the bottom figure to match it on the head curve. We measured both signals at a similar frequency of 90 Hz, and hence the time between any two consecutive frames was 11 ms.

Head Motion Type

In this study, I compared discrimination of transformations during real head motion with during a constant (virtual) camera rotation simulating a head motion. In trials with a constant rotation of the camera (a camera pan), we set the rotation of the camera to 16.2 degrees per second. The participants' average active head rotation velocities are shown in Figure 4.7. The average of participants head motion velocities during the block where they had to make a head motion was 15.88 degrees per second, which was not significantly less than the 16.2 degrees per second simulated motion ($t(15)=-0.9759$, $p=0.3457$).

In the simulated head rotation condition, we observed that participants had a tendency to rotate their head, even though they were continuously (before and during the experiment) instructed to keep their head still. The average rotation speed of the participants' head was still small and had an average speed of 3.4 degrees per second, with a range of [0.2 - 4.2] degrees per second across all eight participants. This amount of head rotation is negligible in our experiment. On the other hand, in the conditions with no head active or simulated head rotation, the participants found it easier to keep their heads still. In these conditions the average head rotation velocity was very close to zero (0.89 degrees per second, and a range of [0.2-1.6] degrees per second), across all eight participants.

By analysing the eye movement data, we noticed that in almost all trials, participants made their largest saccade in response to the object displacement in the scene, as they were directed. The trials were short in duration, and participants focused

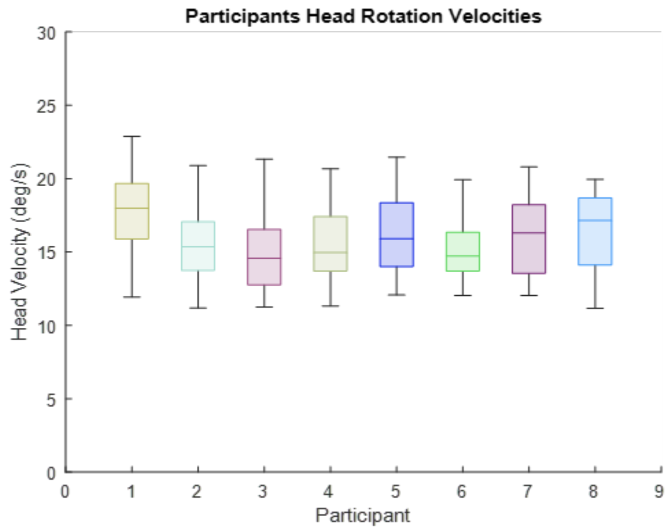


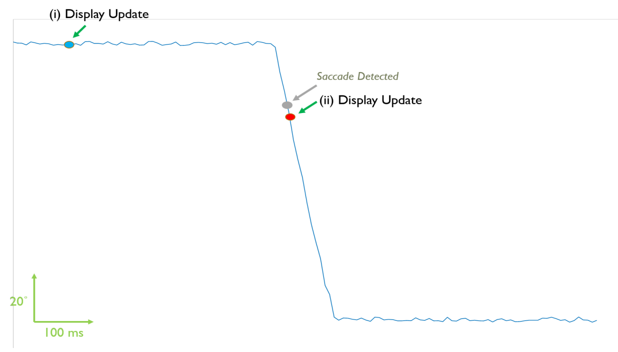
Figure 4.7: Horizontal head rotation velocities during trials with users' head motion. Box plot whiskers show the minimum to maximum of head rotations for each participant, across all trials. The central mark on each box indicates the median, and the top and bottom edges of it are the 25th and 75th percentiles, respectively.

on the task as to when they should generate their saccade. Nonetheless, they still made smaller accidental saccades or eye blinks. In our analysis we included trials with saccades that were in the range [18.5-20.5] visual degrees in size, and excluded other sizes. As explained earlier, trials in which a saccade was detected 100 ms after the target onset, or before the target jumped, were automatically not included. The average duration of included saccades was 78±5.7 ms. Figure 4.8, shows changes of eye movement in degrees over time with the time points when scene shifts were applied, as well as the saccades detected/undetected within a specified time frame after target displacement.

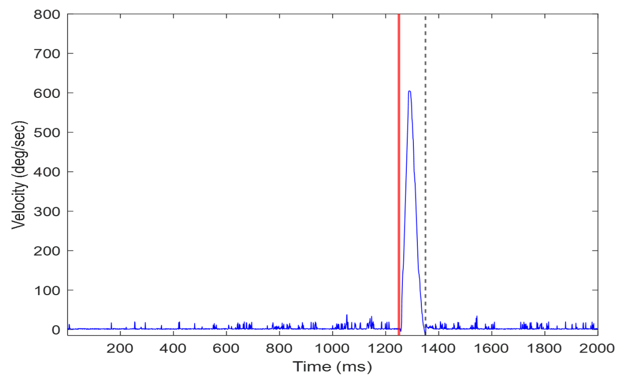
In the trials with a transsaccadic shift, the average latency of the saccades that were made in response to the scene shift were 168 ms [139-230], and their average amplitude was 19.1 degrees. The duration of every trial was short (2 seconds), and participants had a clear task as to what to expect and when to make a saccade, and therefore the saccade were fast with low latencies. In addition, during the intersaccadic scene shift trials, we looked at the motion of the eye when the scene shift occurred. The analysis of the gaze data of our participants across all intersaccadic trials did not show any significant difference of the eye movement behaviour between conditions with various scene shift sizes or directions. We further noticed that as the trials were proceeding, the saccades were landing with better precision on the target stimuli, and the spatial error decreased.

The intersaccadic shifts, that occurred before a saccade, were significantly more likely to be discriminated correctly compared to the transsaccadic ones, during real head motion ($\chi^2(1)=64.44$, $p<0.001$), simulated head motion ($\chi^2(1)=60.95$, $p<0.001$), and the static head trials ($\chi^2(1)=60.50$, $p<0.001$). We noticed that the rates of correct discriminations were lower when there was a simulated head motion in the scene, compared to when the subjects moved their head themselves, although the difference was not significant. This was also reported verbally by participants. These findings can be seen in Figure 4.9. Participants often reported they noticed some changes in the image even though most subjects could not distinguish the exact directions on that axis (left vs right). Thus chance performance corresponding to guessing predicts 50% correct, on average, for this analysis.

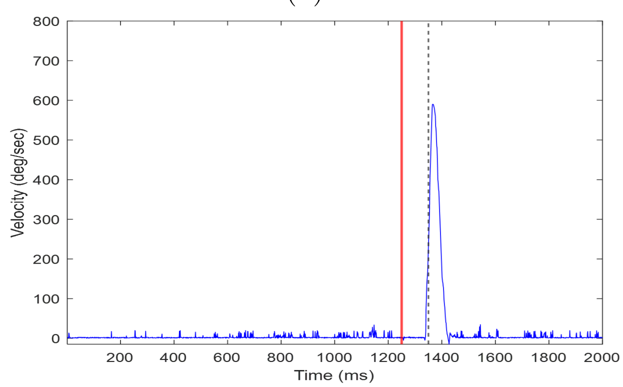
The transsaccadic scene shifts were significantly less visible to subjects during real head rotation trials, than trials with no head rotations ($\chi^2(1)=4.16$, $p<0.05$). Likewise, they were significantly less sensitive during blocks with a simulated head rotation, than the trials with a static head ($\chi^2(1)=4.65$, $p<0.05$). However, the difference in correct scene shift discriminations was not significant between real and



(a)



(b)



(c)

Figure 4.8: Sub-figure (a): Sample saccade. The changes in gaze direction that during a 20° saccades. Point (i) shows timing of a typical intersaccadic scene shift, and point (ii) shows a transsaccadic scene shift. The sub-figure (b) and (c) show the timing of two saccade events versus their velocity. The red vertical line shows the time point that the target object was displaced, and the user was expected to start their saccade. We allowed a 100 ms time window after the object displacement for the subject's saccade to be detected. The vertical dashed black line shows the end of this interval. In (b), a detected saccade is shown and in (c) the saccade was not detected and therefor no seen shifts were made. Such trials were repeated until a saccade was detected within the specified time frame.

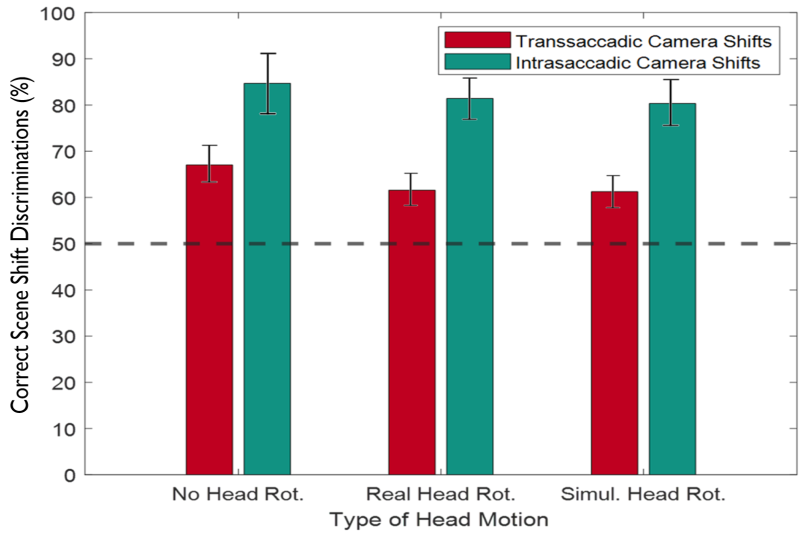


Figure 4.9: Percentage of correct scene shift discriminations during blocks with different head rotations

simulated head rotations ($\chi^2(1)$=0.13, p=0.90). Figure 4.9 shows this.

Transsaccadic Scene Shifts

During the trials where the scene shift occurred during a participant's saccade, there was a pattern in the discrimination rates of scene shifts based on their direction and their size. The horizontal transsaccadic scene shifts (right or left) were significantly harder to discriminate correctly, than the vertical ones (up or down), during real head rotation ($\chi^2(1)$=5.18, $p<0.05$), simulated head rotation ($\chi^2(1)$=4.45, $p<0.05$), and no head rotation trials ($\chi^2(1)$=4.38, $p<0.05$).

In addition, when comparing discrimination rates for the two sizes of scene shifts that occurred during saccades, the larger size scene shifts (1°) were significantly easier to discriminate during trials with subject's active head rotation ($\chi^2(1)$=12.78, $p<0.001$). Similarly, larger scene shifts were easier to discriminate than smaller ones, during trials with a simulated head rotation ($\chi^2(1)$=16.45, $p<0.0001$), and

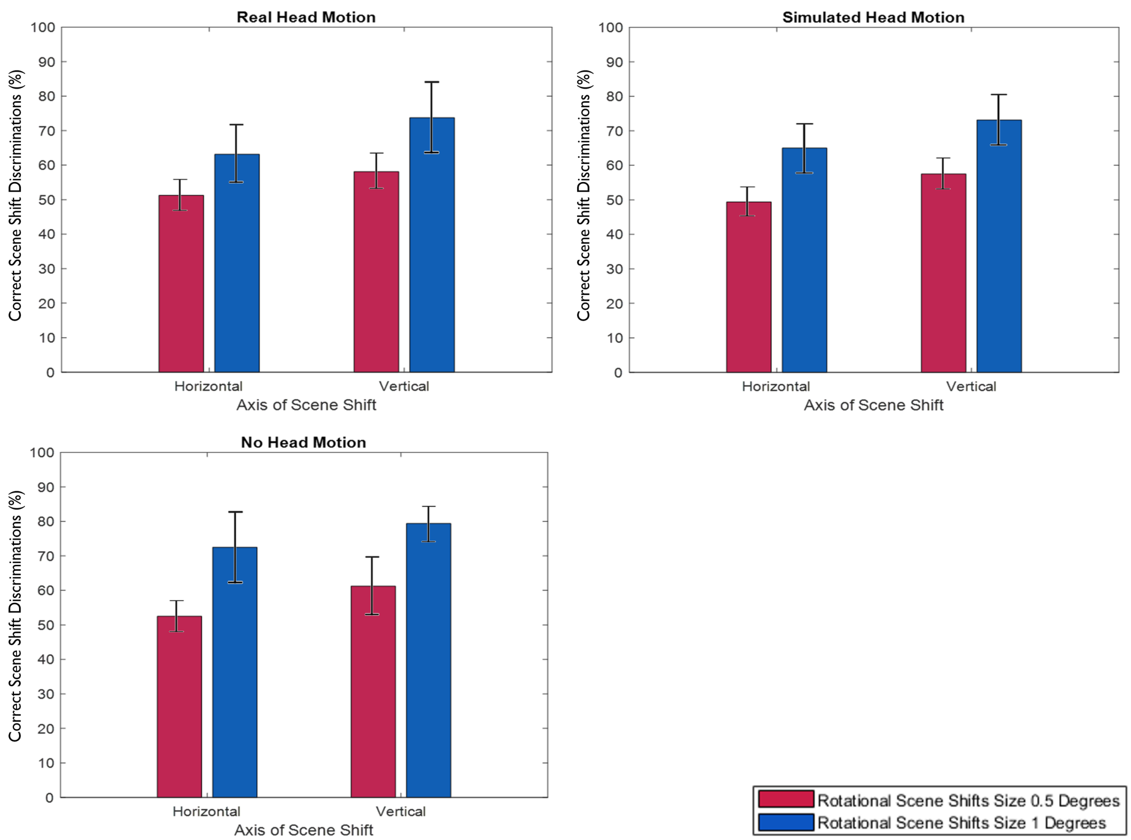


Figure 4.10: Correct responses for the scene shifts occurring transsaccadically, for each head motion condition, and based on the size of the scene shift and its axis of rotation.

trials with a fixed head ($\chi^2(1)$=26.06, p<0.0001). Figure 4.10 shows the percentage of correct discriminations, during each head motion type, grouped by axis and size of scene shift. Regardless of type of head rotation, there was significant difference ($\chi^2(1)$=53.72, p<0.001) in discrimination performance between 0.5° and 1° transsaccadic scene shifts.

Intersaccadic Scene Shifts

In the trials when the scene transition occurred during a fixation, participants could detect the directions significantly more easily than when it occurred during a saccade. For these shifts, we also used the same size and direction that we used for transsaccadic shifts. In our analysis, we removed the trials in which there were accidental saccades or blinks during these intersaccadic scene shifts. Our results showed significant difference between intersaccadic horizontal and vertical scene shifts during real head rotations ($\chi^2(1)$=6.45, p<0.05), during simulated head rotations ($\chi^2(1)$=3.95, p<0.05), and on trials with no head rotations ($\chi^2(1)$=12.34, p<0.001).

Also, the 1° size scene shifts were significantly easier to discriminate by the subjects than the 0.5° size scene shifts during real head rotations ($\chi^2(1)$=35.88, p<0.001), during simulated head rotations ($\chi^2(1)$=37.98, p<0.001), and trials with no head rotations ($\chi^2(1)$=25.49, p<0.001).

Catch trials

The trials in which no saccade was detected upon the displacement of the target stimuli, were counted as catch trials. There was no rotational scene change in those

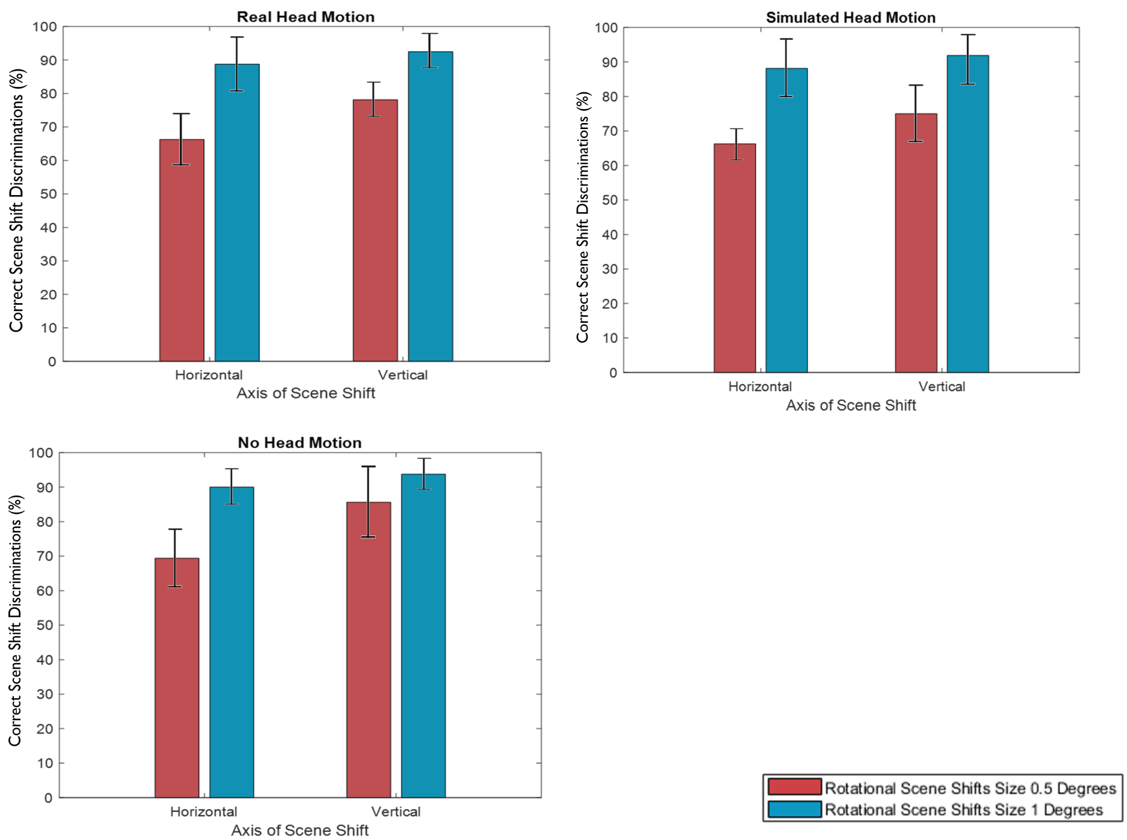


Figure 4.11: Correct responses for the scene shifts occurring intersaccadically, for each head motion condition, and based on the size of the scene shift and its axis of rotation.

trials. Subjects were unaware that there was no background shift in the trials and still responded to the task. We looked at subjects' responses during these trials, as shown in Figure 4.12. We looked at the patterns of response guesses for three types of head motion blocks, between right and left saccade directions. We noticed that when the head motion was a pan to the right, subjects tended to select more 'Right' responses than 'Left', and also selected more 'Left' responses when the head motion was to the left. The difference between Right and Left selection was significant during trials with a simulated head rotation ($\chi^2(1)$=4.06, p<0.05), and the direction that was in the same direction to the head rotation direction had a higher selection. Additionally, selecting an 'Up' and a 'Down' response depended on the direction head rotation in the trial ($\chi^2(1)$=6.70, p<0.01). In that case, subjects selected 'Up' direction when their head rotated rightwards, and selected more 'Down' direction when their head rotated leftwards. The difference between selection of responses during other head motion types was not significant.

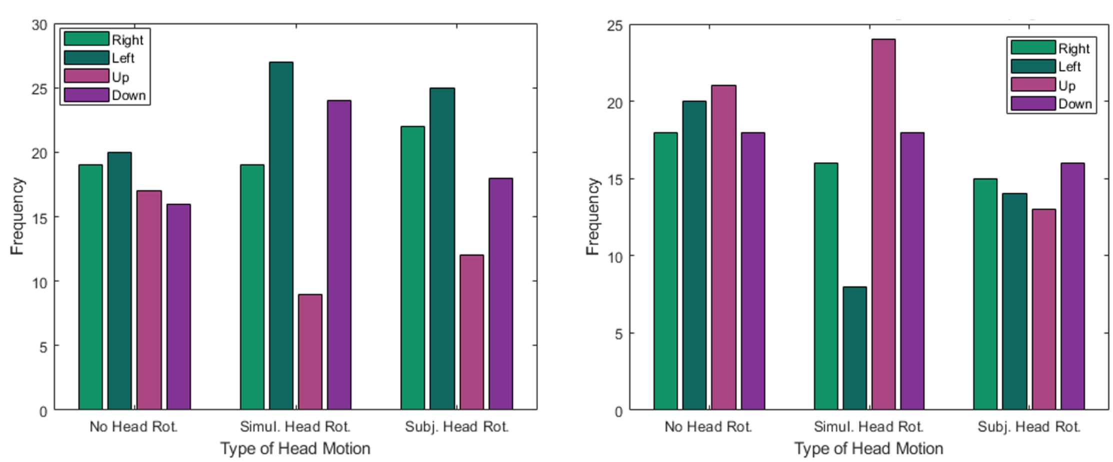


Figure 4.12: Subjects' selected responses during catch trials, for scenes with a right-to-left target movement (left image) and a left-to-right target movement (right image).

4.2.7 Discussion and Conclusions

In this experiment, I looked at how sensitivity to scene shifts during saccades and before saccades changes across different head motion types in a virtual environment. One finding was that intersaccadic scene shifts were easier to correctly discriminate than the transsaccadic ones, regardless of whether the head moved or was fixed. This is consistent with the previous studies (Allison et al., 2010), and implies the existence of a stronger suppression during the saccadic eye movements.

Previous studies showed that viewers were more sensitive to transsaccadic image transitions that were in the same direction as the saccade, than the ones in opposite direction (McConkie and Rayner, 1975; McConkie and Currie, 1996; Currie et al., 2000). However there were other studies that showed the opposite effect or no particular difference at all (Bridgeman and Macknik, 1995; Allison et al., 2010). Further, horizontal scene shifts seemed to be masked better than vertical shifts, during the horizontal saccades (and also during intersaccadic trials). This suggests when making a horizontal saccade in a VR, horizontal rotations of the virtual camera can be better hidden from the users. This is in agreement with the previous studies (Allison et al., 2010; Keyvanara and Allison, 2019). In addition, as a follow up to our previous research, we compared the difference between conditions where a user made a head rotation and when they keep their head still and we added a camera rotation in the VR scene to resemble a head rotation. When the saccade was generated during the movement of the head, we found a stronger suppression than when the head was still. We noticed how there was a higher chance of scene shifts not being correctly discriminated when there was real or simulated head rotation. This was in line with our previous study discussed in chapter 3 (Keyvanara and Allison, 2020).

Overall, our results suggest that when a saccade is accompanied by a head rotation, there is a higher chance that other display updates occurring during saccade may go unnoticed. In research on monkeys, saccade dynamic features during a restrained head were indistinguishable from saccades made during free head movement (Morasso et al., 1973), using eye movements alone or using a combined eye-head movement, not just for animals but also humans (Morasso et al., 1973). In a VR study (Jerald et al., 2008), the results showed that system latency in a head mounted display is less likely to be noticed when the scene moves with the head. This means when a head yaw begins and the scene moves with head, additional scene motion can be added to VR. It will be less noticeable compared to when slowing down a head yaw or changing head rotation direction.

Head rotations are important in different applications in VR. Users of a VR can be triggered to rotate their head to avoid them hitting obstacles in a small physical space. A rotation gain can reset the orientation of a user by involving them in a task that requires head rotations (Coelho et al., 2022). In this study, we also found that if small size rotation shifts were applied during head motion and parallel to saccade direction, even if they are intersaccadic (small horizontal intersaccadic scene shifts), there is a chance they can go unnoticed by users. It should be mentioned that, in this case, we verified that the these intersaccadic shifts did not occur during blinks or saccades. We also noticed that for larger intersaccadic scene shifts, we cannot make such conclusion. Further research can be conducted to investigate whether head motion direction, and hence head tracking alone, would be sufficient for applying small shift in the same direction as head motion. This may mean very small and subtle image transitions can be applied to the VR image during a head rotation, without requiring eye tracking, and go unnoticed by the viewers. Previous studies have shown how the interaction of head and gaze is integrated in a VR viewing setting. In (Sitzmann et al., 2018), it was suggested that head orientation alone (with no eye tracking) could be sufficient for predicting accurate saliency. In addition, a head-rotation based saccade prediction method was used for redirected walking. A study (Joshi and Poullis, 2022) utilized a trained neural network to perform a real-time saccade prediction, during apparent head rotations in VR. The deep neural network was trained on head rotation data. Their approach doesn't require eye tracking hardware, and showed good results in redirected walking user studies. Overall, these studies reveal the importance of head rotations in interactive applications VR .

In addition, during catch trials with no scene shifts, there was a preference to select a response that was on the same axis and direction as the constant camera pan (simulating a head rotation). When stationary observers view a large moving visual stimulus such as when the main camera in the scene moves, they may experience an illusion of self motion. Therefore, our participants, believed that they saw a scene shift, in the same direction as the virtual camera showing the scene.

When implementing gaze-contingency on an HMD, the VR display should be updated with minimal latency to ensure the updates happen as soon as possible after the saccade onset. There are different sources of delay for a saccade-contingent display update. these include the tracker delay, the saccade detection speed (which depends on the speed and robustness of algorithm and the sampling frequency), and also the display refresh rate Communications for example with a wifi connection could further add latency. To reduce the latency in our experiment, we used a multi-threading implementation for our gaze tracking part (to get the maximum number of eye samples from our tracker in Unity), having a fast processor, no wifi connection, and elicited a big saccade size. Although Stein et al (2021, report HTC Vive Pro Eye's end-to-end latency to be high for a gaze-contingent application (79 ms for the Vive's native SDK, and 80 ms when using Tobii XR SDK), their eye sampling frequency was at 88.3Hz, and their evaluation was conducted at the very beginning of when the headset was released (Stein et al., 2021). Since then, the eye tracking SDK provided by HTC Vive has been updated and with the right programming it allows a 120Hz sampling rate. In Sipatchin (2021), a value of 58.1ms is reported for the latency of HTC Vive Pro Eye, which is an acceptable value in the design of some gaze-contingent applications. Both studies showed that there was no difference between the eye tracking SDKs of Tobii Pro and the SRanipal (by HTC Vive) (Sipatchin et al., 2020). Moreover, in (Albert et al., 2017) the tolerable latency for a foveated rendering application in virtual reality is reported to be a total system latency of 50–70ms, which is the same as the values reported in the two studies above. Nevertheless, there are other ways that can help in improving the end-to-end latency of a VR HMD, such as predicting where a saccade will land (Arabadzhiyska et al., 2017; Morales et al., 2021; Griffith et al., 2018; Morales et al., 2018). In our experiment, we used a saccade amplitude of 20° with a long duration which allowed a longer time for the transform to be completed. We implemented our eye sampling based on a 120Hz frequency, and kept our scene layouts relatively simple, to reduce the amount of processing, and therefore latency.

4.3 Experiment II : Inattentional blindness during multimodal interaction

Human vision is one of the main sources of information from our environment (Burns et al., 2005). An added sense, such as haptics or hearing, can help increase our understanding of our surroundings and gain more sensory input. Multimodal cues help us have a unified experience of our world (Shams and Kim, 2010). In many virtual environments, the users are actively engaged in some activity or have a task or goal to accomplish. Every task completion requires a certain amount of attentional engagement, and different senses to interact with the virtual environment (Witmer and Singer, 1998). In many VR applications, a tracked controller can be used to allow a physical movement of the hand and to act as a user's physical hand in the virtual world and enable virtual object selection or manipulation (Lee et al., 2019; Ali and Cardona-Rivera, 2020).

There are different techniques that allow hand interaction in VR, such as ray casting for selection of distant objects, virtual gloves, leap motion, use of a controller and attachable gear on hand (Khademi et al., 2014; Bowman and Hodges, 1997; Perret and Vander Poorten, 2018; Kim et al., 2020). In (Poupyrev et al., 1998), the authors compared two interaction means: a virtual hand with a virtual pointer, that were used in object selection and positioning. Their experiments showed that both interaction methods were suitable for different interaction scenarios. Another study (Chen et al., 2021), also found that a virtual hand, whether in close proximity to the user or at a distance, had a lower error rate of moving target selection, but a slower selection time, compared to a laser pointer selector.

Inattentional blindness is a perceptual experience where we fail to see highly visible objects that are in our field of view, because our attention is elsewhere (Mack and Rock, 1998; Mack, 2003). Change blindness is another failure of visual awareness in which an observer fails to notice an obvious change, and has been studied in many VR studies. One study compared how change blindness differed between vision and touch. They showed that unimodal attention during change blindness, visual only or haptic only, showed better performance compared to crossmodal conditions where both the vision and tactiles were involved. Their results showed that change blindness is a multisensory process, and that vision and haptics differ in their encoding and memory limits (Auvray et al., 2007).

Inattentional blindness and its applications in VR have gained more attention in the recent years (Vasquez-Caballero, 2020; Joshi and Poullis, 2020b). Currently there is no previous research in VR to compare how different perception of the virtual environment seems when a user is focused on a task, where two versus one sense are involved. Our study is innovative in exploring and comparing various combinations of scene shifts and a second task for participants to focus their attention on.

4.3.1 Hypothesis and Objective

The main hypothesis in this experiment is that performing a task involving a hand movement will make background image transitions less detectable compared to when there is only gaze involved. Previous literature shows that when more than one sense is involved and our attention is focused on another task, we cannot detect the other changes that occur in our field of view (Blouin et al., 1995).

The objective of this study was to investigate whether using a hand movement to displace a virtual object affects the detection and discrimination of background changes in the scene. We were interested in comparing this with using only gaze to watch an object being displaced. Inattentional blindness provides an opportunity to hide unobtrusive updates in VR (Joshi and Poullis, 2020a; Joshi and Poullis, 2020b), which was a great motivation to investigate vision and touch interaction in two VR tasks. Specifically, we wanted to see what kind of image shifts could be better hidden from a user of VR without them noticing a reorientation of their viewpoint.

4.3.2 Stimuli and Apparatus

Our stimuli were 3D scenes that we designed in Unity3D. For this experiment, we created a 3D room, with some furniture and closed on all sides by walls and a ceiling. We placed a box on a table, with two cubical stands. The sample scene can be seen in Figure 4.13. All 3D scenes were created and rendered in real-time on a 64 bit Windows 10 desktop computer with Intel Core i7 6700 CPU, and 16GB RAM. The visual environments were designed in Unity3D and programmed in C#

scripts, and were presented on a Dual OLED 3.5-inch diagonal 3D HTC Vive Pro Eye headset display, with a resolution of 1440H * 1600V pixels per eye (2880 x 1600 pixels combined). The headset was connected to the desktop computer through a USB and video connection. The users' eye movements were recorded with a Tobii eye tracker, which was integrated inside the headset. The sampling frequency of the eye tracker was 120 Hz.

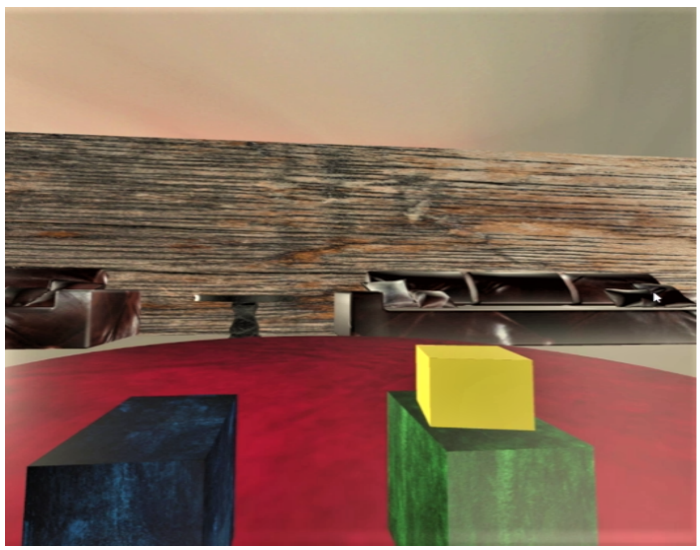


Figure 4.13: Experiment Stimuli: The 3D scene used in all blocks of this experiment.

4.3.3 Experiment Design and Procedure

I used Unity3D and C# scripting for the design of this experiment. I followed a similar use of packages as the previous section. Mainly the packages included SteamVR and the eye tracking API provided by Tobii eye tracking (SRanipal), and some other imported packages such as SteamVR, SteamvRInput, SteamVRResources, as well as TobiiXR and ViveSR.

the subject was to evaluate how focusing on a task, such as moving an object can change how we perceive changes applied to a scene compared to passive viewing. I created two variants of a scene in Unity. The layout of both versions was the same. They included a 3D room with a table in the middle, and two pillars on it. There was a box on the right pillar. During the trial this box was moved from one pillar to the other. The difference between the two versions was how the box on the pillar was moved: it was either displaced by the subjects' physical hand by grabbing it using the trigger button on a controller, and then it moving it with the controller, or it would be moved automatically through scripts (The two scenes were similar, but the interaction differed.)

Figure 4.14: Vive Controller used for moving the target object.

We conducted this study in one group of participants, and two parts for each individual. Both parts were conducted in one visit, and took almost one hour altogether. To control for a learning effect, we broke up the two parts into four blocks (two blocks in each part). Upon arrival, each participant was given some background and instructions about the study. Participants signed a written consent form approved by the Human Participants Review Committee of York University and read a covid-related hygiene information sheet, before beginning the experiment. Because the headset's cushion directly touches the participant's skin, disposable hygiene face covers were offered to the participants. At the start, we ran a couple of trials of the experiment to train the participant, until we made sure they understood the experiment procedure and the task. These trials were excluded from our data analysis. Once they were familiarized with the procedure of the experiment, we started the first block. For the duration of the experiment, the users were seated on a chair within the Vive's tracking space to keep the environment constant for all participants. Participant setup can be seen Figure 4.15

Figure 4.15: A participant, seated on a chair with no wheels, completing the experiment.

Hand Movement Blocks

The subjects used the controller to move the box from the green pillar on the right to the blue pillar on the left. They touched the box with the controller, and pressed the trigger button, shown in Figure 4.18, to pickup the object. They held onto the trigger button (as it would maintain their hold on the box connected to the controller), and released the trigger button when they wanted to set the box on the blue pillar on the left.

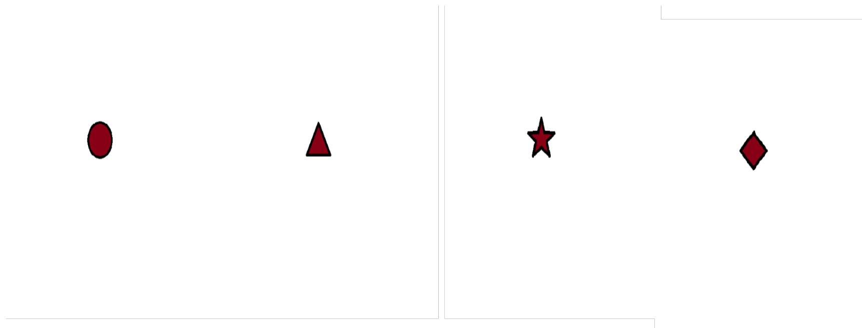


Figure 4.16: Shapes that appeared on the cube while it was displaced

As the target object (the box) was being moved, its 3D position and orientation obtained from the Vive tracker were recorded in a file. They were used later (in the gaze-only block) for object displacement. While the subjects were moving the box, two things happened. There was an instantaneous (across one frame) scene rotation in one of the directions of up, down, right or left, in one of sizes of 0.75° 1.5° or 2.25°. The participants were asked to indicate this direction after each trial. To make sure the participants kept their eyes on the target box while moving it, we placed a pattern on the box at some random time before or after the scene shift. The pattern was one of the four shapes shown in Figure 4.16. The pattern would appear on the box, remain for 100 ms, and then disappear. Participants would be asked to report this shape at the end of each trial. This was also a 2AFC, with two buttons, one that showed the correct shape as well another one from the remaining three shapes. When the participants selected a choice for both tasks, they could proceed to the next trial.

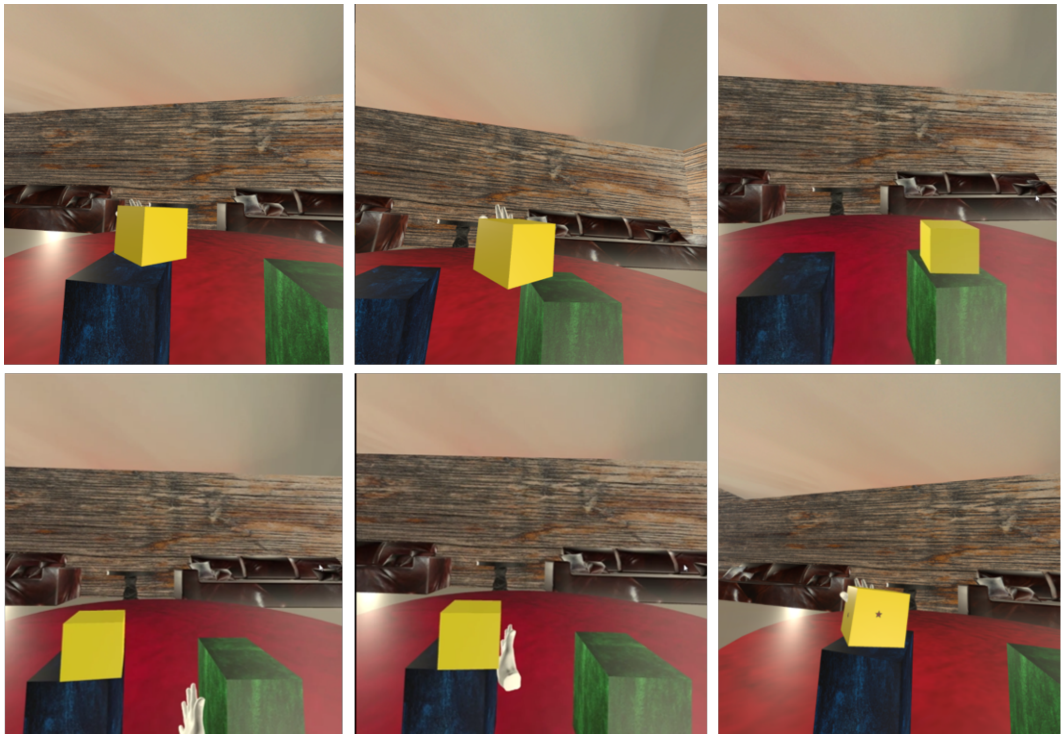


Figure 4.17: Frame sequences of a sample trial in the experiment block

Gaze Only Blocks

In this experimental block, during each trial the participants only watched the recorded displacement of the box from the right pillar to the left pillar. During these trials, there was no hand interaction of the subjects with the scene objects during the trials. The discrimination tasks were identical to those in the hand blocks. They responded to two tasks, one about the direction of the scene shift and the other about the shape that appeared on the box while it was being moved.

The rest of the procedure was the same for both types of block: Each trial took 2300 ms, but subjects could take as long as needed to make their responses. The times taken to respond to the tasks were recorded for each trial. As explained, we ran the study in four blocks for each subject. Because of the design of the experiment, a block with the hand movement (box displacement by user) had to be completed first, followed by a gaze only block. Then we ran the gaze only block again (with the box displacement data of the first block), and then finished with another hand movement block.

Figure 4.18: Snapshot of the experiment tasks, where participants had to select the correct scene shift they saw and the pattern that appeared on the box. Subjects used the controller, and pressed its trigger button to select a button.

I ran the experiment blocks in a reverse counterbalancing order, as in Figure 4.19. The reason was that I recorded the hand displacement of the target stimuli, and had to use that for the displacement of the target in the gaze-only block. Each block started with a calibration procedure and IPD setting, for which Tobii's SDK for HTC Vive Pro Eye headset was used. In this experiment, the calibration was performed successfully for all participants, except one. That participant removed their glasses, confirmed they could see clearly and the calibration completed successfully. After calibration the first trial started. The trials in each block, consisted of 12 different conditions that were presented to the participants in random order. The conditions included (3 sizes of scene rotations) * (4 directions of scene rotations). Every unique condition was randomly repeated 10 times for each participant, and therefore we had 120 trials in both the hand and gaze conditions for each subject, that we broke into four blocks of 60 trials each. Every block took approximately 10 minutes, and so the four blocks plus the training and the breaks in between the session took approximately one hour for every participant, and were completed in one visit of the participant to the lab. We cleaned the headgear using sanitizing wipes after each participant.

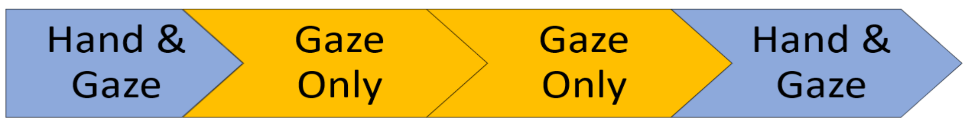


Figure 4.19: The order of running the experiment blocks.

4.3.4 Data Collection

The data collection included the data of the responses of the participant for the two tasks. We collected separate data files for each participant. The data files included data for the participant's eye movement features, and head rotation and translation coordinates in every frame. The other data we saved was the responses subjects selected during the 2-AFC tasks for the scene shift and the pattern on the box. Each line of this file included values such as the trial number, subject's selected response for scene shift direction discrimination, and the correct response. These files were then used for data analysis.

4.3.5 Subjects

For this experiment, I recruited eight participants, four female and four male, with an average age of 28.63 years (range [24-33]). Five of the participants had normal vision, and the rest had corrected to normal vision, but none wore or needed glasses for the experiment. Five subjects had previous experience using VR, and were quite familiar with using a handheld controller and VR headsets. The participants were all graduate students at York University, and received a $15 financial compensation for their participation. Before beginning the experiment, they signed an informed consent form, and read an information sheet about the covid19 related health and safety protocols followed in the lab.

4.3.6 Results

To analyse the data I collected in this experiment, I used MATLAB and python scripts to read the data files and perform statistical analysis. The timing of the two events that occurred across all trials and participants was recorded during all trials. On average, the scene shifts took place at 990±56 ms into the trial, and the shape appeared on the box at 1058±48 ms, during the gaze-only block. For the trials with a hand movement, the average timing of the scene rotations was 996±46 ms and the appearance of the shape on the box was 1068±33 ms from the start of the trial.

Trials with success in the shape identification task

Participants were instructed to keep their eyes on the box throughout all trials. Overall, the rate of correct shape on the box identifications was 92.60% for the blocks with gaze only, and 91.67% for the blocks with the gaze and hand movement.

Figure 4.20 shows the percentage of correct shape on the box identifications, and percentage of scene shifts identified correctly with and without correct identifications of shape on the box.

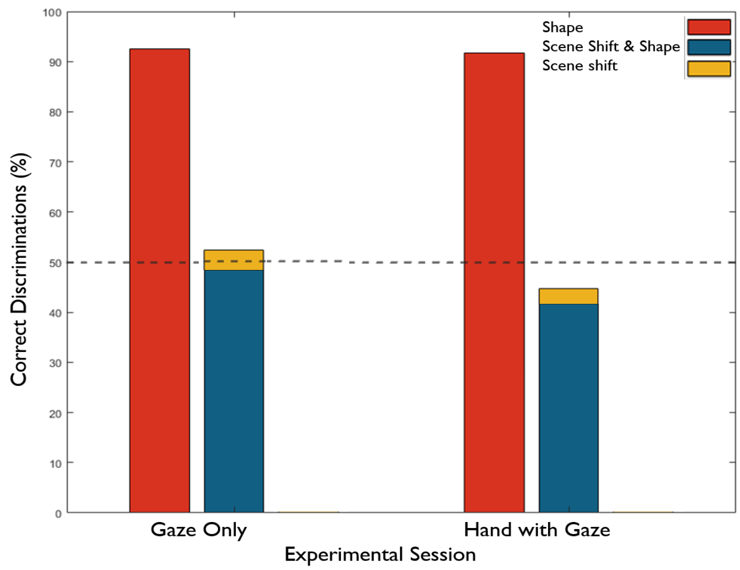


Figure 4.20: Percentage of correct identifications for shape on the box (red), scene shifts and shape on the box (blue), scene shifts (yellow), during both experimental blocks.

Since this was a dual task study, I further looked at trials where observers correctly responded to both tasks at the same time. This indicated that the subjects were focused on the scene shift task and also kept their eyes and attention on the box. First, I looked at how the correct discriminations of scene shifts differed between gaze and hand movement blocks. For this repeated measures comparison I used a McNemar statistical test to compare the response of subjects across the two interaction conditions of experiment. For scene rotations to the right, the difference in correct discriminations was significant ($p<0.01$, McNemar) between the two hand and gaze blocks, and users tended to correctly select more 'Right' responses during the block with a hand movement. However there were no significance differences for left rotations, up rotations, or down rotations.

I also compared the correct responses for different sizes of scene rotations between the hand and gaze-only blocks. The McNemar test showed no significant difference for either the size 0.75° or 1.5° scene rotations between the two experiment blocks. However, for the larger size of 2.25° scene rotations, there were significantly more correct discrimination during the gaze-only ($p<0.01$), than hand and gaze condition.

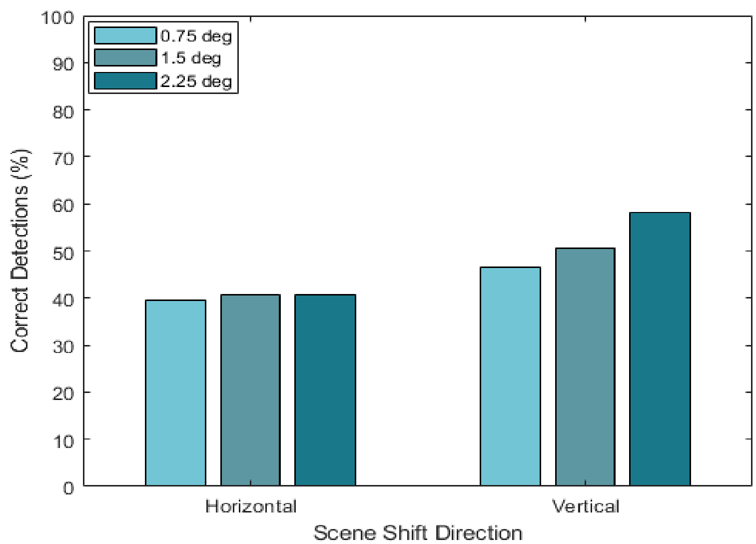


Figure 4.21: Percentage of correct scene shift discriminations during trials in which participant's hand and gaze were used for interaction.

Trials with hand and gaze interaction During these trials, the participants used a controller and their physical hand to move a virtual object, while also keeping their eyes on the object. The participants' selection of correct responses during horizontal scene shifts was not significantly different for the three scene shift sizes. Figure 4.21 demonstrates these detections for the trials that subjects used their physical hand for object displacement.

Trials with gaze only interaction During these trials, the participants only watched the box displace. The motion of the box was set by the recorded position and rotations of the box during the block that subjects moved the virtual box themselves. The results showed that the larger scene shift (i.e. 2.25°) was significantly easier to discriminate than the two smaller sizes of scene shifts of 1.5° ($\chi^2(2)=4.46$, $p<0.05$), and 0.75° ($\chi^2(2)=4.20$, $p<0.05$). Figure 4.22 shows these results for the trials when subjects only used their gaze to follow the target object.

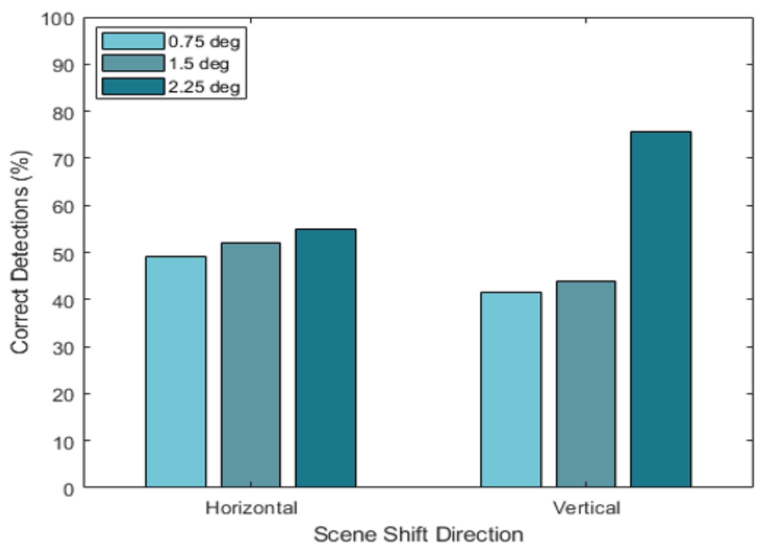


Figure 4.22: Percentage of correct scene shift discriminations during trials in which only participant's gaze was used to follow the target stimuli.

Trials with failure in the shape identification task

In the approximately 8% trials where the shape on the box was not discriminated correctly, there was 63% and 47% of correct responses for scene shifts, during the gaze-only and hand-gaze blocks respectively. Figure 4.23, shows the correct discriminations for each size of scene rotation, for the two blocks. We found the rate of correct discriminations in this case to be significantly different between the gaze-only block and the hand-gaze block ($\chi^2(1)$=6.88, p<0.05).

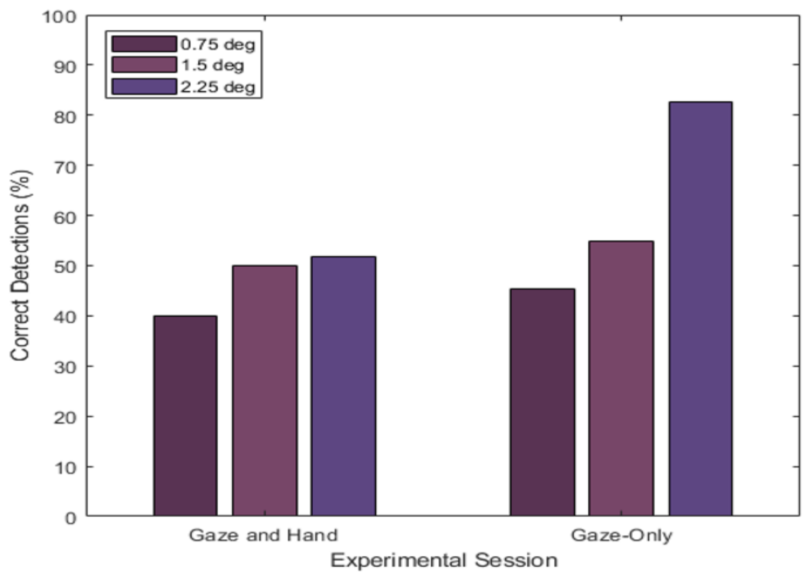


Figure 4.23: Percentage of correct scene shift discriminations during trials in which the shape on the box was not identified correctly (Failure in shape on the box identification).

4.3.7 Conclusions and Discussion

This research explored a perceptual phenomenon, known as inattentional blindness, to modify participants' sensitivity to image shifts. This was conducted in a similar way, as saccadic suppression, with an aim to hide graphical updates from the user. Overall, our results showed that when the users were involved in a primary task (focusing on the displacement of the box), whether they used one sense or two senses, they had reduced sensitivity to rapid changes of the VR viewpoint in different directions. Although the pattern we watched was that when they were only watching the target stimuli (and not using their hand), they were more likely to notice larger shifts that were orthogonal axis to the displacement of the box (as Figure 4.22 shows), which is consistent with our studies in chapter 3. This indicates that subjects were able to use their covert attention for larger changes of the 3D scene to identify the direction of reorientation, which is similar to previous studies (Richards et al., 2012). However, we also noticed that subjects still selected correct responses for scene shift discriminations. When subjects were using both their hand and gaze to move and track the object, they were less sensitive to image rotations, than when they only used their gaze.

There are many cases where we may need to make a hand movement in VR and move objects in a virtual scene. These include in gaming, rehabilitation, interaction and selection, and many more (Pereira et al., 2020; Lee et al., 2015; Singh et al., 2021). Therefore this study provides valuable results for many applications in VR, including games or interactive scenarios where multitasking is important. We suggest that when users need to multi task and have attention on two tasks, the background changes may be of larger size before they are noticed. On the other hand, small changes in the scene can well be hidden from user while the user is involved and focused on a hand interaction task that requires full attention. This means subtle reorientations and changes of viewpoint in the scene virtual camera can easily go unseen from the users, when the user's attention is on another task.

Our results revealed that when users were involved in a primary task (shape identification task), whether they used one or two senses, many image transitions that occurred in their field of view were not clearly detectable. On the other hand, when these scene rotations were larger in size, they were more easily discriminated during the gaze versus hand and gaze trials. Even for the largest rotations observers still found it hard to discriminate the image changes. The shape identification task in our experiment was quite demanding, as the participants did not know when the shape would appear, and it would disappear quickly. Therefore they had to keep their gaze and attention on the box to succeed. Performance was good because they were instructed to focus on the shape identification task. In Simons and Jensen (2009), it was found that the demands of the task can affect inattentional blindness rates, where more difficult tasks had lower detection rates of the unexpected objects. They did not find any effects of individual differences on inattentional blindness (Simons and Jensen, 2009).

We noticed that even when participants did not respond correctly to shape identification task, they still did not necessarily select the correct response for scene shift discriminations either. Failure in shape identification does not necessarily mean that participants should succeed in scene shift detection. One reason for incorrect identification of the shape on the box could be because participants were paying attention to the scene shift detection. However, even though the subjects were not able to discriminate the correct shape on the box, they were still, in some cases, unable to correctly detect the scene shifts. This may mean that failure to discriminating the shape on the box, does not imply that the subjects were attending more to the background scene. Based on the findings in our previous experiments, it could be that the events occurred during subjects' blinks or saccades, and due to the suppression of visual information, were not visible. It might also be that subjects forgot what the shape was, or that they failed both tasks because the two events appeared at random times and participants didn't know when to expect each one, and so failed to properly attend.

In a recent study(Coelho et al., 2022), four different tasks with mental or physical loads, were compared to see which ones were best at masking user redirection. They found low detection rates for redirection during all tasks, but more so in the physically demanding ones. On the other hand, there were studies that showed that cognitive performance, and particularly resilience to inattentional blindness, can improve after exercise, but starts to decline after exceeding a work intensity threshold (Hüttermann and Memmert, 2012; Davranche and McMorris, 2009). Other research (Schmelter et al., 2021), found that five interactive tasks, including ”Focus, Pick Up, Throw, Shoot and Fight” could be used as distractors during redirected walking, to apply discrete manipulations. These approaches suggest a possibility of subtly redirecting VR users by involving them in an attentive task.

4.4 Summary

In this chapter I discussed two experiments, one based on saccadic suppression of image displacements and the other on inattentional blindness. Both studies were implemented and conducted on the HTC Vive Pro Eye headset. The goal in both experiments was to investigate how the two perceptual concepts (saccadic suppression and inattentional blindness) could be applied in a VR setting to hide graphical updates from the users. Saccadic suppression of image displacement on a VR HMD can be used as an effective method, especially when co-ocurrs with head rotations, and is a promising method to enable hidden display updates. Likewise, involving the user on an attentionally demand task, was another good way of making the display updates go unseen from the users of a VR.

Chapter 5

Conclusions

The overall aim of this dissertation was to use features of human visual perception in designing better virtual reality, and gaze-contingent content. We mainly focused on saccadic suppression of image displacements, and leveraged the suppression of visual information during our saccades to improve the experience for VR users. This has potential for more effective updating of graphical displays, including re-orientations of the display's viewpoint, during short moments of trans saccadic insensitively. Research in this field is very broad and it is advancing rapidly. Therefore, the results obtained in the experiments discussed in this dissertation contribute to improved gaze interaction in VR, and applications such as redirected walking that create the illusion of large virtual worlds inside small physical spaces. For a more immersive virtual experience, the virtual environments must be rendered in real time within a very tight computational budget, while overcoming constraints on user's limited physical work space in the real world, and without causing negative health consequences such as eye strain and nausea.

5.1 Summary and Conclusions

In chapter 1, I identified several research questions that I was interested in investigating further. Chapter 2 aimed to build on the introduction, and presented some background on the fundamental theories of human visual perception and some of the recent work on gaze-contingent displays and eye tracking in virtual reality. Chapter 3 described two of the gaze contingent experiments I conducted using a research grade eye tracking system. The first experiment focused on saccadic suppression of different image transformations, including rotations and translations on different axes. I studied six translational and six rotational scene changes to explore any directional differences in sensitivity. During horizontal saccades, the most recognizable changes were rotations along the roll axis. The study was very controlled in terms of parameters as the direction and magnitude of user's saccade was known, but the results are valuable and generalizable in VR display designs. Further, I found that the extent of suppression depends on the size and direction of the image transformations that occur during a saccade. The second experiment of Chapter 3, added the variable of a simulated head rotation while the saccade occurred. The results of that study showed that the direction of the head rotation in a virtual scene, has a stronger effect than the saccade direction, on the discrimination of the transsaccadic image shifts. Both experiments were conducted using the precise, high sampling rate, low latency Eyelink eye tracker on a 3D monitor set at a fixed position from the user.

In Chapter 4, I described details of two more experiments that were conducted on a a HTC Vive pro Eye headset a in 3D space. The first experiment of this chapter focused on comparing different transsaccadic and intersaccadic image transformations during head motion. Based on a quantitative analysis of our experiments, it can be concluded that direction of head rotation may be a more important factor (than the direction of a saccade) to consider when designing and targeting imperceptible image shifts in virtual reality. I also found that when the head is still, a person is more sensitive to image transformations. The second experiment of chapter 4 investigated the phenomenon of inattentional blindness during a task that involved movement of the physical hand that moved a virtual object from one location to another. Again, I imposed various image transformations and compared how their detectability differed when a participants used their hand and gaze to move and watch an object compared to when they only watched it move. The results confirmed that when two senses (vision and touch) are involved in a primary task, another visual stimuli can go more unnoticed than when only one sense (vision) is involved. Overall, while there were specific control conditions in my experiments, the approaches and various combinations of scene rotations, saccade directions, head rotations and tasks that I studied can well be generalized in broader VR studies, and can provide new insight into the field of more efficient VR content design.

5.2 Limitations

One technical challenge in current gaze-contingent displays is latency. The real-time synchronization of eye movements and the image displayed is crucial in such applications. Latencies are caused due to the frame rate of the content being displayed, saccade detection and prediction algorithms, eye tracker delays and the refresh rate of the display (Arabadzhiyska et al., 2017; Stein et al., 2021). This may be even more apparent when viewers make fast saccadic eye movements during which the information about gaze location is significantly delayed and the mismatch can be noticed on saccade landing. One solution that may help in reducing is using foveated rendering. In a foveated rendering display, when the eye moves, the location of this foveal area changes and the changing resolution is potentially perceptible. In general, these changes occur during a saccade and thus saccadic suppression helps reduce the visibility of the changes in image detail (Patney, 2017; Albert et al., 2017). Also, programmed saccades are easier to detect and analyse than natural saccades which could occur at any time and with any duration. In this case, one possible solution is predicting saccade landing position, which has been studied with different methods (Arabadzhiyska et al., 2022; Arabadzhiyska et al., 2017; Zhang et al., 2018; Morales et al., 2018). In applications of redirected walking, foveated rendering can be used to render the part of the graphical scene that the user is looking at in full resolution. Also, the display can be rendered in low resolution during a saccade as it can be hidden from the user. By predicting the landing point of a generated saccade, the VR viewpoint can be modified before the saccade has landed.

5.3 Future Research Directions

In future research, I would like to experiment with more natural viewing of a virtual environment with less controlled conditions with saccades of arbitrary sizes and directions, and also test in real applications of VR, such as games or tours that require a lot of navigation, or multiplayer VR. Although using a research grade high-sampling-frequency eye tracker gives more accurate results with a high number of eye samples per second, most VR setting require the user to wear an HMD, which often have lower frequencies for sampling eye data. As some users in experiments of chapter 3 reported guessing the direction of movement after the saccade by the static visual cues in the scene, we can also have conditions that reduce these visual cues in the scene such as occlusion of objects, shadows or arrangement of the objects. Also, I can look at effect of visual depth cues, such as occlusion, shadows, straight lines or arrangement of the objects in the scene on the detection of the scene transitions.

Further research may be required to determine the effect of other parameters on the amount of saccadic suppression. For example, head translations, and the effect of motion parallax in a VR scene could be explored. Based on these conclusions, designers of VR environments should consider how they can take advantage of head rotations, saccadic eye movements, and task designs. There should different ways to trigger saccades with specific features (for example, larger ones for more suppression), and head movements on certain axes or directions. Also generalizing this research to both saccades and blinks, and taking advantage of both could be beneficial in VR. Therefore, more gaze-contingent paradigms such as depth-contingent and auditory GCDs can be studied. Correct use of depth cues, such as blur or accommodation, need to be considered in the design of such paradigms. Moreover, due to the real-time performance requirement of gaze-contingent displays and high frame rates of VR applications, the possibility of using artificial intelligence and machine learning techniques in such displays has not been explored in depth. As well as deep learning approaches to offline modelling of eye movements patterns and gaze predictions, improvements in real-time interface using deep learning methods may provide improved saccade detection and estimation in the future.

Moreover, further research and future studies can combine inattentional blindness tasks, with saccade or blink-based display changes to increase the imperceptible thresholds and create a more natural redirection or reorientation in VR. Also, we can investigate whether the type of the task affects this sensitivity. In addition, a follow up experiment can test inattentional blindness during non-visual activities, such as an audio-related tasks, or during conversations with a character in the scene, and see how that affects the scene shift discriminations. For example, the results of (Lee et al., 2022) showed that auditory and olfactory attractors could be used to redirect users of VR to the desired path. The auditory attractor provided more natural and immersive reorientation, while the olfactory lead to more turn changes and larger rotational gains.

Virtual Reality and Augmented Reality are being used in many different fields, and are not just limited to entertainment and games. Human factors such as gaze behaviour and eye movements are important in the progress of these technologies. Eye tracking and haptics are two very common approaches for sensing behavioral actions of humans in VR. Designers of VR displays that create a high sense of presence for their users, and allow natural interaction should consider perceptual concepts, such as saccadic or blink suppression, and inattentional or change blindness.

References

Abbasov, I. B. (2019). Psychology of visual perception. *Amazon Digital Services LLC*.

Abrams, R. A., Meyer, D. E., and Kornblum, S. (1990). Eye-hand coordination: oculomotor control in rapid aimed limb movements. *Journal of experimental psychology: human perception and performance*, 16(2):248.

Albert, R., Patney, A., Luebke, D., and Kim, J. (2017). Latency requirements for foveated rendering in virtual reality. *ACM Transactions on Applied Perception (TAP)*, 14(4):1–13.

Albert, R. A., Godinez, A., and Luebke, D. (2019). Reading speed decreases for fast readers under gaze-contingent rendering. In *ACM Symposium on Applied Perception 2019*, pages 1–6.

Ali, M. and Cardona-Rivera, R. E. (2020). Comparing gamepad and naturally-mapped controller effects on perceived virtual reality experiences. In *ACM Symposium on Applied Perception 2020*, pages 1–10.

Allison, R. S., Howard, I. P., and Zacher, J. E. (1999). Effect of field size, head motion, and rotational velocity on roll vection and illusory self-tilt in a tumbling room. *Perception*, 28(3):299–306.

Allison, R. S., Schumacher, J., Sadr, S., and Herpers, R. (2010). Apparent motion during saccadic suppression periods. *Experimental brain research*, 202(1):155–169.

Andersson, R., Larsson, L., Holmqvist, K., Stridh, M., and Nyström, M. (2017). One algorithm to rule them all? an evaluation and discussion of ten eye movement event-detection algorithms. *Behavior research methods*, 49(2):616–637.

Angelaki, D. E. and Hess, B. J. (2005). Self-motion-induced eye movements: effects on visual acuity and navigation. *Nature Reviews Neuroscience*, 6(12):966–976.

Anson, E. R., Bigelow, R. T., Carey, J. P., Xue, Q.-L., Studenski, S., Schubert, M. C., Weber, K. P., and Agrawal, Y. (2016). Aging increases compensatory saccade amplitude in the video head impulse test. *Frontiers in neurology*, 7:113.

Arabadzhiyska, E., Tursun, C., Seidel, H.-P., and Didyk, P. (2022). Practical saccade prediction for head-mounted displays: Towards a comprehensive model. *arXiv preprint arXiv:2205.01624*.

Arabadzhiyska, E., Tursun, O. T., Myszkowski, K., Seidel, H.-P., and Didyk, P. (2017). Saccade landing position prediction for gaze-contingent rendering. *ACM Transactions on Graphics (TOG)*, 36(4):1–12.

Auvray, M., Gallace, A., Tan, H. Z., and Spence, C. (2007). Crossmodal change blindness between vision and touch. *Acta Psychologica*, 126(2):79–97.

Bahill, A. T., Clark, M. R., and Stark, L. (1975). The main sequence, a tool for studying human eye movements. *Mathematical biosciences*, 24(3-4):191–204.

Bellet, M. E., Bellet, J., Nienborg, H., Hafed, Z. M., and Berens, P. (2019). Human-level saccade detection performance using deep neural networks. *Journal of neurophysiology*, 121(2):646–661.

Berencsi, A., Ishihara, M., and Imanaka, K. (2005). The functional role of central and peripheral vision in the control of posture. *Human movement science*, 24(5-6):689–709.

Berthoz, A. and Droulez, J. (1982). Linear self motion perception. In *Tutorials on motion perception*, pages 157–199. Springer.

Binda, P. and Morrone, M. C. (2018). Vision during saccadic eye movements. *Annual review of vision science*, 4:193–213.

Bittencourt, P., Wade, P., Smith, A., and Richens, A. (1981). The relationship between peak velocity of saccadic eye movements and serum benzodiazepine concentration. *British journal of clinical pharmacology*, 12(4):523–533.

Blouin, J., Bridgeman, B., Teasdale, N., Bard, C., and Fleury, M. (1995). Visual stability with goal-directed eye and arm movements toward a target displaced during saccadic suppression. *Psychological Research*, 58(3):169–176.

Bohrn, I., Carbon, C.-C., and Hutzler, F. (2010). Mona lisa's smile—perception or deception? *Psychological science*, 21(3):378–380.

Bolte, B. and Lappe, M. (2015). Subliminal reorientation and repositioning in immersive virtual environments using saccadic suppression. *IEEE transactions on visualization and computer graphics*, 21(4):545–552.

Bombari, D., Schmid Mast, M., Canadas, E., and Bachmann, M. (2015). Studying social interactions through immersive virtual environment technology: virtues, pitfalls, and future challenges. *Frontiers in psychology*, 6:869.

Boraston, Z. and Blakemore, S.-J. (2007). The application of eye-tracking technology in the study of autism. *The Journal of physiology*, 581(3):893–898.

Born, S. (2019). Saccadic suppression of displacement does not reflect a saccade-specific bias to assume stability. *Vision*, 3(4):49.

Bowman, D. A. and Hodges, L. F. (1997). An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In *Proceedings of the 1997 symposium on Interactive 3D graphics*, pages 35–ff.

Bowman, D. A., Koller, D., and Hodges, L. F. (1997). Travel in immersive virtual environments: An evaluation of viewpoint motion control techniques. In *Proceedings of IEEE 1997 Annual International Symposium on Virtual Reality*, pages 45–52. IEEE.

Bremmer, F., Kubischik, M., Hoffmann, K.-P., and Krekelberg, B. (2009). Neural dynamics of saccadic suppression. *Journal of Neuroscience*, 29(40):12374–12383.

Bridgeman, B. (2012). Eye movements. *Encyclopedia of human behavior*, 2:160–166.

Bridgeman, B. and Fisher, B. (1990). Saccadic suppression of displacement is strongest in central vision. *Perception*, 19(1):103–111.

Bridgeman, B., Hendry, D., and Stark, L. (1975). Failure to detect displacement of the visual world during saccadic eye movements. *Vision research*, 15(6):719–722.

Bridgeman, B. and Macknik, S. L. (1995). Saccadic suppression relies on luminance information. *Psychological research*, 58(3):163–168.

Bridgeman, B., Van der Heijden, A., and Velichkovsky, B. M. (1994). A theory of visual stability across saccadic eye movements. *Behavioral and Brain Sciences*, 17(2):247–258.

Bruder, G. and Langbehn, E. (2017). Subliminal rotations during eye blinks for redirected walking. *Journal of Vision*, 17(10):1266–1266.

Bruno, A., Brambati, S. M., Perani, D., and Morrone, M. C. (2006). Development of saccadic suppression in children. *Journal of neurophysiology*, 96(3):1011–1017.

Burns, E., Razzaque, S., Panter, A. T., Whitton, M. C., McCallus, M. R., and Brooks, F. P. (2005). The hand is slower than the eye: A quantitative exploration of visual dominance over proprioception. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.*, pages 3–10. IEEE.

Burr, D. C. and Morrone, M. C. (2010). Vision: keeping the world still when the eyes move. *Current Biology*, 20(10):R442–R444.

Carbon, C.-C. (2014). Understanding human perception by human-made illusions. *Frontiers in human neuroscience*, 8:566.

Cartwright-Finch, U. and Lavie, N. (2007). The role of perceptual load in inattentional blindness. *Cognition*, 102(3):321–340.

Chakravarthula, P., Zhang, Z., Tursun, O., Didyk, P., Sun, Q., and Fuchs, H. (2021). Gaze-contingent retinal speckle suppression for perceptually-matched foveated holographic displays. *IEEE Transactions on Visualization and Computer Graphics*, 27(11):4194–4203.

Chappell, R. (2015). System and method for eye tracking authentication. *International Patent application, publication number: US20150227735 A*, 1.

Chen, S., Duinkharjav, B., Sun, X., Wei, L.-Y., Petrangeli, S., Echevarria, J., Silva, C., and Sun, Q. (2022). Instant reality: Gaze-contingent perceptual optimization for 3d virtual reality streaming. *IEEE Transactions on Visualization and Computer Graphics*, 28(5):2157–2167.

Chen, Y., Sun, J., Xu, Q., Lank, E., Irani, P., and Li, W. (2021). Empirical evaluation of moving target selection in virtual reality using egocentric metaphors. In *IFIP Conference on Human-Computer Interaction*, pages 29–50. Springer.

Chien, K.-P., Tsai, C.-Y., Chen, H.-L., Chang, W.-H., and Chen, S. (2015). Learning differences and eye fixation patterns in virtual and physical science laboratories. *Computers & Education*, 82:191–201.

Coelho, D. N., Steinicke, F., and Langbehn, E. (2022). Design of mentally and physically demanding tasks as distractors of rotation gains. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pages 620–621. IEEE.

Collewijn, H. and Smeets, J. B. (2000). Early components of the human vestibulo-ocular response to head rotation: latency and gain. *Journal of Neurophysiology*, 84(1):376–389.

Connelly, L., Jia, Y., Toro, M. L., Stoykov, M. E., Kenyon, R. V., and Kamper, D. G. (2010). A pneumatic glove and immersive virtual reality environment for hand rehabilitative training after stroke. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(5):551–559.

Crevecoeur, F. and Kording, K. P. (2017). Saccadic suppression as a perceptual consequence of efficient sensorimotor estimation. *Elife*, 6:e25073.

Currie, C. B., McConkie, G. W., Carlson-Radvansky, L. A., and Irwin, D. E. (2000). The role of the saccade target object in the perception of a visually stable world. *Perception & psychophysics*, 62(4):673–683.

Da Silva, M. P., Courboulay, V., and Prigent, A. (2007). Gameplay experience based on a gaze tracking system. In *"Gaze-based Creativity, Interacting with Games and On-line Communities" INPROCEEDINGS in proceedings of CO-GAIN 2007 (Communication by Gaze Interaction IST FP6 European Project)*, pages 25–28.

Davranche, K. and McMorris, T. (2009). Specific effects of acute moderate exercise on cognitive control. *Brain and cognition*, 69(3):565–570.

Daye, P. M., Roberts, D. C., Zee, D. S., and Optican, L. M. (2015). Vestibulo-ocular reflex suppression during head-fixed saccades reveals gaze feedback control. *Journal of Neuroscience*, 35(3):1192–1198.

Deubel, H., Schneider, W. X., and Bridgeman, B. (1996). Postsaccadic target blanking prevents saccadic suppression of image displacement. *Vision research*, 36(7):985–996.

Diamond, M. R., Ross, J., and Morrone, M. C. (2000). Extraretinal control of saccadic suppression. *Journal of Neuroscience*, 20(9):3449–3455.

Duchowski, A. (2007). Eye tracking techniques. In *Eye tracking methodology*, pages 51–59. Springer.

Duchowski, A. T. (2002). A breadth-first survey of eye-tracking applications. *Behavior Research Methods, Instruments, & Computers*, 34(4):455–470.

Duchowski, A. T. and Çöltekin, A. (2007). Foveated gaze-contingent displays for peripheral lod management, 3d visualization, and stereo imaging. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 3(4):1–18.

Duchowski, A. T., Cournia, N., and Murphy, H. (2004). Gaze-contingent displays: A review. *CyberPsychology & Behavior*, 7(6):621–634.

Duchowski, A. T. and Duchowski, A. T. (2017). *Eye tracking methodology: Theory and practice*. Springer.

Duchowski, A. T., House, D. H., Gestring, J., Wang, R. I., Krejtz, K., Krejtz, I., Mantiuk, R., and Bazyluk, B. (2014). Reducing visual discomfort of 3d stereoscopic displays with gaze-contingent depth-of-field. In *Proceedings of the acm symposium on applied perception*, pages 39–46.

E. Jacobs, D., Gallo, O., A. Cooper, E., Pulli, K., and Levoy, M. (2015). Simulating the visual experience of very bright and very dark scenes. *ACM Transactions on Graphics (TOG)*, 34(3):1–15.

Eizenman, M., Frecker, R., and Hallett, P. (1984). Precise non-contacting measurement of eye movements using the corneal reflex. *Vision research*, 24(2):167–174.

Elias, Z. M., Batumalai, U. M., and Azmi, A. N. H. (2019). Virtual reality games on accommodation and convergence. *Applied ergonomics*, 81:102879.

Ellis, S. R., Adelstein, B. D., Baumeler, S., Jense, G., and Jacoby, R. H. (1999). Sensor spatial distortion, visual latency, and update rate effects on 3d tracking in virtual environments. In *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)*, pages 218–221. IEEE.

Engbert, R. and Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision research*, 43(9):1035–1045.

Engbert, R. and Mergenthaler, K. (2006). Microsaccades are triggered by low retinal image slip. *Proceedings of the National Academy of Sciences*, 103(18):7192–7197.

Fahmi, F., Tanjung, K., Nainggolan, F., Siregar, B., Mubarakah, N., and Zarlis, M. (2020). Comparison study of user experience between virtual reality controllers, leap motion controllers, and senso glove for anatomy learning systems in a virtual reality environment. In *IOP Conference Series: Materials Science and Engineering*, volume 851, page 012024. IOP Publishing.

Findlay, J. M., Findlay, J. M., Gilchrist, I. D., et al. (2003). *Active vision: The psychology of looking and seeing*. Number 37. Oxford University Press.

Franke, I. S., Günther, T., and Groh, R. (2014). Saccade detection and processing for enhancing 3d visualizations in real-time. In *International conference on human-computer interaction*, pages 317–322. Springer.

Freedman, E. G. and Sparks, D. L. (1997). Eye-head coordination during head-unrestrained gaze shifts in rhesus monkeys. *Journal of neurophysiology*, 77(5):2328–2348.

Freedman, E. G. and Sparks, D. L. (2000). Coordination of the eyes and head: movement kinematics. *Experimental brain research*, 131(1):22–32.

Frens, M. and Erkelens, C. (1991). Coordination of hand movements and saccades: evidence for a common and a separate pathway. *Experimental Brain Research*, 85(3):682–690.

Gancarz, G. and Grossberg, S. (1999). A neural model of saccadic eye movement control explains task-specific adaptation. *Vision research*, 39(18):3123–3143.

Gao, P., Matsumoto, K., Narumi, T., and Hirose, M. (2020). Visual-auditory redirection: Multimodal integration of incongruent visual and auditory cues for redirected walking. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 639–648. IEEE.

Gibaldi, A. and Banks, M. S. (2019). Binocular eye movements are adapted to the natural environment. *Journal of Neuroscience*, 39(15):2877–2888.

Gibson, J. J. (1950). The perception of the visual world.

Gloriani, A. H. and Schütz, A. C. (2019). Humans trust central vision more than peripheral vision even in the dark. *Current Biology*, 29(7):1206–1210.

Gori, M., Del Viva, M., Sandini, G., and Burr, D. C. (2008). Young children do not integrate visual and haptic form information. *Current Biology*, 18(9):694–698.

Gregory, R. L. and Zangwill, O. L. (1963). The origin of the autokinetic effect. *Quarterly Journal of Experimental Psychology*, 15(4):252–261.

Gremmler, S. and Lappe, M. (2017). Saccadic suppression during voluntary versus reactive saccades. *Journal of vision*, 17(8):8–8.

Griffith, H., Biswas, S., and Komogortsev, O. (2018). Towards reduced latency in saccade landing position prediction using velocity profile methods. In *Proceedings of the Future Technologies Conference*, pages 79–91. Springer.

Grossberg, S., Roberts, K., Aguilar, M., and Bullock, D. (1997). A neural model of multimodal adaptive saccadic eye movement control by superior colliculus. *Journal of Neuroscience*, 17(24):9706–9725.

Grossman, G. E., Leigh, R. J., Abel, L. A., Lanska, D. J., and Thurston, S. (1988). Frequency and velocity of rotational head perturbations during locomotion. *Experimental brain research*, 70(3):470–476.

Guitton, D. and Volle, M. (1987). Gaze control in humans: eye-head coordination during orienting movements to targets within and beyond the oculomotor range. *Journal of neurophysiology*, 58(3):427–459.

Han, D. T., Suhail, M., and Ragan, E. D. (2018). Evaluating remapped physical reach for hand interactions with passive haptics in virtual reality. *IEEE transactions on visualization and computer graphics*, 24(4):1467–1476.

Hansen, D. W. and Ji, Q. (2009). In the eye of the beholder: A survey of models for eyes and gaze. *IEEE transactions on pattern analysis and machine intelligence*, 32(3):478–500.

Helmholtz, H. v. (1985). Treatise on physiological optics, volume iii: the perceptions of vision (jpc southall, trad.). *Birmingham, Inglaterra: Gryphon.(Original publicado em 1867).*

Henderson, J. M. (2007). Regarding scenes. *Current directions in psychological science*, 16(4):219–222.

Herpers, R., Schumacher, J., and Allison, R. S. (2004). Hiding graphic updates during long saccadic suppression periods. *Dyn. Percept*, pages 77–82.

Hopp, J. J. and Fuchs, A. F. (2004). The characteristics and neuronal substrate of saccadic eye movement plasticity. *Progress in neurobiology*, 72(1):27–53.

Howard, I. P. and Childerson, L. (1994). The contribution of motion, the visual frame, and visual polarity to sensations of body tilt. *Perception*, 23(7):753–762.

Hubel, D. H. (1995). *Eye, brain, and vision.* Scientific American Library/Scientific American Books.

Hüttermann, S. and Memmert, D. (2012). Moderate movement, more vision: Effects of physical exercise on inattentional blindness. *Perception*, 41(8):963–975.

Ibbotson, M. and Krekelberg, B. (2011). Visual perception and saccadic eye movements. *Current opinion in neurobiology*, 21(4):553–558.

Idrees, S. (2021). *Saccadic suppression by way of retinal image processing.* PhD thesis, Universität Tübingen.

Janky, K. L., Patterson, J., Shepard, N., Thomas, M., Barin, K., Creutz, T., Schmid, K., and Honaker, J. A. (2018). Video head impulse test (vhit): The role of corrective saccades in identifying patients with vestibular loss. *Otology & neurotology: official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 39(4):467.

Jensen, M. S., Yao, R., Street, W. N., and Simons, D. J. (2011). Change blindness and inattentional blindness. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(5):529–546.

Jerald, J., Peck, T., Steinicke, F., and Whitton, M. (2008). Sensitivity to scene motion for phases of head yaws. In *Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization*, pages 155–162.

Johansson, R. S., Westling, G., Bäckström, A., and Flanagan, J. R. (2001). Eye–hand coordination in object manipulation. *Journal of Neuroscience*, 21(17):6917–6932.

Joshi, Y. and Poullis, C. (2020a). Dynamic foveated rendering for redirected walking in virtual reality. In *ACM SIGGRAPH 2020 Posters*, pages 1–2.

Joshi, Y. and Poullis, C. (2020b). Inattentional blindness for redirected walking using dynamic foveated rendering. *IEEE Access*, 8:39013–39024.

Joshi, Y. and Poullis, C. (2022). Saccadenet: Towards real-time saccade prediction for virtual reality infinite walking. *arXiv preprint arXiv:2205.15846*.

Just, M. A. and Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive psychology*, 8(4):441–480.

Karlsson, P., Allsop, A., Dee-Price, B.-J., and Wallen, M. (2018). Eye-gaze control technology for children, adolescents and adults with cerebral palsy with significant physical disability: Findings from a systematic review. *Developmental neurorehabilitation*, 21(8):497–505.

Kasneci, E., Black, A. A., and Wood, J. M. (2017). Eye-tracking as a tool to evaluate functional ability in everyday tasks in glaucoma. *Journal of ophthalmology*, 2017.

Kellnhofer, P., Didyk, P., Myszkowski, K., Hefeeda, M. M., Seidel, H.-P., and Matusik, W. (2016). Gazestereo3d: Seamless disparity manipulations. *ACM Transactions on Graphics (TOG)*, 35(4):1–13.

Keyvanara, M. and Allison, R. (2019). Saccadic suppression of natural image transformations. In *Perception*, volume 48, pages 71–71. SAGE PUBLICATIONS LTD 1 OLIVERS YARD, 55 CITY ROAD, LONDON EC1Y 1SP, ENGLAND.

Keyvanara, M. and Allison, R. (2020). Effect of a constant camera rotation on the visibility of transsaccadic camera shifts. In *ACM Symposium on Eye Tracking Research and Applications*, pages 1–8.

Khademi, M., Mousavi Hondori, H., McKenzie, A., Dodakian, L., Lopes, C. V., and Cramer, S. C. (2014). Free-hand interaction with leap motion controller for stroke rehabilitation. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*, pages 1663–1668.

Khakhar, R., You, F., Chakkalakal, D., Dobbelstein, D., and Picht, T. (2021). Hands-free adjustment of the microscope in microneurosurgery. *World Neurosurgery*, 148:e155–e163.

Khundam, C., Vorachart, V., Preeyawongsakul, P., Hosap, W., and Noël, F. (2021). A comparative study of interaction time and usability of using controllers and hand tracking in virtual reality training. In *Informatics*, volume 8, page 60. Multidisciplinary Digital Publishing Institute.

Kim, J., Jeong, Y., Stengel, M., Akşit, K., Albert, R., Boudaoud, B., Greer, T., Kim, J., Lopes, W., Majercik, Z., et al. (2019). Foveated ar: dynamically-foveated augmented reality display. *ACM Transactions on Graphics (TOG)*, 38(4):1–15.

Kim, S.-W., Kim, S. H., Kim, C. S., Yi, K., Kim, J.-S., Cho, B. J., and Cha, Y. (2020). Thermal display glove for interacting with virtual reality. *Scientific reports*, 10(1):1–12.

Klein, S. A. (2001). Measuring, estimating, and understanding the psychometric function: A commentary. *Perception & psychophysics*, 63(8):1421–1455.

Konrad, R., Angelopoulos, A., and Wetzstein, G. (2020). Gaze-contingent ocular parallax rendering for virtual reality. *ACM Transactions on Graphics (TOG)*, 39(2):1–12.

Koulieris, G. A., Drettakis, G., Cunningham, D., and Mania, K. (2016). Gaze prediction using machine learning for dynamic stereo manipulation in games. In *2016 IEEE virtual reality (VR)*, pages 113–120. IEEE.

Kruchinina, A., Chertopolokhov, V., Grigoryan, G., and Belyaev, V. (2020). 54-4: Prediction of saccadic eye movement for foveated rendering. In *SID Symposium Digest of Technical Papers*, volume 51, pages 803–806. Wiley Online Library.

Kuss, M., Jäkel, F., and Wichmann, F. A. (2005). Bayesian inference for psychometric functions. *Journal of Vision*, 5(5):8–8.

Lahiri, U., Warren, Z., and Sarkar, N. (2011). Design of a gaze-sensitive virtual social interactive system for children with autism. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 19(4):443–452.

Land, M. F. (2004). The coordination of rotations of the eyes, head and trunk in saccadic turns produced in natural situations. *Experimental brain research*, 159(2):151–160.

Langbehn, E., Bruder, G., and Steinicke, F. (2016). Subliminal reorientation and repositioning in virtual reality during eye blinks. In *Proceedings of the 2016 symposium on spatial user interaction*, pages 213–213.

Langbehn, E., Steinicke, F., Lappe, M., Welch, G. F., and Bruder, G. (2018). In the blink of an eye: leveraging blink-induced suppression for imperceptible position and orientation redirection in virtual reality. *ACM Transactions on Graphics (TOG)*, 37(4):1–11.

Lanman, J., Bizzi, E., and Allum, J. (1978). The coordination of eye and head movement during smooth pursuit. *Brain research*, 153(1):39–53.

Lappe, M., Bremmer, F., and Van den Berg, A. (1999). Perception of self-motion from visual flow. *Trends in cognitive sciences*, 3(9):329–336.

Lappe, M., Hoffmann, K.-P., et al. (2000). Optic flow and eye movements. *International review of neurobiology*, pages 29–50.

Lappi, O. (2016). Eye movements in the wild: Oculomotor control, gaze behavior & frames of reference. *Neuroscience & Biobehavioral Reviews*, 69:49–68.

Lee, J., Hwang, S., Kim, K., and Kim, S. (2022). Auditory and olfactory stimuli-based attractors to induce reorientation in virtual reality forward redirected walking. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, pages 1–7.

Lee, J., Sinclair, M., Gonzalez-Franco, M., Ofek, E., and Holz, C. (2019). Torc: A virtual reality controller for in-hand high-dexterity finger interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–13.

Lee, P.-W., Wang, H.-Y., Tung, Y.-C., Lin, J.-W., and Valstar, A. (2015). Transection: hand-based interaction for playing a game within a virtual reality game. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, pages 73–76.

Lefèvre, P., Bottemanne, I., and Roucoux, A. (1992). Experimental study and modeling of vestibulo-ocular reflex modulation during large shifts of gaze in humans. *Experimental brain research*, 91(3):496–508.

Leigh, R. J. and Zee, D. S. (2015). *The neurology of eye movements*. Contemporary Neurology.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical society of America*, 49(2B):467–477.

Levoy, M. and Whitaker, R. (1990). Gaze-directed volume rendering. In *Proceedings of the 1990 symposium on Interactive 3D graphics*, pages 217–223.

Li, W. and Matin, L. (1990). The influence of saccade length on the saccadic suppression of displacement detection. *Perception & Psychophysics*, 48(5):453–458.

Lieberman, H. R. and Pentland, A. P. (1982). Microcomputer-based estimation of psychophysical thresholds: the best pest. *Behavior Research Methods & Instrumentation*, 14(1):21–25.

Liversedge, S. P. and Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in cognitive sciences*, 4(1):6–14.

Loschky, L. C. and McConkie, G. W. (2000). User performance with gaze contingent multiresolutional displays. In *Proceedings of the 2000 symposium on Eye tracking research & applications*, pages 97–103.

Loschky, L. C. and McConkie, G. W. (2002). Investigating spatial vision and dynamic attentional selection using a gaze-contingent multiresolutional display. *Journal of Experimental Psychology: Applied*, 8(2):99.

Lünenburger, L., Kutz, D., and Hoffmann, K.-P. (2000). Influence of arm movements on saccades in humans. *European Journal of Neuroscience*, 12(11):4107–4116.

Luo, S., Nguyen, A., Song, C., Lin, F., Xu, W., and Yan, Z. (2020). Oculock: Exploring human visual system for authentication in virtual reality head-mounted display. In *2020 Network and Distributed System Security Symposium (NDSS)*.

Luxenburger, A., Prange, A., Moniri, M. M., and Sonntag, D. (2016). Medicalvr: Towards medical remote collaboration using virtual reality. In *Proceedings of* *the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, pages 321–324.

Mack, A. (2003). Inattentional blindness: Looking without seeing. *Current directions in psychological science*, 12(5):180–184.

Mack, A. and Rock, I. (1998). Inattentional blindness: Perception without attention. *Visual attention*, 8:55–76.

Madigan, R. and Williams, D. (1987). Maximum-likelihood psychometric procedures in two-alternative forced-choice: Evaluation and recommendations. *Perception & Psychophysics*, 42(3):240–249.

Malpica, S., Serrano, A., Gutierrez, D., and Masia, B. (2020). Auditory stimuli degrade visual performance in virtual reality. *Scientific reports*, 10(1):1–9.

Mardanbegi, D., Hansen, D. W., and Pederson, T. (2012). Eye-based head gestures. In *Proceedings of the symposium on eye tracking research and applications*, pages 139–146.

Martinez-Conde, S. (2006). Fixational eye movements in normal and pathological vision. *Progress in brain research*, 154:151–176.

Martinez-Conde, S., Macknik, S. L., and Hubel, D. H. (2004). The role of fixational eye movements in visual perception. *Nature reviews neuroscience*, 5(3):229–240.

Matin, E. (1974). Saccadic suppression: a review and an analysis. *Psychological bulletin*, 81(12):899.

Mauderer, M., Conte, S., Nacenta, M. A., and Vishwanath, D. (2014). Depth perception with gaze-contingent depth of field. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 217–226.

Mauderer, M., Flatla, D. R., and Nacenta, M. A. (2016). Gaze-contingent manipulation of color perception. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 5191–5202.

Maus, G. W., Duyck, M., Lisi, M., Collins, T., Whitney, D., and Cavanagh, P. (2017). Target displacements during eye blinks trigger automatic recalibration of gaze direction. *Current biology*, 27(3):445–450.

McConkie, G. W. and Currie, C. B. (1996). Visual stability across saccades while viewing complex pictures. *Journal of Experimental Psychology: Human perception and performance*, 22(3):563.

McConkie, G. W. and Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, 17(6):578–586.

Mihali, A., van Opheusden, B., and Ma, W. J. (2017). Bayesian microsaccade detection. *Journal of vision*, 17(1):13–13.

Morales, A., Costela, F. M., Tolosana, R., and Woods, R. L. (2018). Saccade landing point prediction: A novel approach based on recurrent neural networks. In *Proceedings of the 2018 International Conference on Machine Learning Technologies*, pages 1–5.

Morales, A., Costela, F. M., and Woods, R. L. (2021). Saccade landing point prediction based on fine-grained learning method. *IEEE Access*, 9:52474–52484.

Morasso, P., Bizzi, E., and Dichgans, J. (1973). Adjustment of saccade characteristics during head movements. *Experimental Brain Research*, 16(5):492–500.

Munoz, D. P. (2002). Commentary: saccadic eye movements: overview of neural circuitry. *Progress in brain research*, 140:89–96.

Nardini, M., Jones, P., Bedford, R., and Braddick, O. (2008). Development of cue integration in human navigation. *Current biology*, 18(9):689–693.

Navalpakkam, V. and Churchill, E. F. (2014). Eye tracking: a brief introduction. In *Ways of Knowing in HCI*, pages 323–348. Springer.

Negi, S. and Mitra, R. (2020). Fixation duration and the learning process: an eye tracking study with subtitled videos. *Journal of Eye Movement Research*, 13(6).

Nguyen, A. and Kunz, A. (2018). Discrete scene rotation during blinks and its effect on redirected walking algorithms. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, pages 1–10.

Nguyen, A., Rothacher, Y., Lenggenhager, B., Brugger, P., and Kunz, A. (2018a). Individual differences and impact of gender on curvature redirection thresholds. In *Proceedings of the 15th acm symposium on applied perception*, pages 1–4.

Nguyen, A., Yan, Z., and Nahrstedt, K. (2018b). Your attention is unique: Detecting 360-degree video saliency in head-mounted display for head movement prediction. In *Proceedings of the 26th ACM international conference on Multimedia*, pages 1190–1198.

Nilsson, N. C., Peck, T., Bruder, G., Hodgson, E., Serafin, S., Whitton, M., Steinicke, F., and Rosenberg, E. S. (2018). 15 years of research on redirected walking in immersive virtual environments. *IEEE computer graphics and applications*, 38(2):44–56.

Nilsson, S., Gustafsson, T., and Carleberg, P. (2009). Hands free interaction with virtual information in a real environment: Eye gaze as an interaction tool in an augmented reality system. *PsychNology Journal*, 7(2).

Norouzi, N., Bölling, L., Bruder, G., and Welch, G. (2019). Augmented rotations in virtual reality for users with a reduced range of head movement. *Journal of rehabilitation and assistive technologies engineering*, 6:2055668319841309.

Nyström, M. and Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior research methods*, 42(1):188–204.

Otto, S. and Weinzierl, S. (2009). Comparative simulations of adaptive psychometric procedures. *Jahrestagung der Deutschen Gesellschaft für Akustik*, pages 1276–1279.

Patney, A. (2017). Perceptual insights into foveated virtual reality. In *NVIDIA GPU Technology Conference 2017 Talks*.

Pentland, A. (1980). Maximum likelihood estimation: the best pest. *Perception & psychophysics*.

Pereira, M. F., Prahm, C., Kolbenschlag, J., Oliveira, E., and Rodrigues, N. F. (2020). Application of ar and vr in hand rehabilitation: A systematic review. *Journal of Biomedical Informatics*, 111:103584.

Perret, J. and Vander Poorten, E. (2018). Touching virtual reality: a review of haptic gloves. In *ACTUATOR 2018; 16th International Conference on New Actuators*, pages 1–5. VDE.

Pfeil, K., Taranta, E. M., Kulshreshth, A., Wisniewski, P., and LaViola Jr, J. J. (2018). A comparison of eye-head coordination between virtual and physical realities. In *Proceedings of the 15th ACM Symposium on Applied Perception*, pages 1–7.

Pinson, E., Pietroszek, K., Sun, Q., and Eckhardt, C. (2020). An open framework for infinite walking with saccadic redirection. In *26th ACM Symposium on Virtual Reality Software and Technology*, pages 1–2.

Piumsomboon, T., Lee, G., Lindeman, R. W., and Billinghurst, M. (2017). Exploring natural eye-gaze-based interaction for immersive virtual reality. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 36–39. IEEE.

Poupyrev, I., Ichikawa, T., Weghorst, S., and Billinghurst, M. (1998). Egocentric object manipulation in virtual environments: empirical evaluation of interaction techniques. In *Computer graphics forum*, volume 17, pages 41–52. Wiley Online Library.

Prablanc, C., Echallier, J., Komilis, E., and Jeannerod, M. (1979). Optimal response of eye and hand motor systems in pointing at a visual target. *Biological cybernetics*, 35(2):113–124.

Prado, J., Clavagnier, S., Otzenberger, H., Scheiber, C., Kennedy, H., and Perenin, M.-T. (2005). Two cortical systems for reaching in central and peripheral vision. *Neuron*, 48(5):849–858.

Rayner, K. and Pollatsek, A. (1992). Eye movements and scene perception. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 46(3):342.

Razzaque, S. (2005). *Redirected walking*. The University of North Carolina at Chapel Hill.

Reder, S. M. (1973). On-line monitoring of eye-position signals in contingent and noncontingent paradigms. *Behavior Research Methods & Instrumentation*, 5(2):218–228.

Richards, A., Hannon, E. M., and Vitkovitch, M. (2012). Distracted by distractors: Eye movements in a dynamic inattentional blindness task. *Consciousness and cognition*, 21(1):170–176.

Rietzler, M., Deubzer, M., Dreja, T., and Rukzio, E. (2020). Telewalk: Towards free and endless walking in room-scale virtual reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–9.

Riggs, L. A., Armington, J. C., and Ratliff, F. (1954). Motions of the retinal image during fixation. *JOSA*, 44(4):315–321.

Riggs, L. A., Volkmann, F. C., and Moore, R. K. (1981). Suppression of the blackout due to blinks. *Vision Research*, 21(7):1075–1079.

Rucci, M. and Poletti, M. (2015). Control and functions of fixational eye movements. *Annual review of vision science*, 1:499–518.

Rucci, M. and Victor, J. D. (2015). The unsteady eye: an information-processing stage, not a bug. *Trends in neurosciences*, 38(4):195–206.

Ruhland, K., Peters, C. E., Andrist, S., Badler, J. B., Badler, N. I., Gleicher, M., Mutlu, B., and McDonnell, R. (2015). A review of eye gaze in virtual agents, social robotics and hci: Behaviour generation, user interaction and perception. In *Computer graphics forum*, volume 34, pages 299–326. Wiley Online Library.

Salvucci, D. D. and Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the 2000 symposium on Eye tracking research & applications*, pages 71–78.

Schmelter, T., Hernadi, L., Störmer, M. A., Steinicke, F., and Hildebrand, K. (2021). Interaction based redirected walking. *Proceedings of the ACM on Computer Graphics and Interactive Techniques*, 4(1):1–16.

Schöne, B., Sylvester, R. S., Radtke, E. L., and Gruber, T. (2021). Sustained inattentional blindness in virtual reality and under conventional laboratory conditions. *Virtual Reality*, 25(1):209–216.

Schumacher, J., Allison, R. S., and Herpers, R. (2004). Using saccadic suppression to hide graphic updates. In *EGVE*, pages 17–24.

Schweitzer, R. and Rolfs, M. (2020). An adaptive algorithm for fast and reliable online saccade detection. *Behavior research methods*, 52(3):1122–1139.

Shams, L. and Kim, R. (2010). Crossmodal influences on visual perception. *Physics of life reviews*, 7(3):269–284.

Sidenmark, L. and Gellersen, H. (2019). Eye&head: Synergetic eye and head movement for gaze pointing and selection. In *Proceedings of the 32nd annual ACM symposium on user interface software and technology*, pages 1161–1174.

Simons, D. J. and Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *perception*, 28(9):1059–1074.

Simons, D. J. and Jensen, M. S. (2009). The effects of individual differences and task difficulty on inattentional blindness. *Psychonomic Bulletin & Review*, 16(2):398–403.

Simpson, W. A. (1988). The method of constant stimuli is efficient. *Perception & psychophysics*, 44(5):433–436.

Singh, A. K., Gramann, K., Chen, H.-T., and Lin, C.-T. (2021). The impact of hand movement velocity on cognitive conflict processing in a 3d object selection task in virtual reality. *NeuroImage*, 226:117578.

Sipatchin, A., Wahl, S., and Rifai, K. (2020). Eye-tracking for low vision with virtual reality (vr): testing status quo usability of the htc vive pro eye. *bioRxiv*.

Sitzmann, V., Serrano, A., Pavel, A., Agrawala, M., Gutierrez, D., Masia, B., and Wetzstein, G. (2018). Saliency in vr: How do people explore virtual environments? *IEEE transactions on visualization and computer graphics*, 24(4):1633–1642.

Špakov, O. and Majaranta, P. (2012). Enhanced gaze interaction using simple head gestures. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 705–710.

Stark, L., Kong, R., Schwartz, S., Hendry, D., and Bridgeman, B. (1976). Saccadic suppression of image displacement. *Vision Research*.

Stein, N., Niehorster, D. C., Watson, T., Steinicke, F., Rifai, K., Wahl, S., and Lappe, M. (2021). A comparison of eye tracking latencies among several commercial head-mounted displays. *i-Perception*, 12(1):2041669520983338.

Steinicke, F., Bruder, G., Jerald, J., Frenz, H., and Lappe, M. (2009). Estimation of detection thresholds for redirected walking techniques. *IEEE transactions on visualization and computer graphics*, 16(1):17–27.

Stewart, E. E., Hübner, C., and Schütz, A. C. (2020). Stronger saccadic suppression of displacement and blanking effect in children. *Journal of Vision*, 20(10):13–13.

Strandvall, T. (2009). Eye tracking in human-computer interaction and usability research. In *IFIP Conference on Human-Computer Interaction*, pages 936–937. Springer.

Sun, Q. (2021). Leveraging human visual perception for an optimized virtual reality experience. *IEEE Computer Graphics and Applications*, 41(6):164–170.

Sun, Q., Patney, A., and Steinicke, F. (2020). Redirected walking in vr. In *Real VR–Immersive Digital Reality*, pages 285–292. Springer.

Sun, Q., Patney, A., Wei, L.-Y., Shapira, O., Lu, J., Asente, P., Zhu, S., McGuire, M., Luebke, D., and Kaufman, A. (2018). Towards virtual reality infinite walking: dynamic saccadic redirection. *ACM Transactions on Graphics (TOG)*, 37(4):1–13.

Tanriverdi, V. and Jacob, R. J. (2000). Interacting with eye movements in virtual environments. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 265–272.

Tchalenko, J. (2007). Eye movements in drawing simple lines. *Perception*, 36(8):1152–1167.

Tomlinson, R. and Bahra, P. (1986). Combined eye-head gaze shifts in the primate. ii. interactions between saccades and the vestibuloocular reflex. *Journal of neurophysiology*, 56(6):1558–1570.

Triesch, J., Sullivan, B. T., Hayhoe, M. M., and Ballard, D. H. (2002). Saccade contingent updating in virtual reality. In *Proceedings of the 2002 symposium on eye tracking research & applications*, pages 95–102.

Unema, P. J., Pannasch, S., Joos, M., and Velichkovsky, B. M. (2005). Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual cognition*, 12(3):473–494.

Vasquez-Caballero, C. V. (2020). inattentional blindness: A virtual reality experiment.

Vater, C., Kredel, R., and Hossner, E.-J. (2017). Detecting target changes in multiple object tracking with peripheral vision: More pronounced eccentricity effects for changes in form than in motion. *Journal of Experimental Psychology: Human Perception and Performance*, 43(5):903.

Vercher, J.-L., Magenes, G., Prablanc, C., and Gauthier, G. (1994). Eye-head-hand coordination in pointing at visual targets: spatial and temporal analysis. *Experimental brain research*, 99(3):507–523.

Vinnikov, M. and Allison, R. S. (2014). Gaze-contingent depth of field in realistic scenes: The user experience. In *Proceedings of the symposium on eye tracking research and applications*, pages 119–126.

Vinnikov, M., Allison, R. S., and Fernandes, S. (2017). Gaze-contingent auditory displays for improved spatial attention in virtual reality. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 24(3):1–38.

Volkmann, F. C., Riggs, L. A., and Moore, R. K. (1980). Eyeblinks and visual suppression. *Science*, 207(4433):900–902.

Volkmann, F. C., Riggs, L. A., White, K. D., and Moore, R. K. (1978). Contrast sensitivity during saccadic eye movements. *Vision research*, 18(9):1193–1199.

Volokitin, A., Gygli, M., and Boix, X. (2016). Predicting when saliency maps are accurate and eye fixations consistent. In *Proceedings of the ieee conference on computer vision and pattern recognition*, pages 544–552.

Wagner, M., Baird, J. C., and Barbaresi, W. (1981). The locus of environmental attention. *Journal of Environmental Psychology*, 1(3):195–206.

Walker, R., Walker, D. G., Husain, M., and Kennard, C. (2000). Control of voluntary and reflexive saccades. *Experimental Brain Research*, 130(4):540–544.

Wang, Y., Wu, Y., Chen, C., Wu, B., Ma, S., Wang, D., Li, H., and Yang, Z. (2021). Inattentional blindness in augmented reality head-up display-assisted driving. *International Journal of Human–Computer Interaction*, pages 1–14.

Watson, A. B. and Pelli, D. G. (1983). Quest: A bayesian adaptive psychometric method. *Perception & psychophysics*, 33(2):113–120.

Wei, L., Okuyama, F., and Sakamoto, Y. (2020). Fast calculation method with saccade suppression for a computer-generated hologram based on fresnel zone plate limitation. *Optics express*, 28(9):13368–13383.

Wexler, M. and Collins, T. (2014). Orthogonal steps relieve saccadic suppression. *Journal of vision*, 14(2):13–13.

Whiteside, T. C., Graybiel, A., and Niven, J. I. (1963). Visual illusions of movement.

Technical report, NAVAL SCHOOL OF AVIATION MEDICINE PENSACOLA FL.

Witmer, B. G. and Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence*, 7(3):225–240.

Wu, T. L., Gomes, A., Fernandes, K., and Wang, D. (2019). The effect of head tracking on the degree of presence in virtual reality. *International Journal of Human–Computer Interaction*, 35(17):1569–1577.

Yang, Q., Bucci, M. P., and Kapoula, Z. (2002). The latency of saccades, vergence, and combined eye movements in children and in adults. *Investigative Ophthalmology & Visual Science*, 43(9):2939–2949.

Yarbus, A. L. (1967). Eye movements during perception of complex objects. In *Eye movements and vision*, pages 171–211. Springer.

Young, L. R. and Sheena, D. (1975). Survey of eye movement recording methods. *Behavior research methods & instrumentation*, 7(5):397–429.

Zhang, M., Ma, K. T., Lim, J. H., Zhao, Q., and Feng, J. (2018). Anticipating where people will look using adversarial networks. *IEEE transactions on pattern analysis and machine intelligence*, 41(8):1783–1796.

Zhdanov, A. D., Zhdanov, D. D., Bogdanov, N. N., Potemin, I. S., Galaktionov, V. A., and Sorokin, M. I. (2019). Discomfort of visual perception in virtual and mixed reality systems. *Programming and Computer Software*, 45(4):147–155.

Ziat, M., Hayward, V., Chapman, C. E., Ernst, M. O., and Lenay, C. (2010). Tactile suppression of displacement. *Experimental brain research*, 206(3):299–310.

Zietz, D. and Hollands, M. (2009). Gaze behavior of young and older adults during stair walking. *Journal of motor behavior*, 41(4):357–366.

Zuber, B. and Stark, L. (1966). Saccadic suppression: elevation of visual threshold associated with saccadic eye movements. *Experimental neurology*, 16(1):65–79.