Adaptive Rendering for Improving Locomotive Experiences in Virtual and Augmented Realities

January 2023

Guanghan Zhao

Adaptive Rendering for Improving Locomotive Experiences in Virtual and Augmented Realities

Submitted to
Graduate School of Information Science and Technology
Osaka University

January 2023

Guanghan Zhao

Thesis Committee:

Prof. Haruo Takemura (Osaka University)
Prof. Tatsuhiro Tsuchiya (Osaka University)
Prof. Noriyuki Miura (Osaka University)
Assoc. Prof. Jason Edward Orlosky (Osaka University and Augusta University)

List of Publications

Journals

1. G. Zhao, J. Orlosky, S. Feiner, P. Ratsamee, and Y. Uranishi. Mitigation of VR Sickness during Locomotion with a Motion-Based Dynamic Vision Modulator. *IEEE Transactions on Visualization and Computer Graphics.*, Vol. 89, 1–10, 2022.

International Conferences

Peer-reviewed Posters

1. G. Zhao, J. Orlosky, and Y. Uranishi. Evaluating Presence in VR with Self-Representing Auditory-Vibrotactile Input. *IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 577-578, Mar. 2021.

Abstract

Motion is a fundamental part of interaction in virtual reality (VR) and Augmented reality (AR). Challenging situations occur more often to the users while they are moving virtually or physically in the environment since the view or body is no longer stable. In addition, the motions cause distraction, sickness and loss of self-orientation. Thus as the user starts to move virtually or physically, ordinary approaches of interactions and rendering may not be reliable compared with stable scenarios. To date, several methods have been proposed to support users to gain instant and accessible information and visual effects while moving such as real-time visual guidance, field of view (FoV) modifications, auto-placed interfaces and hands-free interactions. In some practical use cases, however, those approaches were not well designed to counterbalance interactions and user experiences. Moreover, most current methods only focus on motive parameters such as speed, and further exploration on the application of interactive sensors and environment-adaptive rendering is still barely touched. To tackle these issues, this work focuses on exploring innovative methods for interactions in locomotive scenarios with motion tracking, eye tracking, environment sensing and real-time rendering technologies. The focus can be distinguished into: a) virtual passive motions (locomotion) in VR, and b) physical active motions in AR.

In virtual reality, VR sickness resulting from continuous locomotion via controllers or joysticks is still a significant problem. In this work, I present a set of algorithms to mitigate VR sickness that dynamically modulate the user's field of view by modifying the contrast of the periphery based on movement, color, and depth. In contrast with previous work, this vision modulator is a shader that is triggered by specific motions known to cause VR sickness, such as acceleration, strafing, and linear velocity. Moreover, the algorithm is governed by delta velocity, delta angle, and average color of the view. I conducted two experiments with different washout periods to investigate the effectiveness of dynamic modulation on the symptoms of VR sickness, in which this approach is compared against baseline and pitch-black field of view restrictors. The first experiment made use of a just-noticeable-sickness design, which can be useful for building experiments with a short washout period. In the second experiment the methods were tested with a fashionable long washout experimental design for further comparison.

For physical active motions, this work focused on applying AR techniques to cycling interfaces. During cycling activities, cyclists often monitor a variety of information such as heart rate, distance, and navigation using a bike-mounted phone or cyclocomputer. In many cases, cyclists also ride on side-

walks or paths that contain pedestrians and other obstructions such as potholes, so monitoring information on a bike-mounted interface can slow the cyclist down or cause accidents and injury. In this work, I present HazARdSnap, an augmented reality-based information delivery approach that improves the ease of access to cycling information and at the same time preserves the user's awareness of hazards. To do so, we implemented real-time outdoor hazard detection using a combination of computer vision, motion, and position data from a head mounted display (HMD). We then developed an algorithm that snaps information to detected hazards when they are also viewed so that users can simultaneously view both rendered virtual cycling information and the real-world cues such as depth, position, time to hazard, and speed that are needed to assess and avoid hazards. Results from a study with 24 participants that made use of real-world cycling and mixed reality (MR) hazards showed that both HazARdSnap and forward-fixed augmented reality user interfaces (UIs) can effectively help cyclists access virtual information without having to look down, which resulted in fewer collisions (51% and 43% reduced compared to baseline, respectively) with virtual hazards.

Additionally, cyclists often focus on pedestrians, vehicles or road conditions in front of their bicycle. Thus, approaching vehicles from behind can easily be missed, which can result in accidents, injury, or death. Although rear information can be accessed by applying rear-view mirrors or monitors, having to look down at this kind of small interface distracts the cyclist from other hazards. To help address this problem, a peripheral information delivery approach that enhances awareness of rear-approaching vehicles and at the same time preserves forward vision, which was named ReAR Indicator, is presented. ReAR uses computer vision applied to a rear-facing RGB-D camera with position data from an HMD for real-time vehicle detection. Then, an algorithm that delivers information to the periphery such that the user can simultaneously view forward information but still use cues that provide information about hazard distance, width, and probability of collision was developed. Results from a VR-based experiment with 20 participants showed that the ReAR Indicator can effectively help cyclists maintain focus on their forward view while still avoiding collisions with virtual rear vehicles.

The findings of this work provide implications and insights on the implementation of locomotive interactions, the design of locomotion-related experiments in VR and AR, interfaces that adapt to environments, and applications of motion-based real-time rendering for HMDs. The applied combination of real-time rendering techniques with motion tracking can lead to great improvement of human-HMD interaction in both VR and AR.

Acknowledgments

The writing of this dissertation is an incredible journey and a milestone in my life. I could not have accomplished this endeavor without the tones of support and help from my advisors, friends and family. I would like to give my sincere regards to all of those who supported me during the completion of this work.

This work is done under the supervision of Prof. Haruo Takemura of the Graduate School of Information Science and Technology at Osaka University. I would like to thank him for providing me with continued support and the ideal environment for accomplishing the entire work. I also want to thank the thesis committee for offering me great comments to improve this dissertation.

I want to especially express my sincere gratitude to my supervisor, Assoc. Prof. Jason Orlosky at Augusta University, for his inspiration, assistance and dedication to this work. Throughout my research, he provided me with inspiring ideas, sound opinions, academic experience and encouragement.

I am also grateful to Assoc. Prof. Yuki Uranishi at Osaka University, Lec. Photchara Ratsamee at Osaka Institute of Technology, and Prof. Kiyoshi Koyokawa at Nara Institute of Science and Technology for continuously offering me advice and help during the period of me as a Ph.D student, which made this current work possible.

I also appreciate the opportunity to be able to study and work with a lot of people at Osaka University. I want to thank all the members of Takemura Laboratory for their support. I would like to especially mention and thank Chang Liu, who offered me an overview of the research field the first time I stepped into the world of academia. I also want to thank my schoolmates Tao Tao and Harn Sison for their guidance and help on both research and daily life since I first came to the lab. Furthermore, I express my gratitude to the staff at Cybermedia Center and the Graduate School of Information Science and Technology for their help with various official procedures and paperwork.

Lastly and most importantly, I give my sincere thanks to my parents and grandmother, Min Zhao, Xianghong Chen, and Biyun Liu, for their understanding and sustained support in my whole life. I could not have had this amazing opportunity to study as a Ph.D student and pursue my own interests without their selfless effort. I dedicate this dissertation to them.

The work in this thesis was supported in part by the Office of Naval Research Global grant N62909-18-1-2036 and the Japan Society for the Promotion of Science grant 21H03482.

Guanghan Zhao
Osaka University
January 2023

Contents

List of Tables

List of Figures

List of Acronyms

ANOVA	Analysis of Variance
API	Application Program Interface
AR	Augmented Reality
DNN	Deep Neural Network
FoV	Field of View
GLMMPQL	Generalized Linear Mixed Model
HCI	Human-Computer Interaction
HLSL	High Level Shader Language
HMD	Head-Mounted Display
HUD	Head-Up Display
GLSL	OpenGL Shading Language
MR	Mixed Reality
OST-HMD	Optical See-Through Head-Mounted Display
PC	Personal Computer
QTE	Quick Time Event
ROI	Regions of Interest
SAM	Self-Assessment Manikin
SSQ	Simulator Sickness Questionnaire
UI	User Interface
VE	Virtual Environment
VR	Virtual Reality

Chapter 1
Introduction

To date, virtual reality and augmented reality technologies have been studied by researchers around the world for decades. Owe to their efforts, HMDs of cross realities have been improved and able to handle complicated tasks and provide desirable performances for users. In addition, with the powerful HMDs, more and more algorithms and systems are developed to improve user experience as well as deliver convenient solutions for multiple purposes.

Current VR contents are meanly welcomed by consumers as games, stereo movies, and training simulators since VR technology allows users to gain immersion by realistic looking and moving in the virtual world in stead of traditional monitor-controller based interactions. However, there exists an huge obstacle that hampers VR to spread. As users are moving passively in a virtual environment, their eyes will gain the felling of vection while their body feel stable. This situation triggers visual-vestibular conflict and causes simulation sickness (VR sickness) symptoms. Therefore, novel approaches for mitigating VR sickness during locomotion is needed.

Unlike VR, AR technologies are designed to help users solving real-world problems instead of amusement. Approaches in fields of remote working, E-learning, and object detecting are well studied. However, AR related methods for moving situations are still confined to driving, which usually happens on a flat and clear road. Thus approaches for more complicated environments and moving conditions besides driving is barely touched.

In both VR and AR, interactive developments for moving situations are harder than stable situations since the camera is not stable and the environment keeps changing. In most of the moving situations, excessive user effort can be caused by perceiving and/or interacting with the environment. As such, this work contributes to the improvement of user experience of both VR and AR while moving and interacting in virtual and real world by developing real-time rendering methods that are compatible for modern VR and AR devices.

1.1 Background

This section gives a brief introduction on the applied HMDs, interactive technologies, as well as appropriate graphical approaches as solutions, followed up with definitions of the problems that I have found regarding locomotive interactions. In the end of this section, I discuss the contributions of my studies and provide an overview of my dissertation.

1.1.1 Head-mounted Displays

According to the differences of visual systems, HMDs are usually be categorized as immersive HMDs and optical see-through HMDs (OST-HMDs). In particular, immersive HMDs cover the user's whole view and provide immersive virtual visions. Moreover, the user can view a virtual scene through the HMD's lenses and interact with virtual objects via controllers. On the other hand, OST-HMDs often use optical combiners to project information, interfaces, or virtual objects on the user's view of the real world. In addition, with OST-HMDs, the user is able to perceive the real world with augmented vision as well as interact with additional interfaces delivered by the HMDs.

These categories of HMDs were developed for different proposes, namely for VR and AR. In this work, I take advantage of the features of different HMDs to solve the identified problems: For VR sickness, HTC Vive Pro Eye HMD (Fig. 1.1-a) was used since it has both high resolution (1440 * 1600) and FoV (110 degrees). Furthermore, the HMD offers a high speed cable that allows me to process my custom rendering algorithms with a powerful graphic card for usages on personal computers (PCs). For cycling AR, Microsoft HoloLens 2 HMD (Fig. 1.1-c) was used since it is the most powerful mobile HMD for AR usages. In particular, HoloLens 2 has an eye tracker built in and it allows developers to place 3D contents in the view by hologram technology. However, similar to other AR HMDs/glasses, the FoV of HoloLens 2 is relatively small (43 degrees) and restricts the implementation of visual rendering. Thus in this work, I discuss my ideas about how to overcome the FoV restriction of hardware. Besides, considering safety of user tests, I applied Oculus Quest 2 (Fig. 1.1-b) to evaluate one of my method regarding cycling. Since the HMD is lite and wireless, it can significantly reduce the carry load when participants riding on a bicycle during user tests.

Figure 1.1: Images of the immersive HMDs and OST-HMD applied in this work. (a) The HTC Vive Pro Eye and (b) Oculus Quest 2, which provide the user with a fully-immersive VR experience and hence prevent the user from directly viewing the real world. (c) A user wearing the Microsoft HoloLens 2, which renders digital content directly in the user's field of view overlaid onto the real world.

1.1.2 Interactive Technologies

Controllers. Controller based interaction is the major approach of interacting with the virtual objects and scenes for the immersive HMDs. In addition, by using the joysticks, touchpads, or gyroscopes built in the controllers, the user would able to move or teleport in a virtual environment. This is the most common approach for locomotion in VR.

Head Tracking. Head tracking is an indispensable for all of the immersive HMDs and some of the OST-HMDs. It allows the system to tack the position of the user's head which refers to the body position as well as head facing direction. There are two different popular implementations of head tracking: outside-in and inside-out. For HTC Vive Pro Eye HMDs, the outside-in method is implemented by a pair of laser casting base stations. The HMD receives lasers and calculates positions base on the distance and angle of those lasers. For wireless lite HMDs like Oculus Quest 2 and Microsoft HoloLens 2, they use the inside-out method which casts infrared rays to the surroundings from the HMD and calculates the intensity of infrared reflection. With these head tracking technologies, I was able to collect real-time data of head rotation, facing direction and position during locomotion for adaptive rendering.

Eye Tracking. As one of the most salient features on human face, the action of eyes are very active and noticeable. Therefore, When users are tend to focus on an object or area, it can be easily detected by sensors. On the other hand, active gaze movements can be used as interaction methods such as

pointing (Ware and Mikaelian (1986)) or moving objects (Liu et al. (2020a)). In this work, I take advantage of eye tracking technologies in order to help understanding the user's focus and then render interfaces simultaneously.

1.1.3 Graphical Approaches

Computer Vision. Since the 1970s, the field of computer vision keeps drawing the attention of researchers. Studies about "mimic the human visual system by computers" and "describe what it saw" have been done and pushed the technologies in the field to evolve. To date, powerful neural network models such as YOLO (Redmon et al. (2016)) allow programmers to detect multiple numbers and kinds of objects in real-time with a common webcam. Moreover, these object detecting models also encouraged the development of adaptive rendering approaches for real-world interactive AR systems.

Shaders. A shader is a program that calculates the appropriate levels of light, brightness, contrast and color during the rendering of a 3D scene and it has evolved to perform special effects on cameras and textures. In addition, shaders are supported by major graphics software libraries such as OpenGL (OpenGL Shading Language, GLSL) and Direct3D (High Level Shader Language, HLSL). The most frequently used types of shaders are pixel shaders and vertex shaders. Pixel shaders are also known as fragment shaders, they can not only compute color and other attributes of each pixel but also manipulate the attributes. Regarding 3D shaders, they act on 3D models or other geometry and also access the colors and textures used to draw the model or mesh. In this work, shading algorithms are applied to compute the insight pixels' attributes and then manipulate them for adaptive rendering.

Peripheral Rendering. Peripheral vision the vision that occurs outside the focal area, i.e. away from the center of gaze. Peripheral vision is relatively weak at distinguishing detail, color, and shape as compared with foveal vision. Thus rendering on the peripheral area is less distracting and draw less attention. Based on this feature of peripheral vision, inconspicuous visual modulators and displays were proposed in the fields of VR and AR.

1.1.4 Problem Definition

VR sickness. Virtual reality sickness is also known as cybersickness or motion sickness, which usually induced by passive camera movements in virtual environments. It leads to sickness symptoms such as nausea, dizziness and

vomiting that prevent users to stay longer time in the virtual environments. According to the most widely accepted sensory conflict theory proposed by Reason et. al. (Reason and Brand (1975), Reason (1978)), motion sickness is triggered when self-motion perceived from the view within an HMD is mismatched with the motion perceived by the human vestibular system. Therefore there are two perspectives to solve the VR sickness caused by long distance movements: let the user to physically walk around or modify the vision. Since current technologies are still not able to allow users walk infinitely without a reduce of immersion, developers are relying on controller based locomotion methods for moving in virtual environments and apply visual modifications to mitigate VR sickness. To render appropriate visual modifications that able to counterbalance immersion and sickness, blur effects and black FoV restrictors are frequently used.

Blur effects can ideally preserve immersion in VR, as they did not fully block the optical flows in sight. However, that also means that the user will precept more self-motion and reduce the effectiveness of sickness mitigation. On the other hand, black FoV restrictors are able to block optical flows. Nonetheless, the effectiveness of these restrictors can be easily affected by the HMD's FoV. Besides, although these methods are designed to be correlated with speed only and restricted to the 2D plane of the screen space.

To better counterbalance the conflict between immersion and sickness and take different kinds of movements in 3D space into consideration, this work proposes a motion-based dynamic rendering approach.

Difficulties with information presentation during cycling and mobile activity. As one of the most popular transportation form around the world, comparing with driving, cycling did not get enough attention from researchers. Moreover, cyclists must often ride on sidewalks or traffic lanes without user interfaces that can be easily accessed. Currently, cyclists have to use bike-mounted smartphones or cyclocomputers for checking essential information such as navigation, speed, time, or rearview. Looking down frequently at these interfaces causes distractions potentially leads to accidents (De Waard et al. (2014), Terzano (2013)). To address this problem as well as provide safer interactive interfaces, recent research focused on auditory inputs. As a distinctive feature, auditory inputs do not cause any visual distraction or obstruction. However, the audios can be mixed up with ambient sounds and there is a high potential of danger when they cover the sounds of approaching cars. In addition, although bone conducting headphones can be applied to prevent the block of hearing, the audios delivered by them still mix with

ambient sounds in the user's brain. Besides, vibrotactile inputs are considered as a alternative solution. However, the information load that vibrotactile inputs can deliver is relatively low and vibration may affect the control of bicycle handles.

Considering the drawbacks of auditory and vibrotactile inputs, I believe that applying visual interfaces can be the most appropriate way to help cyclists to access information while cycling. Furthermore, the distraction caused by visual interfaces is a huge barrier on this way. Owe to modern powerful HMDs, I would take advantage of the eye trackers and high-resolution graph cards within. Therefore, in this work, I propose different systems for safe cycling information access with eye tracking and peripheral displaying techniques.

1.2 Contributions

This work proposes adaptive methods to human-HMD interaction for improving the user experience and safety during locomotion. In particular, adaptive rendering algorithms are subsequently composed of three subordinate aspects: 1) to mitigate VR sickness when experiencing locomotion in virtual environments, 2) to provide safe and easy-access information delivery while cycling, and 3) to deliver information of backward hazards while moving forward.

Motion-Based Dynamic Vision Modulator. In VR, VR sickness is one of the most concerned barrier resulting from continuous locomotion and has been studied for decades. Though several well-known visual modification methods have been proposed, there is still no best solution for perfectly preserve immersion while preventing VR sickness to be triggered. This work introduces a novel rendering of vision modulator that dynamically modulate the user's vision with a low-contrast screen effect based on movement, color, and depth during locomotion. In this approach, an algorithm that calculates pixels' angular velocity and changes intensity in response to both the user's virtual motions and the depth of pixels in world space was developed. Moreover, I have designed a separate algorithm for sampling the pixels on the screen on a frame-by-frame basis as the color of the low-contrast effect instead of the popular pitch-black color.

In order to evaluate the Motion-Based Dynamic Vision Modulator, I have conducted two experiments with a novel just-noticeable-sickness design and a traditional long washout period design. The goals for these experiments are: 1) testing the modulator against a well-known FoV-modulation technique and a baseline condition and 2) investigate the effectiveness of the three conditions

with two different designs of the washout period. Results showed that both the FoV-modulation technique and the proposed approach effectively mitigated visually induced VR sickness and enabled users to remain in VEs for a longer period of time than the baseline without a significant awareness of the modulator. In addition, the results of the three conditions showed some differences in the two different designs of experiments, which encourages future studies about the effects of washout periods. In summary, the contributions of this work are:

- The design, testing, and refinement of a rendering technique (vision modulator) designed to reduce VR sickness and without becoming distracting by reducing optical flow, engaging only during potentially sickness-causing motions, and preserving scene coloration.
- Two experiments with different washout periods testing the vision modulator against a well-known FoV-modulation technique and a baseline condition, as well as analysis and discussion of the results.

Gaze-based Augmented Information Delivery. Cyclists must often ride on sidewalks, narrow or rugged paths, or situations in which obstacles can be crash hazards. To stay safe, riders must pay attention to what is in front of them without looking down and reduce their speed to avoid collisions when necessary. Meanwhile, cyclists also have the need of monitoring a variety of information such as heart rate, distance, and navigation using a bike-mounted phone or cyclocomputer. To address this problem, this work introduces a novel approach of rendering user interfaces, referred to as HazARdSnap, that improves the ease of access to cycling information and at the same time preserves the userâs awareness of hazards. For this approach, a subsystem of hazard detection using a combination of computer vision, motion and position data was firstly implemented. Then, I developed another subsystem that renders (snaps) information to the detected hazards when they are also viewed so that users can simultaneously view both rendered virtual cycling information and the real-world cues.

Regarding the evaluation of HazARdSnap, I designed and conducted a novel MR based user test. During the tests, users were asked to ride a real bicycle on a real ground and pass virtual obstacles that delivered by an HMD with tasks. As compared to fashionable VR based tests that using stationary bicycles and full-virtual scenes, this MR based design is able to collect more data from real cycling behaviours such as steering and balancing while preserving safety. The HazARdSnap was compared with a forward-fixed AR

interface and a bike-mounted smartphone in the user test. As results, the participants encountered fewer collisions with the HazARdSnap (51%) and the forward-fixed AR interface (43%) as compared to the bike-mounted smartphone. Moreover, both of the AR methods induced less distractions than the bike-mounted smartphone method. In summary, the contributions of this work are:

- The design, testing, and refinement of a interface rendering technique, HazARdSnap, designed to decrease the chance of collisions when accessing digital information during cycling.
- A bike-mounted system that makes use of computer vision, depth information, and environment tracking to detect oncoming hazards such as potholes and pedestrians.
- An MR based user study that compares cyclist performance while using HazARdSnap, a forward-fixed AR interface, and a bike-mounted smartphone as well as analysis and discussion of metrics like collisions with virtual objects, time spent viewing the road, and subjective ratings.

Peripheral Indicators for Rear-Approaching Vehicles. This study is mainly dedicated to enhance awareness of rear-approaching vehicles and at the same time preserves forward vision while cycling. To do so, a peripheral information delivery approach named ReAR Indicator was developed. For this approach, I firstly applied computer vision to a rear-facing RGB-D camera and took position data from a head mounted display (HMD) into consideration for real-time vehicle detection. Secondly, a customized algorithm that renders information to the periphery such that the user can simultaneously view forward information but still use cues that provide information about rear-approaching vehicles was programmed.

Considering that the main purpose of this study is enhancing awareness of rear-approaching vehicles while moving forwardly, a VR based user study was conducted that pays more attention on the participants' reaction to both the forward and backward. The user study used a within-subjects design that tested the ReAR Indicator against two other conditions, a front-placed virtual monitor and a 3D arrow interface. Analyses of the data collected from the user study revealed that participants with the ReAR Indicator and the front-placed virtual monitor were more likely to encounter fewer collisions from behind and react to the rear-approaching vehicles faster. In addition, the ReAR Indicator showed similar effectiveness as looking at the rearview

delivered by the front-placed virtual monitor. The primary contributions of this work include:

- This work proposes a novel peripheral rendering which enhances awareness of rear-approaching vehicles while moving to and focusing on the forward direction.
- The experimental results reveal that that the ReAR Indicator can effectively help cyclists maintain focus on their forward view while still avoiding collisions with virtual vehicles approaching from the rear.

1.3 Dissertation Overview

The remainder of this dissertation is composed as follows:

In Chapter 2, a survey including previously conducted investigations and proposed approaches relating to locomotion and rendering in the fields of VR and AR is provided. Reviews and discussions about these prior works are also stated for better understanding the necessity and challenges as well as help judging this dissertation.

Chapter 3 introduces a novel rendering method for achieving a better counterbalance of immersion and sickness during locomotive activities in VR. The method is implemented by rendering a low-contrast effect that adapted to depths and angular velocities of pixels in sight. Moreover, a customized algorithm is provided to not only efficiently but also accurately sampling the average color of a large amount of pixels. A comprehensive set of experiments was conducted to verify the efficiency of this method in comparison with a black FoV restrictor condition and a baseline condition. A summary discusses the results and findings from the experiments.

Chapter 4 introduces the design and implementation of a novel augmented interface rendering approach for cycling. The approach is composed by two subsystems: a hazard detecting subsystem and a visual rendering subsystem. An MR based user study was conducted in order to evaluate the effectiveness of this approach on safety and the ease of access. A summary then discusses the experimental results along with the implications.

Chapter 5 introduces a novel method for enhancing the awareness of rear-approaching vehicles while moving and looking forward with peripheral rendering techniques. To implement this method, a rear-fixed RGB-D camera is used to detect vehicles and a peripheral indicator is rendered via an HMD. A user experiment in different virtual scenes is conducted to confirm the effec-

tiveness and safety of the peripheral indicator. In the end of the chapter, I discuss the experimental results and demonstrations as a summary.

Chapter 6 introduces an auditory-vibrotactile rendering method which is beyond the visual domain but related to user experiences in VR. This method take advantage of bone conduction headphones and a chest-mounted vibration speaker to deliver self-representing sounds such as heartbeat and snoring. In the experiment, the method was evaluated with other sound delivering inputs in different virtual scenes. Moreover, experimental results along with my findings is provided.

Chapter 7 summarizes and concludes this dissertation with a detailed overview of the contributions and findings of this work, and discusses the remaining challenges regarding the future design and development of real-time adaptive rendering methods for locomotion in VR and AR.

Chapter 2

Related Work

This chapter gives an introduction on related studies in the domain of real-time rendering and interactive information delivery during locomotion, as well as the further motivation for this work. In addition, since highly related studies are also introduced in the chapter of each specific work, this chapter mainly summarizes related studies in a broad sense.

2.1 Real-Time Rendering

Real-time rendering methods have been studied and proposed, especially for different proposes related to locomotive interactions such as simulation and navigation.

Simulation for Safety and Immersion. Regarding simulation for safety and immersion, Hillaire et al. (2007) proposed a rendering of blur effects that adaptively changes it's density and rendering areas basing on depth-of-field for first-person navigations in virtual environments. The results of their experiments indicated that the visual blur effects has no negative effects on gaming performance and is preferred by users. Langbehn et al. (2016) investigated the effect of rendering a visual blur on the estimation of speed and distance during locomotion and demonstrated that blur effects will not affect human perception of speed and distance. Yamashita et al. (2021) proposed an approach that deforms the rendering area according to the vection illusion which induces a sense of immersion only with images in order to enhance the sense of self-motion along with immersion. Their experiment with a driving simulation environment revealed that the deform and transform of rendering areas enhances the sense of immersion but will cause VR sickness.

Moreover, real-time rendering techniques are considered as a effective solution for mitigating VR sickness and have been well-studied. For example, Nie et al. (2019) studied the effects of a saliency detection-based rendering on the reduction of VR sickness when the display on a user's retina is dynamic blurred (Fig. 2.1). Their experimental results demonstrates that the rendering approach may alleviate VR sickness during locomotion in VR and enable

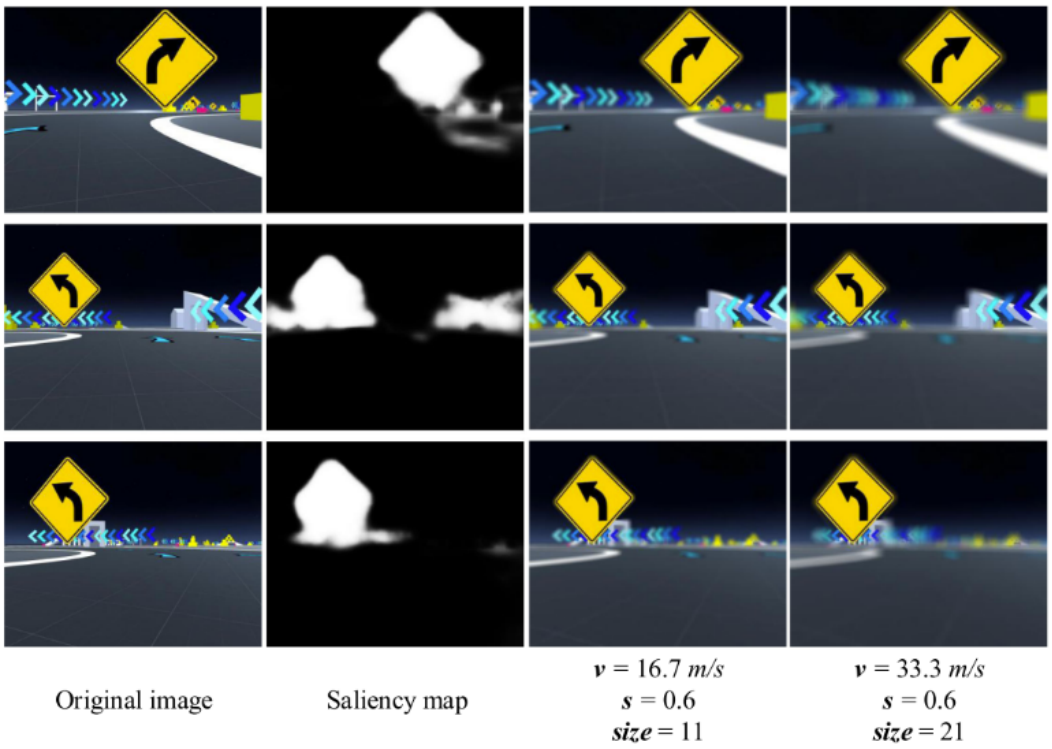

Figure 2.1: Images showing the saliency detection-based rendering method by Nie et al. (2019). Each row corresponds to an example. 1st column: the original VE scene captured by virtual camera of SteamVR in Unity3D. 2nd column: saliency maps. 3rd-4th columns: non-salient areas blurred with 2 different kernel sizes.

users to remain in a virtual environment for a longer period of time. Inspired by Nie et al., Chen et al. (2022) came up with a new approach for rendering blur effects, named as texture blur. The experiment results from different motion tests revealed that the participants with texture blur experienced less sickness. Beside blur effects, the rendering of static and dynamic rest frames have been investigated by Cao et al. (2018). In their work, a dynamic rendering approach which allows the opacity of the rest frame changes in response to visually perceived motion as users virtually traversed the virtual environment is presented. Lim et al. (2021) developed a novel method based on dynamic FoV processing and estimates VR sickness for each video frame by defining the relationship between the motion information and the measured VR sickness.

Augmented Navigation. In 1993, Darken and Sibert (1993) firstly investigated the effectiveness of a toolset of techniques based on principles of navigation derived from real world analogs in virtual environments. They

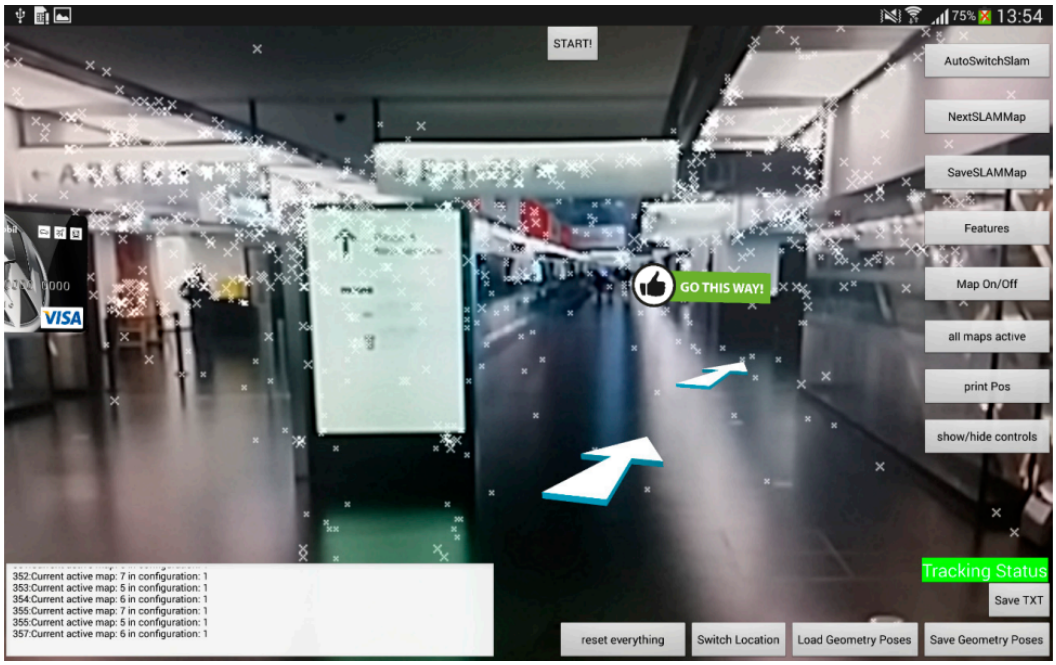

Figure 2.2: An image showing the visualization of the AR navigatior by Gerstweiler et al. (2015).

concluded that rendering navigation base on real world navigation aids such as maps can still be helpful in VR. Since then, real-time rendering methods for navigation in VR have been well-studied by researchers. For example, Fukatsu et al. (1998) proposed a manipulation technique to render and control a âbird's eyeâ overview display and the original virtual vision simultaneously. It enables efficient navigation even in enormous and complicated virtual environments using both views. Furthermore, Wu et al. (2007) evaluated three rendering approaches including a view-in-view map, an animation guide, and a human-system collaboration. Their results showed that while an overview still outperforms the animation guide and the human-system collaboration, the animation guide serves better for most people with ordinary spatial ability and people with superior spatial ability tends to perform better using the human-system collaboration.

On the other hand, the implementation of AR navigators is one of the most popular topics in the field of AR. Gerstweiler et al. (2015) implemented a hybrid tracking system specifically designed for complex indoor spaces, which runs on mobile devices like smartphones or tablets and renders location-based information visualization (Fig. 2.2). Chidsin et al. (2021) proposed a system that adopts the RGB-D camera to observe the surrounding environment in order to build a point cloud map and renders positioning and navigation in-

formation on a hybrid map. The results of a user study demonstrated that their method achieved significantly better performance as compared to a baseline condition. Considering that dangers may occur to people with low vision while moving on stairs, Zhao et al. (2019) designed visualizations of rendering highlights to stairs for different AR platforms. Moreover, Truong-Allié et al. (2021) implemented an adaptive AR guidance for wayfinding by applying user activity and motion detection data to dynamic rendering algorithms. However, their experiments did not demonstrate the effectiveness of this implementation.

2.2 Interactive Information Delivery

In VR and AR, investigating interactive information delivering methods for locomotive situations is also an important topic. More and more approaches have been proposed for walking and transportation such as riding and driving.

Walking Assistance. Interactive interfaces are ideal tools for users to gather additional information while walking in real world. To develop those interfaces, Lucero and Vetek (2014) proposed a novel application that allows a person to receive social network notifications on interactive glasses while walking on a busy street. As the result of their user study, the interface is able to help the participants focus their attention on their surroundings and supports discreet interactions. Gómez-Jordana et al. (2018) investigated the effectiveness of rendering virtual footprints in Parkinson's disease self-directed movements and proofed that the method can significantly improve gait performance in participants with Parkinson's disease. Schweigkofler et al. (2018) developed an augmented interface to contextualize tasks and instructions and provide building components information while walking through constructions. Moreover, Orlosky et al. (2013) evaluated user tendencies when placing virtual text in real street views through an HMD. They found that users have a tendency to place overlayed text in locations near the center of the viewing field, gravitating towards a point just below the horizon while walking on the street, as shown in Fig. 2.3. Further from the findings by Orlosky et al. (2013), Jia et al. (2021) proposed a semantic-aware task-specific label placement method by identifying potentially important image regions through a novel feature map. In particular, given an input image, its saliency information, semantic information and the task-specific importance prior are integrated in the map for the labeling task. Their experiments validated the significant benefits of their method over previous solutions for augmented walking on streets.

Figure 2.3: Image showing the common overlay regions revealed by the experimental results of Orlosky et al. (2013). A plus represents a user's overlay position in the viewing field.

Transportation Assistance. Similar to the walking scenarios, AR based information delivery methods are also developed for accessing information while driving but in a high-speed condition. To do this, Merenda et al. (2018) studied how driver performance and visual attention are affected when using fixed and animated AR HUD interface design approaches in driving scenarios that require top-down and bottom-up cognitive processing (Fig. 2.4). The results of their experiments demonstrated that animated design approaches can produce some driving gains but often come at the cost of response time and distance. Based on the findings by Merenda et al. (2018), Wu et al. (2022) developed an adaptive support system that detects a driverâs gaze on important objects in the traffic scene and adapts a cueing strategy in an AR interface. A following experiment revealed that the system can increase ratio of driversâ fixations on critical objects in their view without significantly increasing dwell time per object. In order to provide a safety aid in left-turn situations, Tran et al. (2013) implemented an AR based left-turn aid that displays a 3-second projected path of the oncoming vehicle in the driverâs environment. In addition, Calvi et al. (2020) tested the effectiveness of a system that warns the driver about the pedestrians when approaching zebra crossing areas. Their investigations confirmed that the AR warning can improve the safety of both drivers and pedestrians. Considering the importance of information delivered by traffic signs, Abdi and Meddeb (2018) implemented a novel AR traffic sign recognition system for improving driving safety and experience based on the Haar cascade and the bag-of-visual-words approach. Their experimental results showed that the method could reach comparable performance of the state-of-the-art approaches with less computational complexity and shorter training time, and it more strongly impacts the allocation of visual attention

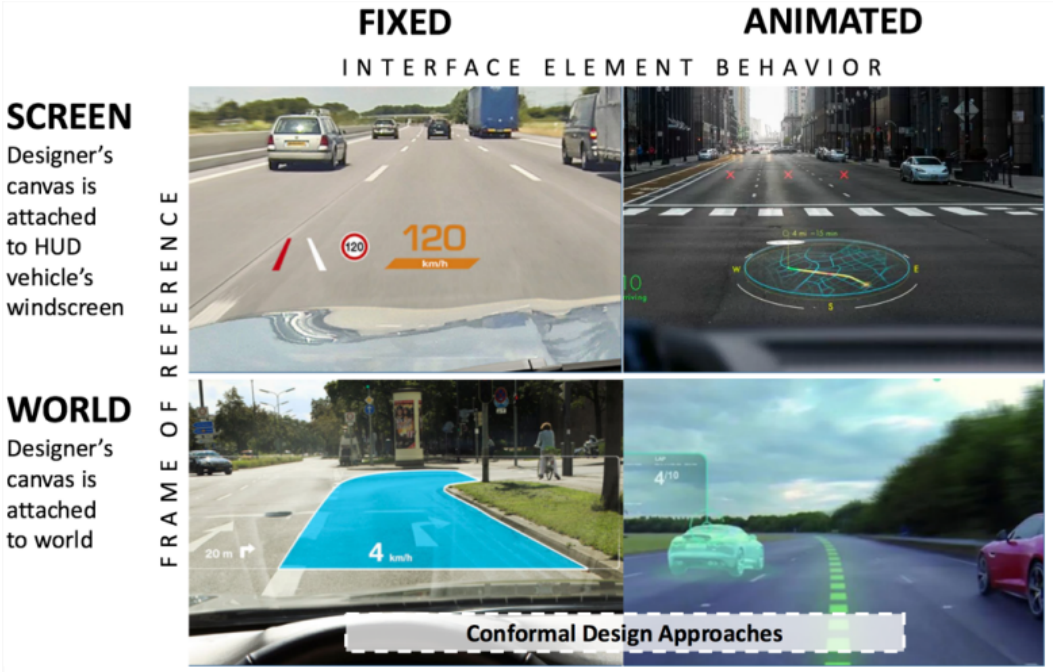

Figure 2.4: Image showing the four classes of AR HUD interface design approaches studied by Merenda et al. (2018), ranging from traditional 2D HUD graphics (e.g., speed and symbology, top-left) to highly dynamic, animated conformal graphics (e.g., virtual car metaphor for navigation, bottom-right).

during decision-making phases.

Beside driving scenarios, AR technologies are also applied to provide safe information delivery for riding: Jenkins and Young (2016) proposed a real-time alerting system which captures and integrates relevant information from an array of public databases and on-motorcycle and environmental sensors. It can fuse the gathered information and apply advanced probabilistic models and reasoning techniques to generate route-based real-time hazard alerts via an AR device. Mantell et al. (2010) developed an audio-AR system that uses sounds rather than speeches and assign them to places in order to create an audio layer placed over the regular soundscape of the city. Ginters (2019) proposed a system that displays a map for navigation and signals cyclist intention. They evaluated their system with a head-up display condition and found that cycling with the gesture projector system is easier to understand and predict the cyclist intention.

2.3 Motivation

The work of this dissertation focuses on applying real-time adaptive rendering approaches specifically to human-HMD interactions when moving in both VR and AR. The focus falls into two aspects: 1) to improve the user experience of locomotion in VR, and 2) to maximize information access and perception of hazards and thus improve the safety of using AR devices for real-world locomotive activities.

2.3.1 Improving the User Experience of Locomotion in VR

While the existing rendering approaches can mitigate VR sickness while the user is moving in virtual environments, most of them use 2D rendering on the screen as visual modifications such as pitch-black FoV restrictors and blurry FoV modifiers. Furthermore, although some of those interfaces take speed as a govern of the size and intensity, there still lacks thorough considerations on motion factors and how they interact with the visual modifications. Thus it is necessary to explore visual rendering methods that not only provide 3D-world based visual modification but also can transform adaptively regarding other motion factors. In this dissertation, Chapter 3 contributes to this topic.

2.3.2 Improving the Safety of Using AR Devices for Real-World Locomotive Activities

In locomotive use cases of AR devices, attentions on both the real-world surroundings and augmented interfaces are often required for achieving better safety. For instance, cyclists need to keep their eyes on the track and check the essential information delivered in augmented vision. As this condition lasts longer, checking information can become distracting and lead to dangers such as crash or fall off. On the other hand, studies that focus on preserving information accessibility and reduce the distraction on real-world perception for locomotion have not been well explored. Therefore, in order to address the issue and contribute to the research field, Chapter 4 and Chapter 5 present two novel approaches.

Chapter 3

Vision Modulator for Mitigating VR Sickness

3.1 Introduction

Real-time adaptive rendering methods can be applied to VR for solving many different issues and virtual reality sickness is one of them. Virtual reality sickness, also known as cybersickness or visually induced motion sickness in virtual environments (VEs), can result in symptoms such as nausea, dizziness and vomiting, which prevent users from staying in the VE for longer periods of time and can persist even after the experience ends (LaViola Jr (2000)). These unpleasant symptoms often lead to users refusing to try VR again. The most commonly accepted explanation of VR sickness is the sensory conflict theory (Reason (1978); Reason and Brand (1975)), which states that self-motion perceived from the view within an HMD is mismatched with the motion perceived by the human vestibular system. This eventually causes the human body to react strongly with discomfort.

To date, HMDs use advanced tracking systems and high refresh rates to prevent VR sickness for general use by eliminating the mismatch of self-motion between the user's visual and vestibular perception when virtual head motion is intended to correspond to tracked physical head motion. However, this does not address users who wish to use a controller or gamepad to move through the virtual world, due to a lack of physical room space or a desire to remain relatively stationary. Since the problem of locomotion in VR has not been perfectly resolved, open-space VR content has relied on various mechanisms for locomotion with controllers such as joysticks or trackpads. Frequently used techniques include teleportation, ratchet-based rotation, blurring (Budhiraja et al. (2017)), or FoV restrictors (Fernandes and Feiner (2016); Bolas et al. (2017)).

Both teleportation and ratcheting are effective for movement and diminish VR sickness by immediately transporting or transforming the user's virtual body to a target position or orientation instead of moving it continuously. This prevents the optical flow of vection in the user's view in order to avoid

sensory conflict. However, moving without optical flow can result in a lack of immersion (Bowman et al. (1997)) and may cause spatial disorientation (Bhandari et al. (2018); Bakker et al. (2003)). Much VR content, especially open-space VR games, makes this tradeoff, applying teleportation as its default locomotion method.

The FoV restrictor is another technique that is widely used to support controller-based continuous locomotion. It helps in reducing vection-induced VR sickness by partly reducing the optical flow in the periphery, while the user's cognition of orientation is significantly affected. This technique, which is also known as tunneling, is commonly implemented by an opaque texture (pitch-black) with a smooth-edged transparent circle in the center of the view, sometimes dynamically controlling the radius of the FoV based on velocity (Fernandes and Feiner (2016)). Studies have demonstrated that covering visual information in peripheral vision is somewhat effective at mitigating VR sickness (Bos et al. (2010); Keshavarz et al. (2011)).

However, decreases in the size of the FoV are correlated with decreases in the sense of presence (Lin et al. (2002); Seay et al. (2001)). As such, it is difficult, if not impossible to find an FoV restrictor size that results both in significantly reduced VR sickness and a high level of presence (Fernandes and Feiner (2016); Runeson (1974)). To overcome the drawbacks of pitch-black FoV restrictors, eye trackers have also been applied to create a foveated FoV restrictor (Adhanom et al. (2020)), though results still did not completely solve the problem of reducing sickness while achieving a high level of presence. In addition, blur effects have been widely explored as FoV modifiers in place of FoV restrictors with the thought that they might reduce optical flow but preserve sense of presence and immersion (Hillaire et al. (2008); Budhiraja et al. (2017); Carnegie and Rhee (2015); Nie et al. (2019); Lin et al. (2020)). Though VR sickness may be reduced, blurring still does not fully block optical flow in the user's FoV, for example for large, high-contrast objects such as pillars or corners, which can still lead to disorientation.

In this work, I introduce a novel set of algorithms that dynamically modulate the user's vision with a low-contrast screen effect. Simply put, my goals are to 1) reduce the amount of optical flow in the peripheral field of view, 2) do this using a shading technique that does not cause noticeable distraction and reduce sense of immersion, and 3) apply this technique only during sickness-inducing motions. To accomplish this, I have designed an algorithm that changes intensity in response to both the user's virtual motions and the depth of pixels in world space. Specifically, the vision modulator described here is correlated with the conflicting delta velocities and angles resulting

from the combination of controller-based linear motions with head rotations and translations. Moreover, instead of implementing a pitch-black restrictor, my strategy was to reproduce the coloration of the environment without generating optical flow. The color of the low-contrast effect is governed by the average color of the user's view, sparsely sampled from pixels on the screen on a frame-by-frame basis. My goal is to mitigate VR sickness in controller-based continuous locomotion by reducing the speed users perceive when their virtual motion changes to make the view look more stable. My contributions in this work are summarized as follows:

- The design, testing, and refinement of a shading technique (vision modulator) designed to reduce VR sickness and without becoming distracting by reducing optical flow, engaging only during potentially sickness-causing motions, and preserving scene coloration.
- Two experiments with different washout periods testing this modulator against a state-of-the-art FoV-modulation technique and a baseline condition, as well as analysis and discussion of the results.

In the following sections, I discuss related work, detail the design process and implementation of the vision modulator, describe experiments, and provide an analysis of the data. I conclude with a discussion of my results and observations.

3.2 Related Work

VR sickness is a major problem in VR that not only stops users from long-term use, but also impedes controller-based continuous locomotion—the easiest and least expensive travel method. This has stood in the way of wide-scale adoption of VR. In this section, I discuss various theories of VR sickness and vection perception, as well as previous proposed solutions that extended from these theories.

3.2.1 Visual Perception of Vection and Speed

The speed of motion can be misperceived by our eyes (Runeson (1974)) in both the real world and VR. Diels and Parkes (2010) observed that speed perceived as driving forward in a driving simulator is positively correlated with FoV size. However, simply reducing FoV size may not always be effective in reducing vection, which can also be influenced by foveal visual stimulation

or by eye movements (Webb and Griffin (2003)). In addition, Banton et al. (2005) found that participants in a VE perceived a walking speed close to their actual walking speed when looking downward/leftward, but a lower speed when looking forward. Regarding this phenomenon, I believe that studies of visual angular velocity could explain it.

The relationship between visual angular velocity (i.e., the angular velocity of textures or objects passing through the view), distance, and relatively perceived speed has been studied by many researchers. For example, Brown (1931) observed that when the angular velocity of an object in view is decreased by decreasing the distance to it, the perceived speed of the object will decrease. Furthermore, Wist et al. (1975) conducted experiments based on Brown's findings, in which they tested perceived speed by a surrounding cylinder that expanded and rotated constantly, suggesting that linear velocity plays an important role along with angular velocity in the perceived speed of object-referred self-motion.

Longuet-Higgins and Prazdny (1980) explained visual angular velocity by thinking of vision as a "hemispherical pinhole camera" in which the pinhole is the focus point moving on the surface of the hemisphere. Therefore, they suggested that if the user is heading or gazing in a different direction than the direction of motion, both the translational and rotational components should be calculated to define relative motion. In addition, the well-known motion parallax theory describes the change of optical flow in the view that results from a change of the observer's viewing position. This leads to a perception of motion of stationary objects during locomotion, which is correlated to "apparent angular velocity" (Helmholtz (1925)). Further studies have investigated the effect of an object's angular velocity passing the FoV on the perception of distance and motion during locomotion (Gibson et al. (1959); Ono and Ujike (2005)). From a biological perspective, Ewert and Hock (1972) demonstrated that increased visual angular velocity leads to increased activation of motion-sensitive neurons in the retina.

On the other hand, perceived speed is also related to contrast, as proposed by Thompson (1982). Stone and Thompson (1992) observed this phenomenon in a 2-D environment found that motion was perceived to be slower as contrast is reduced. Brooks (2001) also studied this finding with binocular images moving in depth and reached the same conclusion.

3.2.2 Sickness-inducing Motions and Motion Estimation

The widely accepted sensory conflict theory (Reason (1978); Reason and Brand (1975)) explains VR sickness by the mismatch of visually perceived self-motion and motion perceived by the vestibular system. With further consideration of whether users experience motion sickness whenever they locomote in real life and in simulators, the effect of "sensory rearrangement" was proposed by Held and Bossom (1961). In addition, the relationship between motion sickness produced by "sensory rearrangement" and that resulting from external motion disturbances was defined by Oman (1982). On the other hand, Perrone and Stone (1994) proposed a model defining how the human visual cortex and brain estimate self-motion in 3-D environments. In addition, this estimation works together with gaze stabilization to produce a feeling of stability as self-motion is perceived. Further studies demonstrated that motion changes that compromise the estimation of self-motion can induce motion sickness in both the real world and VEs (Diels (2014)). These motion changes include change of velocity/acceleration (Alexander et al. (1945); Bubka et al. (2006)) and change of direction (Bonato et al. (2005); Norouzi et al. (2018)). Studying these motions, researchers also observed that there was a nonlinear relationship between the sickness participants felt and the visually perceived changes of self-motion. Young (1973) built a model to define the nonlinear relationship of visual–vestibular interaction and proposed that there is a range of "sensory rearrangement" value to induce sensory conflict.

3.2.3 Mitigation of Controller-Based Continuous Locomotion-Induced VR Sickness

In previous studies, approaches to mitigate VR sickness induced by controller-based continuous locomotion were proposed from both visual and non-visual perspectives.

From a non-visual perspective, Russell et al. (2014) found that paced diaphragmatic breathing training relieves VR sickness symptoms by activating the parasympathetic nervous system. Recent studies have tested physical stimuli to combat VR sickness. Weech et al. (2018) studied whether noisy vestibular stimulation through bone-vibration could reduce VR sickness. Peng et al. (2020) applied vibrotactile feedback to participants' heads, significantly reducing VR sickness while improving presence in virtual walking scenarios. Based on Weech's and Peng's strategies, Liu et al. (2020b) developed a head-mounted air propulsion system and demonstrated that it improved realism and mitigated VR sickness in different high-speed motion simulators. Audi-

tory approaches have also been tested. For example, Dahlman et al. (2008) compared the effectiveness of an artificial sound horizon with non-positioned sounds and demonstrated that non-positioned sounds could be used to mitigate symptoms of motion sickness. With further studies, the results of Keshavarz and Hecht (2014) showed that pleasant music reduces overall severity of visually induced motion sickness.

Compared with non-visual methods, visual methods have advantages including ease of application, no potential side effects on the body, and low cost. The most studied approach of this type is vision restrictors/modulators. Fernandes and Feiner (2016) developed a dynamic FoV restrictor for HMD-based VR in which the FoV was inversely correlated to the mismatch between the user's physical head motion and their virtual head motion. Brument et al. (2020) conducted a study to test different types of FoV restrictors that blacken or blur the periphery during different rotation gains. Wu and Rosenberg (2019) describe FoV restrictors that render a wireframe visualization of the real world geometry in the periphery to serve as a stable reference frame during virtual travel. Based on Fernandes and Feiner's study, Adhanom et al. (2020) developed a foveated FoV restrictor driven by eye tracking, but did not find that it signficantly reduced VR sickness relative to an unfoveated FoV restrictor.

Blur effects have also been studied as an approach to mitigate VR sickness. Carnegie and Rhee (2015) implemented a dynamic depth-of-field modulator with blur and demonstrated that it could reduce overall discomfort in VEs while being difficult for participants to notice. Nie et al. (2019) developed and analyzed a saliency-detection–based blur modulator for a driving simulator, demonstrating that it was effective at mitigating VR sickness. Beside FoVs, Cao et al. (2018) demonstrated the effectiveness of static and dynamic rest frames on mitigating VR sickness. However, an additional study conducted by Zielasko et al. (2019) contradicted the effectiveness of static rest frames. Chardonnet et al. (2020) tested the correlation between VR sickness and navigation parameters such as gestures and distances. Lou and Chardonnet (2019) proposed a geometry-deforming method to mitigate VR sickness.

3.3 Design and Implementation of the Dynamic Vision Modulator

To reiterate, my primary goals with the vision modulator were to 1) reduce the amount of optical flow in the peripheral field of view, 2) to do this using a

Figure 3.1: Images showing the initial blur effects that I tested: (a) average blur, (b) radial blur, and (c) a bilateral filter.

shading technique that minimizes sense of distraction and awareness, and 3) to apply this technique only during sickness-inducing motions. As such, I started by testing different blur effects, since prior work suggested that blurring might be one way to reduce VR sickness. I implemented several shaders using HLSL to generate these effects, such as average blur, radial blur and a bilateral filter, as shown in Fig. 3.1. However, practical testing of these blur effects did not show them to be effective at slowing pixel movement in the FoV. When larger objects with high contrast passed by, the frame-to-frame changes would still create significant optical flow despite any type of blurring.

Because optical flow is a result of the changing contrast of pixels, I next considered how I might be able to change contrast while still preserving scene detail. I first implemented a Unity shader that reduced screen contrast to see if this approach might have potential. The shader was attached to the main camera in the VE, which directly modifies contrast in the HMD. The next step in my design was to ensure that changes to the contrast of the FoV occurred only during actions or motions that generate VR sickness. For example, if a user is traveling at a constant velocity, there is no need to compensate for intersensory conflict, since this would be the equivalent of riding a bike or vehicle at constant speed. The real problems occur during acceleration, strafing, and combinations thereof.

Moreover, different areas of the FoV will have different angular velocities of optical flow relative to the user's motion, so I adapted the effect to process the areas that have higher visual angular velocities. Furthermore, this was triggered by motions caused by controller interaction rather than by physical head motion, including acceleration, strafing, and rotation. Finally, in contrast to other black FoV restrictors, I tried to preserve scene color as best as possible by sampling pixel colors from the areas covered by the effect. This was done in an attempt to decrease the user's awareness of the effect in order

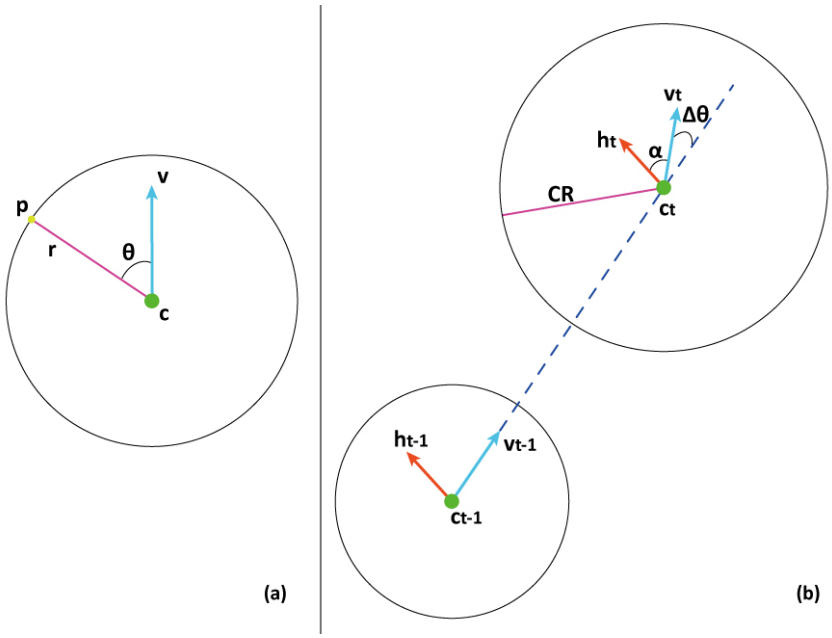

Figure 3.2: (a) Calculation of the angular velocity of a pixel, where v is the linear velocity value, c is the camera, p is the pixel, θ is the angle between the pixel's direction and linear velocity direction, and r is the distance from the camera to the pixel. (b) Parameters: c is the camera, t is the current second, h is the head direction, α is the angle between head direction and linear velocity direction, v is the linear velocity value, $\Delta\theta$ is the variation of linear velocity direction, and the circles stand for covering range(CR).

to reduce the distraction it might cause.

3.3.1 Color Sampling

My next task was to design a method that could replicate scene coloration without generating optical flow. I needed to sample screen pixels in a way that both matched the scene and ran in real time on a commodity graphics card. To accomplish this, I implemented a Unity shader that accessed a 36×36 grid of pixel colors from the camera view (i.e., a fragment of the camera texture). These were sparsely sampled at a distance of 40 pixels to cover the entire camera view and run in real time. The number of sampling pixels was determined by trial and error while monitoring frame rate using an NVIDIA GeForce RTX 2070 graphics card.

After sampling the pixel grid, I calculated the final color using a depth-weighted average, which was calculated with the z-buffer values of every sampling pixel to ensure that the color accurately covered the scene's depth range.

Then I passed the final color to the following pass of the shader, where I obtained the final color and merged it with the colors of pixels using linear interpolation (Lerp). In addition, the parameter of the Lerp was defined by applying a smooth Hermite interpolation between 0 and CR (a value that will be defined in Section 3.3.2) to each pixel's distance. Finally, I compared the luminance of the pixels with the average color to make the highlighted and shadowed area more visible.

3.3.2 Vision Modulator Parameters

After sampling the color for the low-contrast effect, I determined the depth range and intensity. I aimed to cover high angular velocity (ω) areas in the view, so I first applied the angular velocity equation (Fig. 3.2-a):

$$\omega = (v * \sin\theta)/r. \tag{3.1}$$

To make the modulator triggered by motion changes and determine the covering range of depth by motion changes dynamically, the covering range (CR) (Fig. 3.3) was defined based on the angular velocity equation as:

$$CR = k * [(acc * \sin\alpha) + (v * \sin\Delta\theta)], \tag{3.2}$$

where k is the adjusting multiplier to balance the covering range and awareness, acc is the acceleration value, α is the angle between head direction and linear velocity direction, v is the linear velocity, and $\Delta\theta$ is the variation of linear velocity direction during each second. The values of acc, α, v, and $\Delta\theta$ are governed by the camera's motion in Unity (Fig. 3.2-b). To reduce the noticeability of the vision modulator, I squared CR to make the intensity increase nonlinearly.

I conducted a pilot experiment before the primary experiment to determine an appropriate value for k. A long indoor hallway VE was built with the same environmental assets and motion triggering objects (cylinders, half-opened doors, stairs, and corners) with which I was going to build the VEs for the primary experiment. Six healthy participants (6 males, mean age 26.67, SD 3.50, all familiar with VR) were recruited for the pilot experiment. They were asked to go through the hallway with a specific k value and rate the level of disturbance in their vision from the lowest 1 to the highest 7 in each trial. Participants were informed that they were allowed to terminate the experiment at any time point if they felt uncomfortable. To confirm a suitable

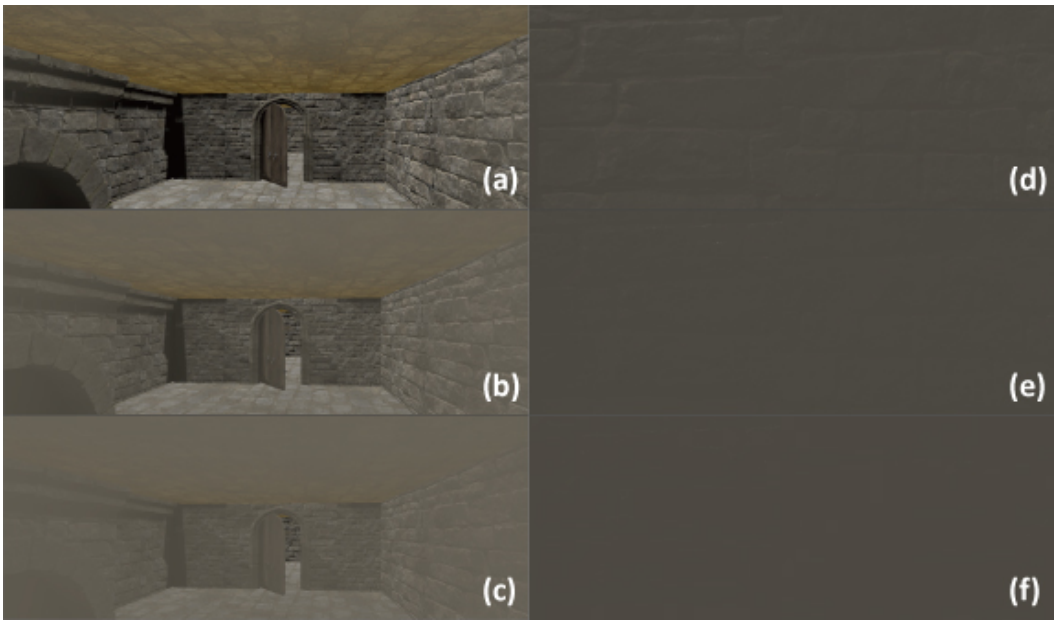

Figure 3.3: The effect of the vision modulator during continuous locomotion with different parameters. The left column shows views when facing distant objects, with parameters of (a) CR=10, (b) CR=20, (c) CR=30. The right column shows the effect when a wall is directly in front of the camera, with parameters of (d) CR=10, (e) CR=20, (f) CR=30.

value of k and minimize the effect of random fluctuation on subjective rating, k was first initialized to 0 with a step increase of 1. If a k value gets a score that is two scores higher than the score when $k = 0$, I then initialize k to a relatively high value (12) with a step decrease of 1. When running this pilot experiment, the disturbance rating typically started to increase at $k = 1$ and decrease at $k = 7$; therefore, I defined the suitable k value to be the average, $k = 4$.

3.4 Experiment 1 (Consecutive Trials)

In Experiment 1, I tested three conditions including a control condition, a dynamic FoV restrictor, and the dynamic low-contrast vision modulator in an outdoor VE and an indoor VE with a novel experiment design (see Sections 3.4.2 and 3.4.6). Participants experienced each VE while standing in place; however, they were allowed to rotate their head or body to change their heading direction in the VE. In addition, the direction of motion achieved using the joysticks was set relative to the heading direction. I measured the level of VR

sickness, the notification level of the modulators, the time participants could have experienced VR sickness, and the position where participants terminated each trial.

I then evaluated the vision modulator against a dynamic FoV restrictor and a control condition with no modulation in these highly interactive VEs to investigate four hypotheses:

- Hypothesis 1: The motion-based vision modulator has a significant effect on mitigating VR sickness induced by controller-based continuous locomotion.
- Hypothesis 2: The motion-based vision modulator is more effective than the other conditions at balancing the trade off between noticeability and VR sickness.
- Hypothesis 3: VEs that have a higher average scene depth of pixels in the participant's view (i.e., the outdoor scene with its skybox) will cause less VR sickness than those that have a lower average scene depth of pixel (i.e., indoor scenes with ceilings).
- Hypothesis 4: The proposed VR-sickness–inducing motions (acceleration, strafing, and linear velocity) are potentially more effective at provoking VR sickness than constant forward motions.

3.4.1 Equipment

I used an HTC VIVE Pro Eye HMD with integrated binocular eye tracker, a PC with Intel Core i7-9750H processor (16GB RAM) and NVIDIA GeForce RTX 2070 graphics card. The VEs for the experiments were developed in and driven by Unity 2019.3.13f1 and SteamVR. The position and orientation of the cameras in the VEs were governed by the 6DoF pose of the HMD tracked using SteamVR base stations. I used HTC Vive controllers, with the trackpad on the left-hand controller set to control motion as a joystick.

3.4.2 Experiment Design

I conducted the experiment with two separate user groups, one with the outdoor VE and one with the indoor VE. This design allowed us to perform a within-subjects analysis for each of the three sickness alleviation methods, and a between-subjects analysis of the indoor vs. outdoor conditions. In each group, all participants experienced all three rendering conditions: a control condition without any visual modification, a condition with a black FoV restrictor, and a condition with my vision modulator.

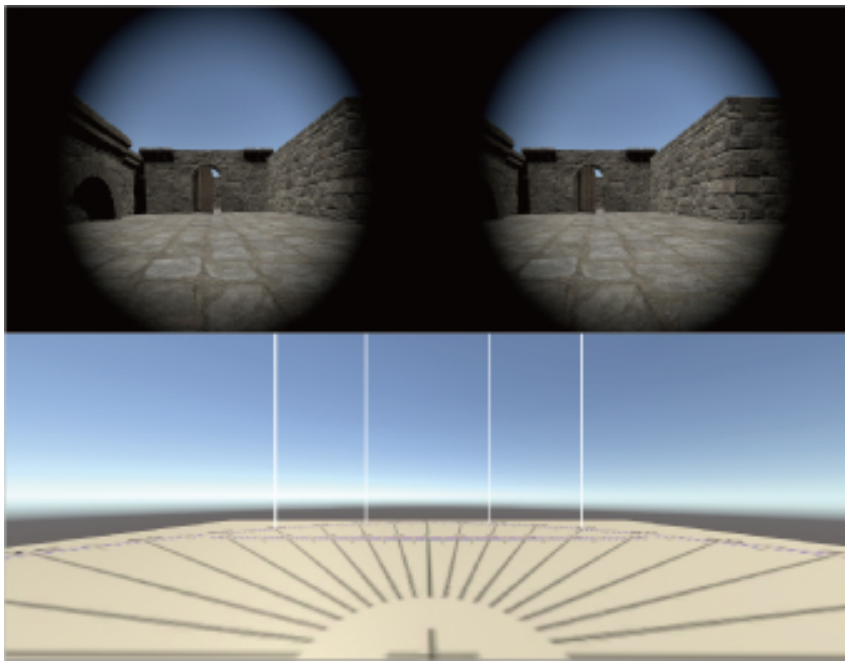

Figure 3.4: Implementation of the black FoV restrictor and the degree tester.

In both VEs, moving speed was set to a maximum of 2.5m/s, the approximate average running speed of a healthy 20 to 40 year-old human [1]. Participants were randomly assigned to the indoor-VE or outdoor-VE group. For each trial, participants traversed the VE of their group as described in Section 3.4.3. Participants were asked to end the trial as soon as they started to feel uncomfortable and the time they spent in the VE up to that point was recorded. After each trial, I calculated the "active time" of each participant as

$$AT = RT - IT, \tag{3.3}$$

where AT is the active time that a participant is moving with the controller and might be experiencing motion-induced VR sickness in the VE, RT is the duration from start to finish, and IT is the time during which the participant was not manipulating the controller. I set the maximum time for each trial to 10 minutes, which I believed to be long enough to induce VR sickness, but short enough that no one would complete the maze task. I recorded velocity, head rotation, and acceleration of each participant to help us understand how their movements were correlated with their reported VR sickness.

[1] https://www.healthline.com/health/how-fast-can-a-human-run

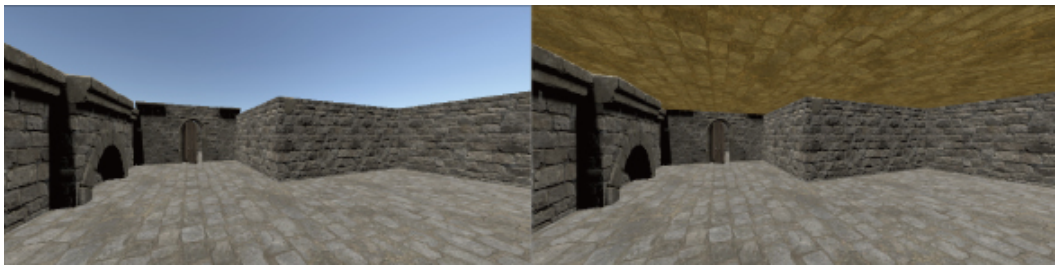

Figure 3.5: Images showing the outdoor (left) and indoor (right) VEs.

For the black FoV restrictor, I reimplemented the algorithm proposed by Fernandes and Feiner (2016) and verified that the angular size of the FoV restrictor was correct by using the FoV testing setup shown in Fig. 3.4. The minimum FoV size was set to 80 degrees, which was demonstrated to be an appropriate value to mitigate VR sickness and reduce awareness in their study. Meanwhile, for the vision modulator condition, I used the proposed algorithm and parameters I defined in Section 3.3.2.

To ensure that participants experienced motions that could increase VR sickness in the VEs, I used cylindrical obstacle objects to trigger linear rotation, half-opened doors to trigger zig-zag motions, and stair objects to trigger pitch, strafing, and acceleration. Participants would also experience acceleration, rotation, and strafing from their own active motion in the VEs.

3.4.3 Virtual Environment

The experiment was implemented with two VEs, one outdoor and one indoor (Fig. 3.5), for variety. I first built a maze scene using parts of a castle asset bundle from the Unity Asset Store [2]. I then added a diverse set of features that would require a variety of different motions and interactions to navigate (Lin et al. (2020)). The maze (Fig. 3.6) included five cylindrical obstacles (Fig. 3.7-c), four pairs of stairs (up and down) (Fig. 3.7-b) and 11 half-opened doors (Fig. 3.7-a). Participants were assigned the task of collecting seven keys (Fig. 3.7-d) to unlock the exit. The maze was large enough to support roaming for ten minutes, even when taking the shortest route to collect all the keys and get to the exit.

I aimed to measure the total time participants were moving in the VE regardless of whether they reached the exit, so I did not build different mazes. I modified the maze scene to produce the two VEs used in the experiment. For

[2]https://assetstore.unity.com/packages/3d/environments/fantasy/lordenfel-castles-dungeons-rpg-pack-175628

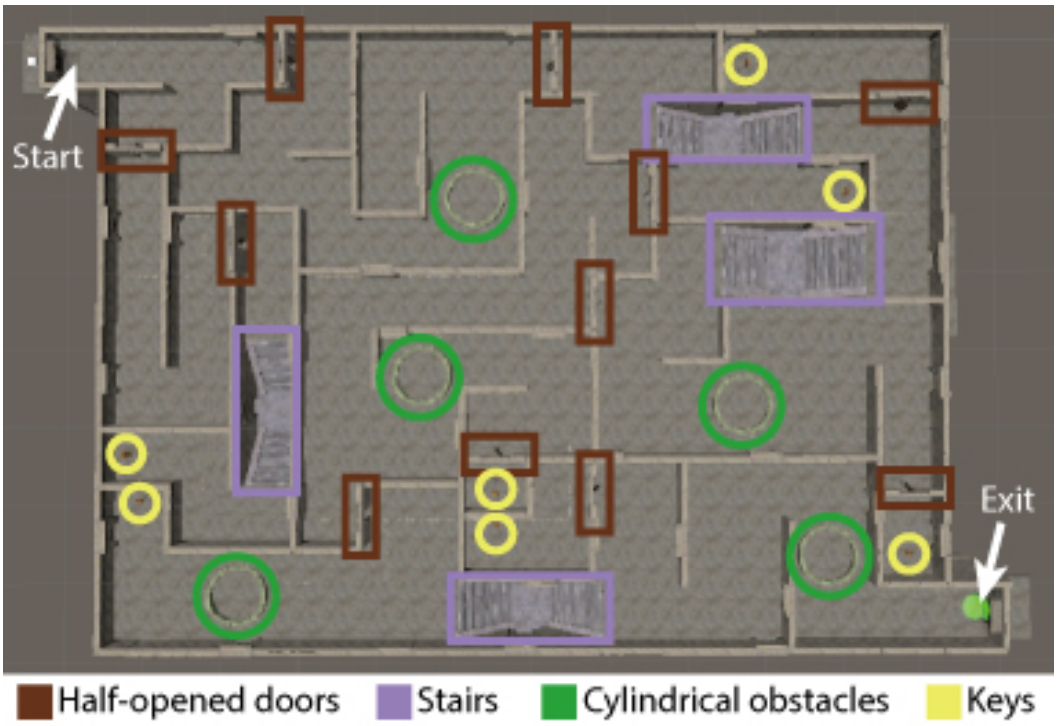

Figure 3.6: The map of the maze marked with the locations of the objects. Brown rectangles are doors, purple rectangles are stairs, green circles are wells, and yellow circles are keys.

the outdoor VE, the height of the walls was set to 3m (standard room height) to prevent participants from seeing the route over the walls and at the same time not block the view of the sky (for which I used the default Unity skybox). For the indoor VE, I used the same height for the walls, and added a roof with the same texture as the ground. Thus, participants were not able to see the sky in this indoor maze. All other features (objects, textures, terrains, routes, etc.) were identical across the two VEs. To critically test the noticeability of the vision modulator, I set all of the objects in the VEs to have different colors, as can be seen in Fig. 3.7 and Fig. 3.5.

3.4.4 Participants

34 people (19 female, mean age 24.03, SD 2.67) were recruited to take part in the experiment. All participants had normal vision, little to no experience with VR or video games and similar health conditions. They were informed that they might experience some VR sickness during the experiment and they were allowed to quit the experiment at any time. Furthermore, they were asked if they had any medical history involving their cranial nerve or vestibular

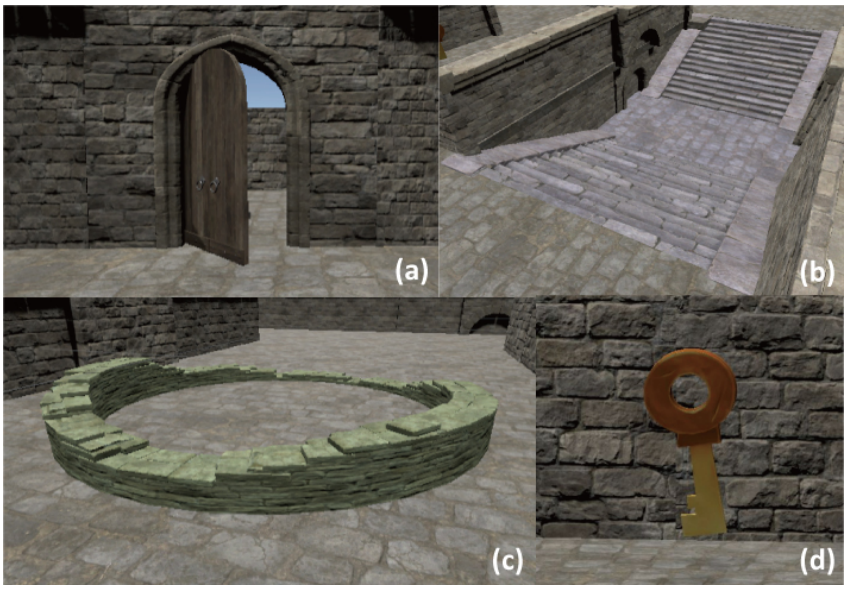

Figure 3.7: Some of the objects in the VEs designed to increase controller motions: (a) a half-opened door, (b) stairs, (c) a cylindrical obstacle, and (d) one of the keys needed to open the exit.

system. The experiment was conducted with the approval of an institutional review board (IRB).

The participants were randomly divided into two groups of 17 participants to be tested in the indoor and outdoor experiments, and the three conditions were ordered with a Latin square design.

3.4.5 Measurements

I gathered data and information from the participants using the Simulator Sickness Questionnaire (SSQ), a post-trial questionnaire, and active time and motion data.

The SSQ (Kennedy et al. (1993)) is a comprehensive questionnaire for perceived sickness and has been widely applied in studies addressing motion sickness, visually induced sickness, simulator sickness and VR sickness. The SSQ produces a total severity score based on the categories of nausea, disorientation, and oculomotor symptoms. The more VR sickness symptoms are perceived, the higher the SSQ score will be. I administered the SSQ before and after each trial to determine a participant's sickness status at each time point. To prevent the situation of "getting a higher SSQ score by a longer exposure time" and compare the results from different trials with the same standard, I use the SSQ score difference value divided by active time to calculate the

sickness per minute, which describes the average increase in sickness score over the exposure time and take that as my main measurement of increased VR sickness of the participant for each trial.

To evaluate the noticeability of each method, I designed a post-trial questionnaire based on a widely used approach in VR studies (Fernandes and Feiner (2016); Peck et al. (2009); Suma et al. (2011)). Participants were asked to rate the following subjective preference questions on a seven-point scale, from 1 ("Strongly disagree") to 7 ("Strongly agree"):

- I saw the virtual environment flicker.
- I saw the virtual environment get brighter or dimmer.
- *I felt immersed in the virtual environment.*
- I saw that something in the virtual environment had changed shape.
- I felt like I was getting bigger or smaller.
- *I was distracted during the test.*
- I felt like the virtual environment got smaller or larger.

The primary outcome measures (italicized here, but not in the actual questionnaire) were surrounded by distractor questions to avoid biasing participant answers. Furthermore, the questionnaire was presented to participants with a note stating that "These phenomena may or may not have happened."

As described in Section 3.4.2, I calculated the "active time" (AT) during which participants were exposed to stimuli that would potentially elicit VR sickness. In previous studies, measuring VR sickness in VEs usually relied on self reporting, and participants often terminated the experiment when they were no longer comfortable. This caused some participants to quit the experiment only after suffering strong VR sickness symptoms, extending the time that they participated. In contrast, I asked that participants end each trial as soon as they started to feel uncomfortable, in order to stop just around their subjective threshold (Jäger and Henn (1981); Shupak and Gordon (2006)) and prevent further aggravating symptoms. With this method, I was able to investigate whether the proposed vision modulator could extend the time spent in VR.

3.4.6 Procedure

Before starting the experiment, I briefly introduced the procedures and informed participants about the tasks and requirements. I emphasized that "If you start to feel any sickness, nausea or dizziness, even a little bit, please use

the designated controller button to end the trial immediately." Before each trial, the participant completed a pre-exposure SSQ, and then the experimental VE was presented with an HMD. The participant tried to find the exit of the maze until they ended the trial or 10 minutes had elapsed, whichever occurred first. The participant was then asked to complete a post-exposure SSQ and a post-trial questionnaire. After a 10-minute rest, the participant started the next trial, first completing a pre-exposure SSQ. This was repeated until all three conditions were tested.

Considering that sickness symptoms may remain for a long period (Dużmańska et al. (2018); Stanney and Kennedy (1998)) or even start to increase approximately 10 minutes after exposure (Min et al. (2004)), a 10-minute rest time was selected. Thus for the pre-exposure SSQ, similar to the approach used by Nie et al. (2019), if the participant's pre-exposure SSQ score was higher than 10, he/she was asked to rest for 10 minutes and fill in another pre-exposure SSQ. If this occurred three times (for a maximum rest time of 30 minutes in total before each trial) and the SSQ score was still higher than 10, the participant was asked to come back at a later date. In the actual tests, all participants were able to finish the experiment without having to return on a different day. Since the participants only experienced "just noticeable sickness" instead of long lasting exposure in each trial, this process, together with the Latin square design, would likely alleviate the influence of the residual VR sickness from other trials despite not having a long washout period.

3.4.7 Results

Here, I describe the results of Experiment 1 with respect to the two methods for mitigating VR sickness (FoV restriction and vision modulation), differences between the outdoor and indoor tests, and the effects of motion on the amount of sickness experienced. I used a level of 0.05 for determining significance.

3.4.7.1 Relative Effectiveness of the Modulator and Restrictor

First, I conducted a Shapiro test and created QQ plots, after which I determined that the sickness per minute and SSQ data were not normally distributed. Therefore, I conducted non-parametric tests on these two kinds of data instead of ANOVA/MANOVA.

A Friedman test was first conducted on the amount of sickness per minute for the three conditions: the control condition, black FoV restrictor, and my vision modulator ($\chi^2 = 16.53$, $p < 0.001$, $W = 0.78$). Next, Wilcoxon paired tests revealed that both the vision modulator ($V = 440$, $p < 0.05$, $W = 0.85$)

and the black FoV restrictor ($V = 453$, $p < 0.01$, $W = 0.85$) outperformed the control condition (Fig. 3.8); however, there was no significant difference between the vision modulator and the black FoV restrictor. I separately analyzed the three sub-scale scores and the total severity of the post-SSQ with Friedman and Wilcoxon tests: For nausea ($\chi^2 = 11.35$, $p < 0.01$, $W = 0.90$) (Fig. 3.9-a), significance differences were found in the control condition vs. the black FoV restrictor ($V = 261$, $p < 0.01$, $W = 0.82$) and the control condition vs. the vision modulator ($V = 272$, $p < 0.01$, $W = 0.83$). For oculomotor ($\chi^2 = 7.61$, $p < 0.05$, $W = 0.62$) (Fig. 3.9-b), a significant difference was found in the control condition vs. the vision modulator ($V = 257$, $p < 0.05$, $W = 0.71$). For disorientation (Fig. 3.9-c), no significant difference was found ($\chi^2 = 5.37$, $p = 0.07$, $W = 0.84$). For total severity ($\chi^2 = 10.97$, $p < 0.01$, $W = 0.74$) (Fig. 3.9-d), significant differences were found in the control condition vs. the black FoV restrictor ($V = 297$, $p < 0.05$, $W = 0.83$) and the control condition vs. the vision modulator ($V = 386$, $p < 0.01$, $W = 0.81$). A MANOVA test on the active time data found no significant differences between the effectiveness of the three conditions ($F(2,32) = 3.09$, $p = 0.87$, $d = 0.43$) (Fig. 3.10).

Furthermore, I analyzed the data from post-trial questionnaires with Friedman tests, and found no significant differences between the three conditions for the "immersion" ($\chi^2 = 1.40$, $p = 0.496$, $W = 0.77$) and "distraction" ($\chi^2 = 0.03$, $p = 0.98$, $W = 0.84$) questions (Fig. 3.11).

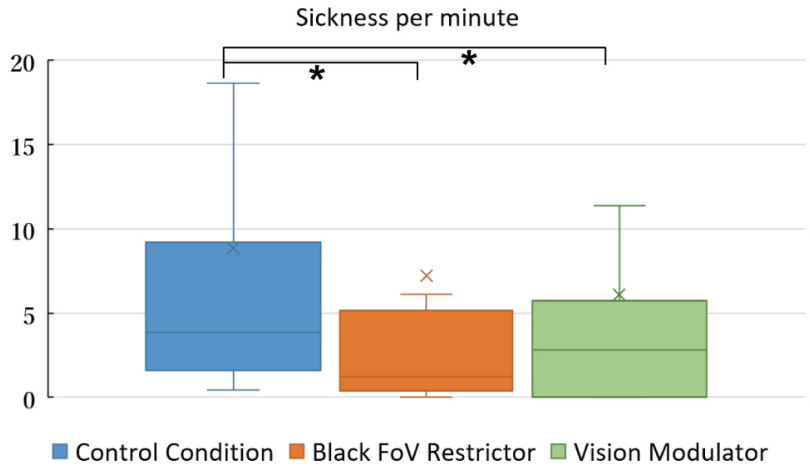

Figure 3.8: Subjective sickness results divided by the time in minutes spent in the VE for each condition.

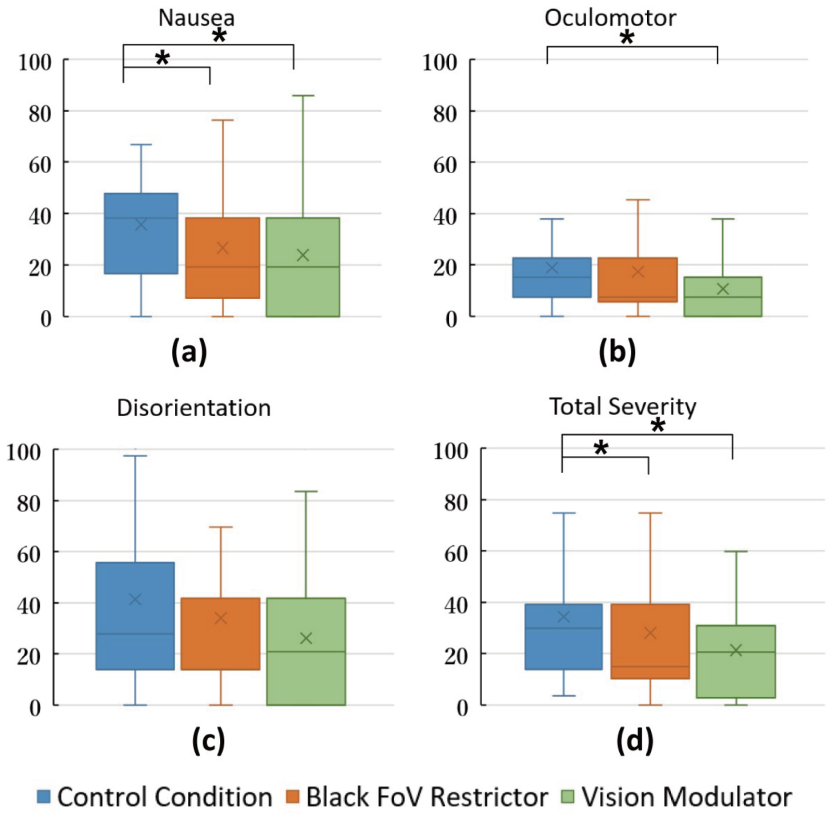

Figure 3.9: Analysis of post-SSQ scoring for (a) nausea, (b) oculomotor, (c) disorientation, and (d) total severity, using the formula from Kennedy et al. (1993).

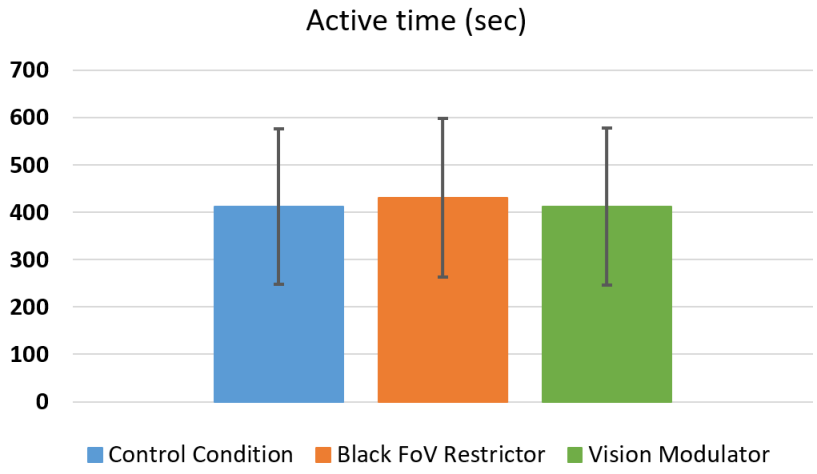

Figure 3.10: Active time that participants spent in the VE for each condition.

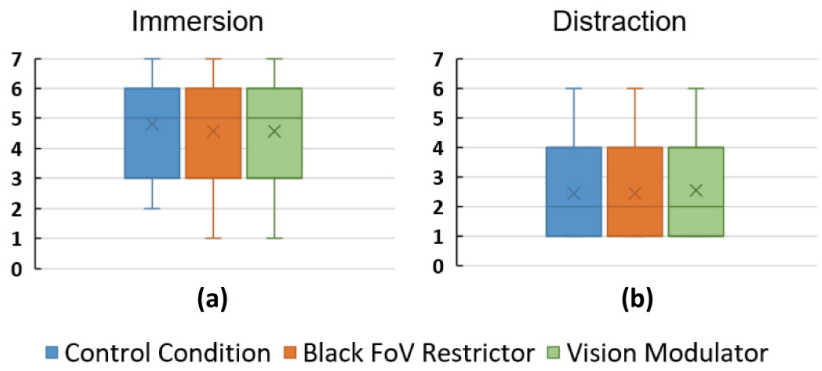

Figure 3.11: Rating results of "immersion" and "distraction" from post-trial questionnaires.

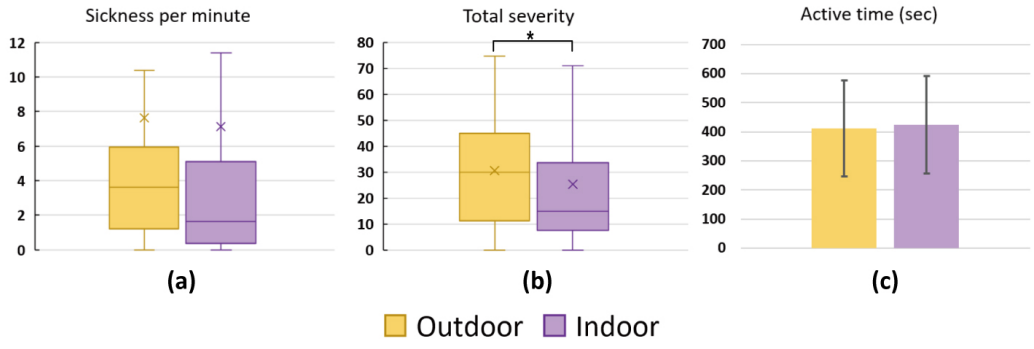

Figure 3.12: Data plots of (a) sickness per minute, (b) Total severity and (c) active time for outdoor vs. indoor.

3.4.7.2 Outdoor vs. Indoor

To determine whether the outdoor and indoor environments were different when holding all else constant, I analyzed the data of sickness per minute (Wilcoxon, $V = 478$, $p = 0.18$, $W = 0.44$), total severity of post-SSQ (Wilcoxon, $V = 320$, $p < 0.05$, $W = 0.48$) and active time (MANOVA, F(1,50) = 3.94, $p = 0.72$, $d = 0.70$). These results revealed that there might be some significant effects of the three conditions in different environments (Fig. 3.12). Therefore, I proceeded to analyze the total severity data from the three conditions separately with Wilcoxon paired tests, and no further significance was found.

3.4.7.3 Effects of Motion

The motion data I collected during the experiment included: linear velocity, acceleration, and the angle between head and velocity directions, which were the most likely to have an effect on VR sickness.

To better analyze how the difference between head angle and direction of motion that occurred during continuous locomotion (both forward motion and strafing) might affect SSQ scores, I collected only the frames with nonzero linear velocity and averaged the angle differences to get the average difference in angle per frame. I compared the motion data with sickness per minute in order to investigate whether relationships existed between motion and VR sickness. I used Pearson Correlation between the motion data and sickness metrics separately for each of the three conditions. In addition, outliers which beyond three times standard deviation (the "three-sigma rule") were removed. Regarding average angle differences: For the black FoV restrictor, there was a positive correlation ($t = 3.24$, $r = 0.51$, $p < 0.01$) (Fig. 3.13-b). For the control condition ($t = 1.83$, $r = 0.31$, $p = 0.08$) (Fig. 3.13-a) and the vision modulator ($t = 1.87$, $r = 0.32$, $p = 0.07$) (Fig. 3.13-c), there were no significant correlations. Regarding average linear velocity: For the control condition, there was no significant correlation ($t = 0.17$, $r = 0.03$, $p = 0.86$) (Fig. 3.14-a). For the black FoV restrictor, there was a negative correlation ($t = -4.37$, $r = -0.62$, $p < 0.001$) (Fig. 3.14-b). For the vision modulator, there was a negative correlation ($t = -3.56$, $r = -0.54$, $p < 0.01$) (Fig. 3.14-c). Regarding average acceleration: For the control condition ($t = -1.35$, $r = -0.23$, $p = 0.19$) (Fig. 3.15-a) and the black FoV restrictor ($t = -2.69$, $r = -0.44$, $p < 0.05$) (Fig. 3.15-b), there were no significant correlation. For the vision modulator, there was a negative correlation ($t = -3.27$, $r = -0.51$, $p < 0.01$) (Fig. 3.15-c).

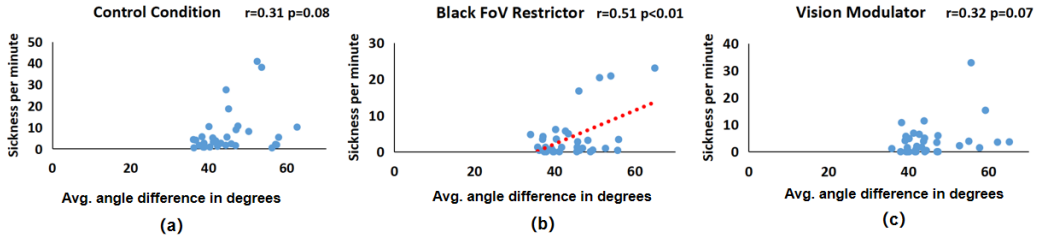

Figure 3.13: The Pearson correlation between sickness per minute and average angular difference in degrees between head angle and direction of motion: (b) For the black FoV restrictor, r=0.51. For the control condition (a) and the vision modulator (c), no significant correlation were found.

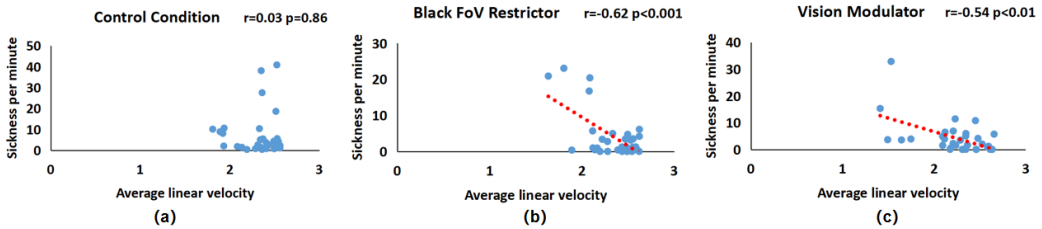

Figure 3.14: The Pearson correlation between sickness per minute and average linear velocity: (a) For the control condition, no significant correlation was found. (b) For the black FoV restrictor, r=-0.62. (c) For the vision modulator, r=-0.54.

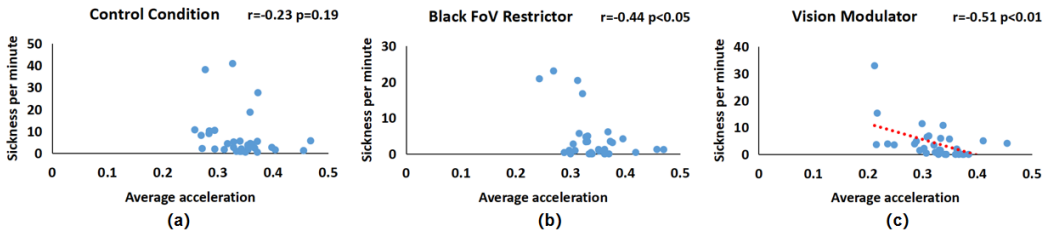

Figure 3.15: The Pearson correlation between sickness per minute and average acceleration: For the control condition (a) and the black FoV restrictor (b), no significant correlation was found. (c) For the vision modulator, r=-0.51.

3.5 Experiment 2 (Multi-Day Trials)

In Experiment 2, I tested the three conditions in the two VEs with a longer washout time; everything else remained the same as in Experiment 1 (Section 3.4).

3.5.1 Participants

22 people (11 female, mean age 24.05, SD 2.17) were recruited to take part in the experiment. All participants had normal vision, little to no experience with VR or video games and similar health conditions. They were informed that they might experience some VR sickness during the experiment and they were allowed to quit the experiment at any time. Furthermore, they were asked if they had any medical history involving their cranial nerve or vestibular system. The experiment was conducted with the approval of an institutional review board (IRB).

The participants were randomly divided into two groups of 11 participants to be tested in the indoor and outdoor experiments, and the three conditions were ordered with a Latin square design.

3.5.2 Procedure

Before starting the experiment, I briefly introduced the procedures and informed participants about the tasks and requirements. I emphasized that "If you start to feel any sickness, nausea or dizziness, even a little bit, please use the designated controller button to end the trial immediately." Before each trial, the participant completed a pre-exposure SSQ, and then the experimental VE was presented with an HMD. The participant tried to find the maze exit until they ended the trial or 10 minutes had elapsed, whichever occurred first. The participant was then asked to complete a post-exposure SSQ and a post-trial questionnaire. The three trials were conducted on separate days (Fernandes and Feiner (2016); Adhanom et al. (2020)).

For the pre-exposure SSQ, similar to the approach used by Nie et al. (2019), if the participant's pre-exposure SSQ score was higher than 10, he/she was asked to rest for 10 minutes and fill in another pre-exposure SSQ. If this occurred three times (for a maximum rest time of 30 minutes in total before each trial) and the SSQ score was still higher than 10, the participant was asked to come back at a later date. In the actual tests, all participants were able to finish each trial without having to return on a different day.

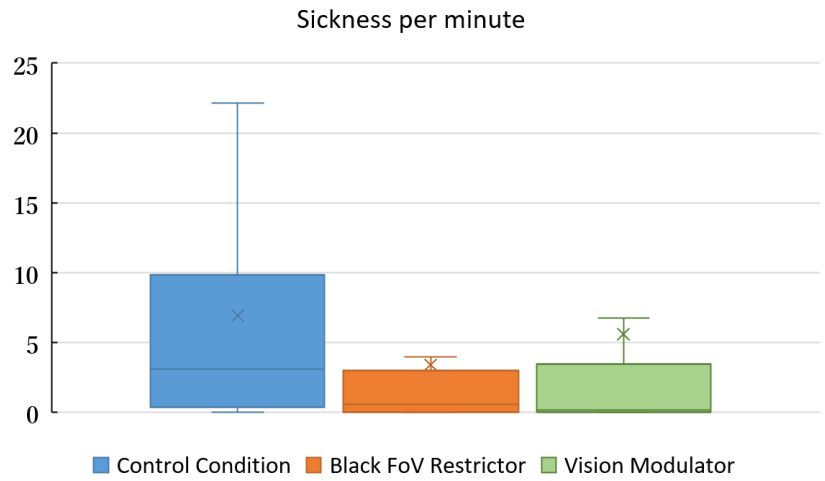

Figure 3.16: Subjective sickness results divided by the time in minutes spent in the VE for each condition.

3.5.3 Results

In this section, I describe the results of Experiment 2 with respect to the two methods for mitigating VR sickness (FoV restriction and vision modulation), differences between the outdoor and indoor tests, and the effects of motions on the amount of sickness experienced. I used a level of 0.05 for determining significance.

3.5.3.1 Relative Effectiveness of the Modulator and Restrictor

Using the same statistical analysis as Experiment 1, I conducted a Shapiro test and drew a QQ plot, which helped us determine that the sickness per minute data and SSQ were not normally distributed. After that, I conducted non-parametric tests on the two kinds of data instead of ANOVA/MANOVA.

A Friedman test was first conducted on the amount of sickness per minute for the three conditions: the control condition, black FoV restrictor, and my vision modulator. No significant difference ($\chi^2 = 3.34$, $p = 0.19$, $W = 0.55$) between the three conditions was found (Fig. 3.16). I separately analyzed the three sub-scale scores and total severity of the post-SSQ with Friedman and Wilcoxon tests: For oculomotor ($\chi^2 = 7.42$, $p < 0.05$, $W = 0.57$) (Fig. 3.17-b), significant differences were found in the control condition vs. the black FoV restrictor ($V = 102$, $p < 0.05$, $W = 0.62$) and the control condition vs. the vision modulator ($V = 114$, $p < 0.05$, $W = 0.60$). Moreover, no significance was found for nausea ($\chi^2 = 2.03$, $p = 0.36$, $W = 0.66$) (Fig. 3.17-

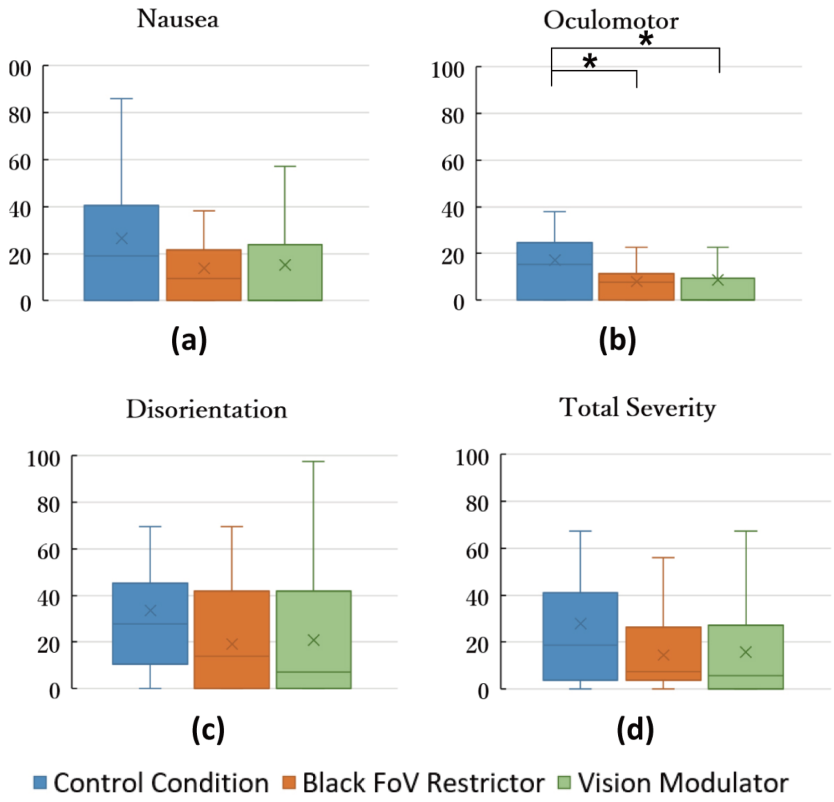

Figure 3.17: Analysis of post-SSQ scoring for (a) nausea, (b) oculomotor, (c) disorientation, and (d) total severity, using the formula from Kennedy et al. (1993).

a), disorientation ($\chi^2 = 2.94$, $p = 0.23$, $W = 0.60$) (Fig. 3.17-c) or total severity ($\chi^2 = 4.46$, $p = 0.11$, $W = 0.61$) (Fig. 3.17-d). I also applied a MANOVA test on the active time data and found no significant differences between the effectiveness of the three conditions (F(2,20) = 3.15, $p = 0.24$, $d = 0.55$) (Fig. 3.18).

Furthermore, I analyzed the data from post-trial questionnaires with a Friedman test, and found no significant difference between the three conditions for the "immersion" ($\chi^2 = 0.72$, $p = 0.70$, $W = 0.22$) and "distraction" ($\chi^2 = 1.10$, $p = 0.58$, $W = 0.29$) questions (Fig. 3.19).

3.5.3.2 Outdoor vs. Indoor

To determine whether the outdoor and indoor environments were different when holding all else constant, I analyzed sickness per minute (Wilcoxon, $V = 192$, $p = 0.95$, $W = 0.63$), total severity of post-SSQ (Wilcoxon, $V = 162$, $p = 0.75$, $W = 0.71$), and active time (MANOVA, F(1,32) = 3.15, $p = 0.24$,

$d = 0.79$). These results revealed that there was no significant effect of the three conditions in different environments (Fig. 3.20).

3.5.3.3 Effects of Motion

The motion data I collected during the experiment included: linear velocity, acceleration, and the angle between head and velocity directions, which were the most likely to have an effect on VR sickness.

The motion data were collected and compared with sickness per minute similar to Experiment 1 (Section 3.4.7.3). Regarding average angle differences: For the black FoV restrictor, there was a positive correlation ($t = 2.51$, $r = 0.50$, $p < 0.05$) (Fig. 3.21-b). For the control condition ($t = -1.32$, $r = -0.29$, $p = 0.20$) (Fig. 3.21-a) and the vision modulator ($t = 1.50$, $r = 0.33$, $p = 0.15$) (Fig. 3.21-c), there were no significant correlation. Regarding average linear velocity: For the control condition, there was a negative correlation ($t = -3.80$, $r = -0.66$, $p < 0.01$) (Fig. 3.22-a). For the black FoV restrictor, there was a negative correlation ($t = -2.52$, $r = -0.50$, $p < 0.05$) (Fig. 3.22-b). For the vision modulator, there was a negative correlation ($t = -3.27$, $r = -0.61$, $p < 0.01$) (Fig. 3.22-c). Regarding average acceleration: For the control condition, there was a negative correlation ($t = -3.38$, $r = -0.61$, $p < 0.01$) (Fig. 3.23-a). For the black FoV restrictor ($t = -1.46$, $r = -0.32$, $p = 0.16$) (Fig. 3.23-b) and the vision modulator ($t = -1.62$, $r = -0.36$, $p = 0.12$) (Fig. 3.23-c), there were no significant correlation.

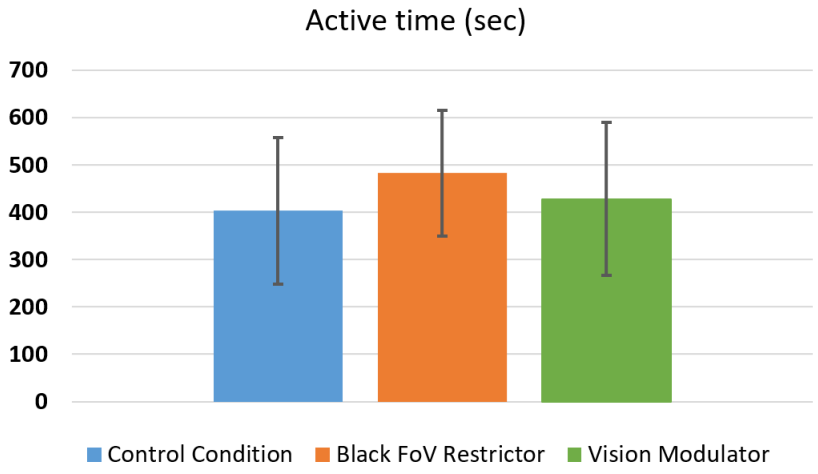

Figure 3.18: Active time that participants spent in the VE for each condition.

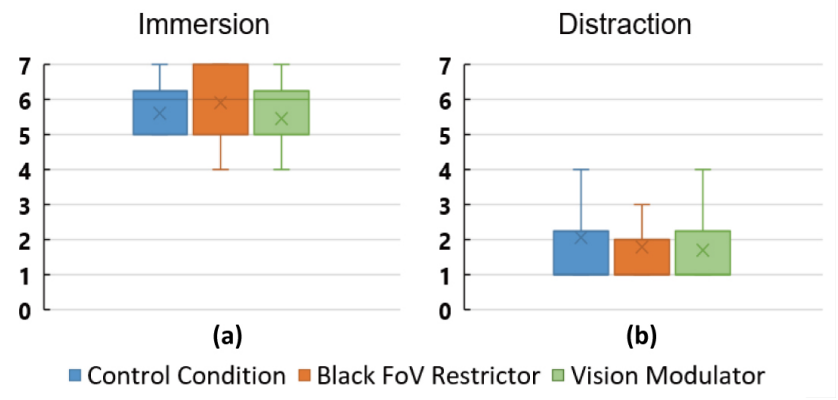

Figure 3.19: Rating results of "immersion" and "distraction" from post-trial questionnaires.

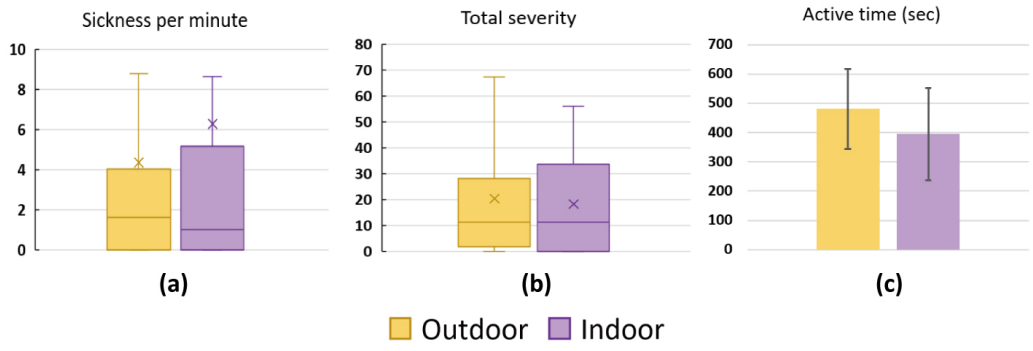

Figure 3.20: Data plots of (a) sickness per minute, (b) Total severity and (c) active time for outdoor vs. indoor.

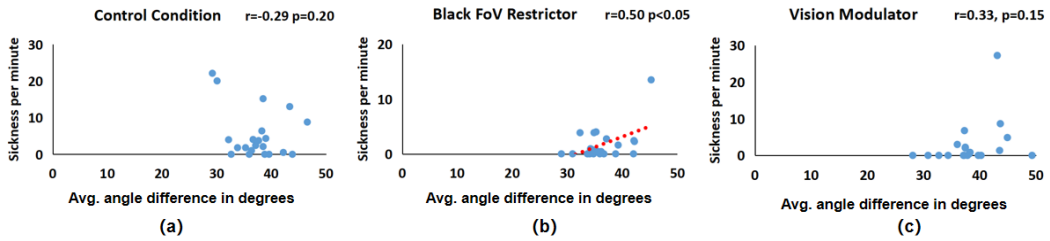

Figure 3.21: The Pearson correlation between sickness per minute and average angular difference in degrees between head angle and direction of motion: (b) For the black FoV restrictor, r=0.50. For the control condition (a) and the vision modulator (c), no significant correlation were found.

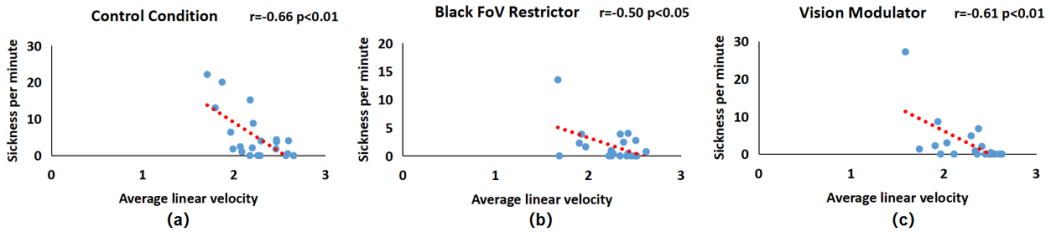

Figure 3.22: The Pearson correlation between sickness per minute and average linear velocity: (a) For the control condition, r=-0.66. (b) For the black FoV restrictor, r=-0.50. (c) For the vision modulator, r=-0.61.

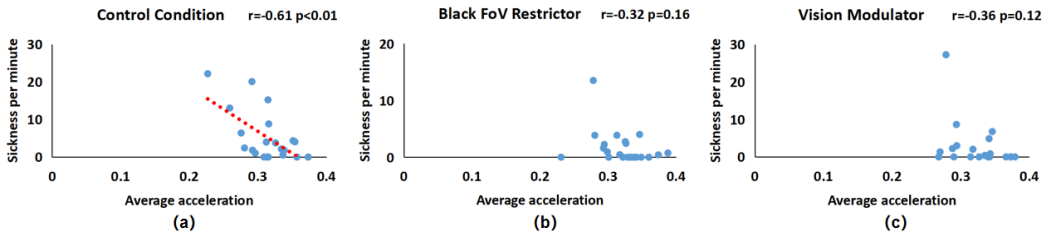

Figure 3.23: The Pearson correlation between sickness per minute and average acceleration: (a) For the control condition, r=-0.61. For the black FoV restrictor (b) and the vision modulator (c), no significant correlation were found.

3.6 Discussion

The results I gained from Experiment 1 demonstrated some effectiveness of the proposed vision modulator. In addition, my findings also indicated that the low-contrast effect is helpful in reducing VR sickness. The experimental results suggested that the visual methods were neither distracting nor noticeable, so they would not likely affect the conscious perception of vection (Riecke et al. (2006)). My goal is to preserve the conscious perception of self motion, but reduce intersensory conflict (i.e., vection in the periphery that is not consciously perceived but contributes to VR sickness), while preserving vection in the central field of view.

3.6.1 Experiment Results (Consecutive Trials)

Based on the sickness per minute data and the total severity data, both the black FoV restrictor and the vision modulator outperformed the control condition, which supports my H1. Although the two methods did not show an overall significant difference, there were some differences. The SSQ oculomotor score and disorientation score indicated that participants experienced significantly less severe oculomotor and disorientation symptoms with the vision modulator. On the other hand, the black FoV restrictor did not have an effect on these two measures.

Unlike FoV restrictors, my vision modulator is not restricted to a 2-D FoV radius, and it can modify pixels both in central vision and the periphery. It also dynamically changes size, depth and intensity based on motion. Therefore, with unblocked peripheral vision, users should not need to turn their eyes and heads as often as with a restricted FoV while moving in a complex VE. After the experiment, several participants orally reported that they thought the black FoV restrictor was effective when they were moving straight ahead and the vision modulator was effective when they were trying to cross the half-opened doors or moving on the stairs. These observations supported the results obtained from the oculomotor and disorientation scores.

Regarding H2, my analysis of the post-trial questionnaires revealed that participants did not feel any significant loss of immersion or strong distraction with both of the VR sickness mitigating methods, which is in line with the demonstration that Zielasko et al. (2018) have made on the distraction of dynamic peripheral rendering. However, this does not mean that the methods were unnoticeable all the time. During the experiment, several participants orally reported that they noticed a "black shadow" in front of their eyes or they felt like there was something getting darker when they were tested with the

black FoV restrictor. On the other hand, several participants orally reported that they noticed "fog" in their view or that they felt like the lenses were unclear when they were tested with the vision modulator. However, these experiences may also be induced by the VEs, since a few participants orally reported perceiving unclearness and shadows in the control condition. I believe that as long as the method is designed to counterbalance visual sensory conflicts, users are likely to notice it at some point in time. Instead of pursuing "unnoticeable" approaches, future studies might focus on making the method less distracting or more immersive by designing it to be more consistent with the VEs or the users' operations.

In H3, I proposed that outdoor VEs may induce less VR sickness than indoor VEs since there are fewer low-depth pixels toward the front of the participant's view. Especially when the skybox is completely blue, pixel movements would not be noticed. Therefore the overall percentage of high visual angular velocity area in outdoor VEs would be lower and produce lower "sensory rearrangement" (Oman (1982)) than the indoor VEs. However, my experimental results did not support this hypothesis. Considering that participants moved freely and looked freely in the experiment, they did not keep the sky in their view all the time when they were traveling through the outdoor VE. On the other hand, since the maze was relatively large, there were also high-depth areas in view of the participants in the indoor VE. In addition, I designed the VEs to be highly interactive and contain a lot of objects, which also affected the visual angular velocity of a participant's view.

Regarding H4, during the experiment, many participants reported that their feeling of VR sickness was accelerated by the half-opened doors and the stairs. This indicates that the objects successfully induced VR sickness by forcing the participant to move in a non-linear fashion. These movements were recorded when there was an angle kept between the velocity vector and the head-facing vector and were calculated as average strafing per frame. The Pearson correlation between strafing and VR sickness revealed a positive correlation for the black FoV restrictor. The Pearson correlation between linear velocity and VR sickness revealed negative correlations for both the black FoV restrictor and the proposed vision modulator, which suggests that they effectively mitigated VR sickness induced by increases in linear velocity. The Pearson correlation between acceleration and VR sickness revealed a negative correlation for the proposed vision modulator, which is reasonable since acceleration was taken into consideration in my algorithm. Lastly, for the control condition, no significant correlation between all the motion data and VR sickness was observed, which opposed my H4.

3.6.2 Experiment Results (Multi-Day Trials)

Experiment 2 was conducted to determine if I could support the findings of Experiment 1, as well as to explore the effectiveness of different washout periods, the reliability of a just-noticeable-sickness design, and the adaptability of designs with different rest periods.

The results from Experiment 2 did not show as many significant differences for either the proposed vision modulator or pitch-black FoV restrictor, though my findings did indicate that the two methods were still helpful in reducing oculomotor symptoms. One reason for this may be that since shorter multi-day trials allowed for longer rest periods, participants simply experienced less sickness overall. Still, the severity of the oculomotor-related effects was significantly decreased between control and vision modulator/FoV restrictors, confirming that the vision modulator can reduce symptoms in a similar fashion to FoV restrictors without affecting other perceptions. This is evidenced by the similar results regarding active time, post-trial questionnaires, outdoor vs. indoor environments, and the correlation between strafing and VR sickness.

For the control condition, the correlation between linear velocity and VR sickness and the correlation between acceleration and VR sickness was negative. This result demonstrated that as linear velocity and acceleration values increased, users may feel less sick during locomotion.

3.6.3 Comparison of Experiments

During Experiment 2, on many occasions participants indicated that differences in physical state (e.g. sleep, hunger, fatigue) on different days likely affected the sickness they experienced. This is a significant trade-off for experiment designs that implement a washout period of a day or more. In addition, participants reported that their standards for judgement when answering the SSQ questions and defining thresholds of sickness may also have varied on different days. In Experiment 1, the just-noticeable-sickness metric was designed to prevent the occurrence of severe symptoms, as well as the transference of symptoms into future trials, which is another trade-off between participant comfort, consistency of judgement, and residual effects of any remaining sickness. It may be beneficial to explore a single-day design in which participants experience trials several hours apart instead of several days apart to reduce the potential variability introduced by different cognitive or physical states for participants on different days. As such, researchers designing future experiments on VR sickness should consider the trade-offs of both experimental designs and make a decision based on the specific needs

and outcomes of the experimental task. If minimization of sickness symptoms is desired, shorter, multi-day trials may be justified.

Both Experiment 1 and Experiment 2 indicated that the black FoV restrictor and the vision modulator did not provoke enough distraction to affect immersion during the exposure time. Meanwhile, both methods did not show any significant help to extend the active time in my experiments, which revealed that although the methods may mitigate (Experiment 1) the feeling of VR sickness, they were not able to lengthen the time users would like to stay in VEs. Therefore, I suggest that future studies on motion-induced VR sickness should consider other motions. Regarding outdoor vs. indoor environments, no statistically significant differences between them were found in either experiment, which does not indicate that average scene depth of a pixel should be considered as a factor in inducing VR sickness. Regarding the correlation between strafing and VR sickness for the black FoV restrictor, both of the experiments indicated a positive correlation, which is similar with the findings by Norouzi et al. (2018). This result might be caused by the smaller FoV, since participants have to rotate their head with a larger range and more often to counterbalance the restricted FoV.

3.7 Chapter Conclusion and Future Work

In this work, I performed two cross-group within-subjects experiments to investigate the performance of a novel motion-based dynamic low-contrast shading technique designed to mitigate VR sickness. This method performs real-time sampling of scene pixels to blend scene color into the user's FoV, integrates user motion so that it engages only during movements that cause intersensory conflict, and uses depth-based rendering. In Experiments 1 and 2, 34 and 22 participants, respectively, were randomly divided into two groups and experienced three conditions using a VR display: a control condition with no modification, an experimental condition with a black FoV restrictor, and an experimental condition with the proposed approach.

The data collected from Experiment 1 demonstrated that both the black FoV restrictor and the proposed approach effectively mitigated visually induced VR sickness and enabled users to remain in VEs for a longer period of time than the baseline without a significant awareness of the modulator. In addition, since the VEs in the experiments were designed with multiple colors in order to test whether my approach could effectively maintain a low sense of distraction, the multiplier k was set to a relatively low value. It may be the case that this type of shading will be more effective in well-designed com-

mercial VEs with more consistent ambient colors. My implementation of FoV modifications revealed similar results to Fernandes and Feiner's study, which showed that their intervention helped participants to experience less VR sickness with no significant distraction. In contrast to other short washout period studies (Golding et al. (2012); Duan et al. (2017)), I provided an additional long washout period experiment to investigate whether washout periods make a big difference.

In future studies, I plan to further explore shading algorithms that can better reproduce scene coloration without generating optical flow. I also plan to conduct experiments testing individual motions such as acceleration and rotation. Moreover, I would like to investigate the effectiveness of the proposed approach with longer exposure times, in other environments such as games, and during more interactive tasks.

I hope that this work will not only encourage the development of other motion-dependent shading techniques that can reduce VR sickness and maintain a sense of immersion, but also encourage the exploration of short-washout-period, just-noticeable-sickness experimental designs.

Chapter 4

Real-Time Visual Augmentation for Safe Information Access

4.1 Introduction

Unlike virtual environments, locomotive scenarios in real-world such as cycling can cause dangers to the user. Thus implementations of augmented information rendering should be more careful on safety. Cycling is a very popular form of transportation around the world, and it is environmentally friendly and generally safe. However, cyclists must often ride on sidewalks, narrow or rugged paths, or situations in which obstacles can be crash hazards. To stay safe, riders must pay attention to what is in front of them without looking down and reduce their speed to avoid collisions when necessary (Vansteenkiste et al. (2013); Chong et al. (2010)). Maintaining visual attention to the forward view can also be challenging when cyclists are distracted by navigation information or smartphone messages (Ren et al. (2021)). In addition, professional cyclists typically make use of a cyclocomputer to access biometric data, which is usually mounted to a bicycle's stem, handlebars, or a forward-projecting extension. Although safety preservation and head-up information delivery approaches for driver assistance have been well-studied in the automotive domain, assistance for cyclists has not been well addressed. To alleviate the dangers of looking down at a smartphone, cyclocomputer, or other digital cycling interface, I propose an information delivery approach for safer and more efficient cycling that combines eye tracking with adaptive augmented visualization.

AR systems have increasingly been implemented with HMDs to augment user vision and provide more accessible digital content. Head-worn technology is an excellent candidate for cycling since riders already wear helmets while riding, and HMDs can easily be integrated into helmets. In addition, bike-mounted screens (e.g. smartphones or cyclocomputers) that function as cycling support systems require cyclists to look down frequently, potentially causing distractions and accidents (De Waard et al. (2014); Terzano (2013)). Optical see-through headworn AR displays have the potential to alleviate this

problem since information can be overlaid in the user's natural cycling field of view. In order to develop an optical see-through interface for cyclists, the detection of, and interaction between, digital content and real-world hazards must be handled by the system in real time. For this purpose, machine learning and deep neural network (DNN) models are well developed for human and object detection, and they have already been widely applied in the fields of unmanned aerial vehicles, robotics, and autonomous driving. Although the performance of the obstacle detecting DNN models remains non-ideal in high-speed scenarios (Falanga et al. (2017), [1]), they can be incredibly useful for bicycles, which travel at relatively low speeds compared to vehicles. In addition, AR approaches are widely applied in driving and hazard avoidance experiments since they can improve immersion while reducing danger to the participants (Lindemann and Rigoll (2017); Yeo et al. (2020); Oliveira et al. (2018)).

In this work, I introduce a novel approach called "HazARdSnap" that not only improves accessibility of bike-related user interfaces (UIs) but also preserves awareness of hazards (e.g. pedestrians, cyclists and potholes). Several images outlining HazARdSnap are shown in Fig. 4.1. First, to determine the location and size of potential hazards, I gather scene data with an RGB-D camera and process scenery in front of the bike through a set of computer vision based algorithms. This data is sent to a HoloLens 2 via TCP/IP socket communication. Then, scene data, motion, gaze, and head position are gathered from the HMD.

Using this data as input, I designed an algorithm that detects hazards and snaps any relevant UI information onto those hazards when they become gaze targets. This allows simultaneous visual access to the information and awareness of the hazard's location and speed. Simply put, I provide better information access to cyclists without compromising their perception of forward-located hazards. In an experiment using real bicycles with virtual hazards (for participant safety), I evaluated HazARdSnap, smartphone-based presentation, and a forward-fixed UI delivered by an HMD.

[1] https://roboticstoday.github.io/

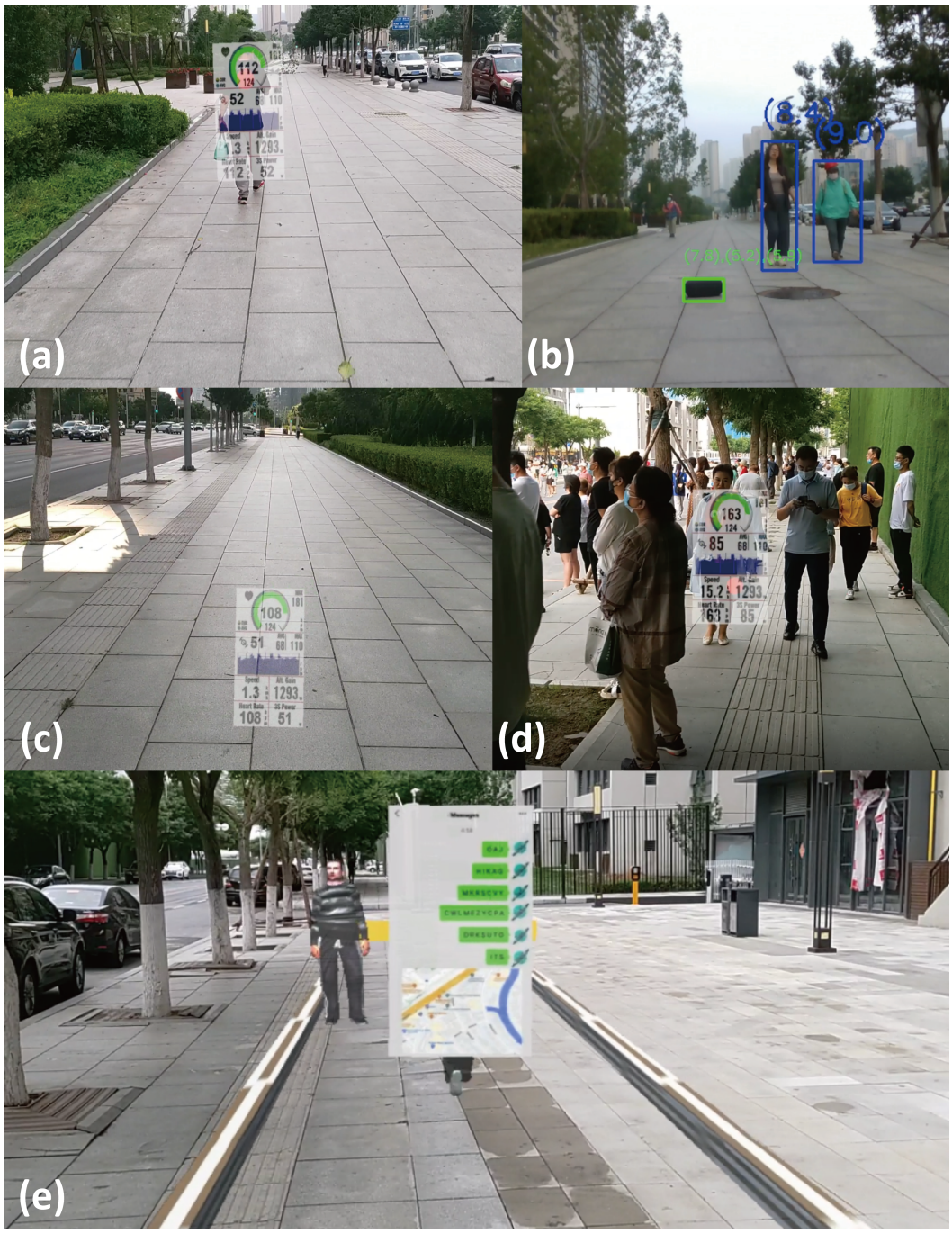

Figure 4.1: Images showing (a) the HazARdSnap snapped to a pedestrian in the gaze path of the cyclist, (b) the computer vision algorithms I designed to detect pedestrians and potholes in real-time, (c) the UI fixed to the forward direction when no hazards are detected, (d) an example of placement in a more complex environment, in this case on a crowded sidewalk, and (e) a screenshot from an MR based user study in which participants avoided collisions with virtual hazards while reading virtual content.

The contributions in this work are summarized as follows:

- The design, testing, and refinement of a UI snapping technique, HazARdSnap, designed to decrease the chance of collisions when accessing digital information during cycling.
- A bike-mounted system that makes use of computer vision, depth information, and environment tracking to detect oncoming hazards such as potholes and pedestrians.
- An MR based user study that compares cyclist performance while using HazARdSnap, a forward-fixed UI, and a bike-mounted smartphone UI as well as analysis and discussion of metrics like collisions with virtual objects, time spent viewing the road, and subjective ratings.

In the following sections, I discuss related work, detail the design process and implementation of HazARdSnap, describe my user study, and provide an analysis of the results. I conclude with a discussion of my findings and observations.

4.2 Related Work

When in motion, cyclists can often be distracted when checking their smartphones or cyclocomputers for navigation instructions, biometric information, or messages, especially when riding on low-quality surfaces or passing pedestrians or other cyclists. To help address these problems, many studies have been conducted for cycling and driving support. In this section, I discuss previously proposed approaches for improving both safety and information accessibility for cyclists and drivers.

4.2.1 Information Delivery

Cycling through narrow tracks or terrain with obstacles often requires high attention to the forward direction, frequent steering and low speed (Vansteenkiste et al. (2013)). These challenges are especially pronounced for people who lack spatial awareness, such as children (Vansteenkiste et al. (2015)). Moreover, Van der Horst et al. (2014) demonstrated that the cyclists' behaviors on sidewalks would be conflicted by pedestrians. On the other hand, using supporting functions (i.e. Navigation) in mobile phones or other displays causes distractions (Ren et al. (2021)) to cyclists and drivers. To solve this problem, Dancu

et al. (2015) evaluated their projection-based interface with a head-up display. Further, AR displays were proposed as a viable solution: Ginters (2019) developed an AR cycling support system which delivers information via a HMD. von Sawitzky et al. (2020) investigated the effectiveness of augmentation concepts with a cycling simulation. With further considerations on usability, Berge et al. (2022) found that HMDs may fulfill cyclists' need for recognition, however the cyclists may hesitate to use such devices. van Lopik et al. (2020) demonstrated that handheld devices provide higher perceived usability while AR glasses improve awareness of hazards.

To better counterbalance the drawbacks of AR based information presenting, suitable content length and delivery timing were studied by Uchiya and Futamura (2021). Regarding driving safety, HMD based information delivery systems were also considered as more competitive than classical head up displays (Langlois and Soualmi (2016)). Park et al. (2013) investigated the performance of a HMD which provides driving-safety information. MR approaches can also be both interactive and immersive for driving support, while efficiency and safety are preserved (Ghiurău et al. (2020); Ono et al. (2005); Matsunaga et al. (2018)).

Beside visual cues, a 3D audio based navigation system with wind noise reduction was proposed and evaluated by Kitagawa and Kondo (2018, 2019). Similar to Kitagawa and Kondo's studies, a bone conduction based navigation system for cyclists was developed by He and Zhao (2020).

Current approaches on AR based information delivery are mostly using fixed UIs on HUDs or taking advantage of auditory outputs. As UI textures being fixed in the field of view, they are inevitably blocking the user's vision and requiring the gaze move away from the forward. On the other hand, although the spatial audios are always delivered by bone conduction headsets, the user still have to process the sounds received from both their ears and bones at the same time, which are likely to cause distraction and confusion.

4.2.2 Hazard Detection

Computer vision based AR methods were widely accepted as efficient and low-cost solutions for detecting pedestrians, thus they were applied to vehicles: Early work by Gavrila (2000) applied a two-step machine learning algorithm on moving vehicles as a pedestrian avoiding support. On the basis of Garvrila's study, Shashua et al. (2004) developed a single-frame classification approach for day-time driving conditions. Hariyono et al. (2014) have improved the performance of moving vehicle pedestrian detection with a his-

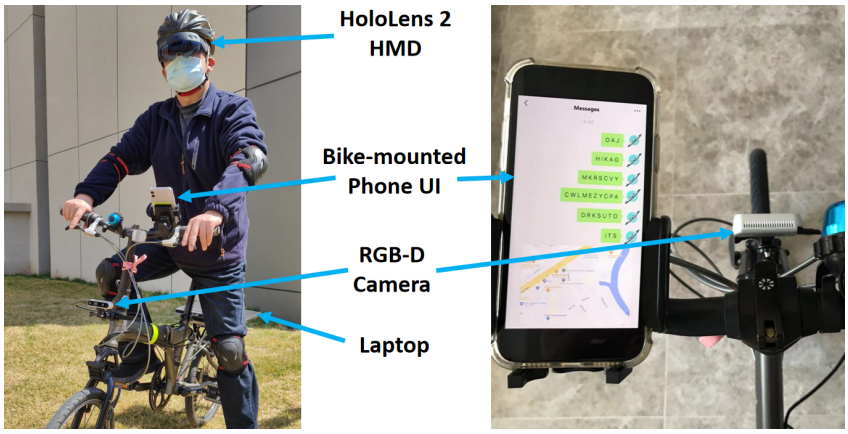

Figure 4.2: Images of the bicycle setup, including the HMD, bike-mounted phone UI, RGB-D camera, and laptop from a front (left image) and top (right image) view.

togram of oriented gradients. To date, DNN models have been applied as an advanced computer vision method for pedestrian detection on vehicles. Wang et al. (2018) demonstrated a DNN based pedestrian detecting system with semantic information of body parts and contextual information. In addition, a fast pedestrian detection system with efficient regions of interest generation and improved decision trees was proposed by Wang et al. (2019). Besides, the performance of a DNN-based pedestrian detection algorithm during night time was evaluated on an urban road by Nataprawira et al. (2020). Regarding the danger caused by low-quality road surfaces with obstacles, Rateke et al. (2019) proposed a convolutional neural network based road quality classifier. Labayrade et al. (2002) implemented a robust computer vision based method for obstacle detection on non flat roads. Peng et al. (2015) have demonstrated the high performance of their 2-D lidar based obstacle detection method.

Furthermore, considering the more interactive and immersive experiences that MR HMDs provided, other approaches were investigated such as that of Ghiurău et al. (2020), which extended the safety support system on Volvo cars to MR and improved users' awareness of upcoming hazards. Matsunaga et al. (2018) developed an MR based assistance system for welfare vehicle users and demonstrated the effectiveness of the system on collision avoidance.

4.3 Design and Implementation

In order to detect potential hazards in front of the bike and place the interface such that it smoothly snaps to each hazard, I developed two subsystems: a *vi-*

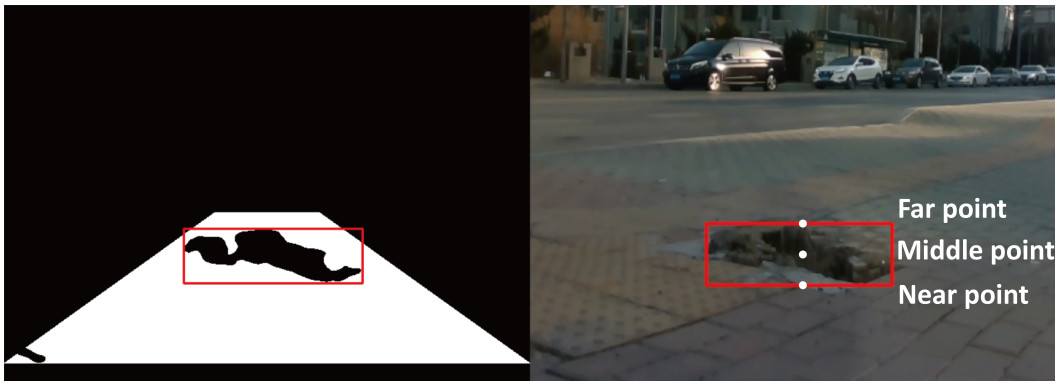

Figure 4.3: Images showing the detection process for a pothole using binary thresholding of the RGB-D camera with contour detection (left) and the resulting detected pothole with the three points indicating the coordinates used for specifying the distance calculation and resulting distance-based pothole detection.

sual recognition subsystem programmed with Python and an *interface delivery subsystem* programmed with C#. The visual recognition subsystem was implemented with an RGB-D camera and a backseat attached laptop, as shown on the left of (Fig. 4.2). The interface delivery subsystem was implemented using an AR HMD. These two subsystems exchange data using asynchronous TCP/IP socket communication through a smartphone WiFi hotspot. Algorithms in these two subsystems run in separate threads in order to ensure smooth parallel processing.

4.3.1 Equipment

In the implementation (Figure 4.2), I applied a Microsoft HoloLens 2 HMD, an Alienware m15 Laptop, an Intel Realsense D435i RGB-D camera, an iPhone 11 smartphone and a Dahon Jetstream P5 bicycle. The algorithms were driven by Unity 2021.1.20f1 and Python 3.9.

The HMD's FoV is 43∗29 degrees. The RGB-D camera has minimum 28cm depth distance, 30 fps frame rate, and 69∗42 degrees FoV.

4.3.2 Visual Recognition Subsystem

To start, I capture both the RGB and depth frames using an RGB-D camera fixed to the front of the bicycle, which faces directly forward with respect to the bike frame. Specifically, the RGB-D camera is fixed to the main frame of the bicycle rather than the handlebar in order to maintain a forward-looking

pose independent of small movements in the handlebar. I align these captured frames with the camera's application program interface and then apply two algorithms: a Caffe-based MobileNet-SSD detection network DNN model [2] for human detection and a customized algorithm for pothole detection. In addition, in the pothole algorithm I first calculate the peak gray value of the ground area using a histogram and then apply a binary threshold with a value of the peak value minus 30. Contours of probable potholes are gathered from the binary threshold results using a contour finder. To filter non-pothole results such as dark stains or manholes, I further compare the average depth of the far point and near point with the depth of the middle point (Fig. 4.3). If the difference is higher than 0.2 meters, the contour will be passed to the next step. With the gathered regions of interest (ROIs) and their vertex coordinates from both algorithms, I calculate the center pixel coordinate and depth of each ROI. In addition, the available depth range is 0.3 meters to 12 meters due to the RGB-D camera, but this covers a wide range of possible hazards and stopping distances (roughly 5 seconds of stopping time at 20km/h at 12 meters) . After gathering the vertex coordinates and center depth, the world coordinates (x, y, z) of each vertex are defined as:

$$z = centerdepth \tag{4.1}$$

$$x = (u - cx) * z/fx \tag{4.2}$$

$$y = -(v - cy) * z/fy \tag{4.3}$$

where the (u, v) is the pixel coordinate of each vertex, and the cx, cy, fx, fy are from the intrinsic matrix of the RGB-D camera. To convert the v value on the top down v axis in pixel coordinates to the y axis in projection coordinates, I calculate $-(v - cy)$. This set of equations are originally from the 2D-3D coordinate transformation models of SLAM. Since I only focus on the front view and the camera is always considered to be the origin in this study, the rotation and transformation matrices are not applied.

Once the world coordinates of the region of interest are defined, I pass them through an asynchronous TCP/IP socket running in another thread. In this socket, I set the subsystem to be the host and the sending interval is set to 0.05 seconds in order to mitigate jitter and packet sticking.

[2] https://github.com/chuanqi305/MobileNet-SSD/

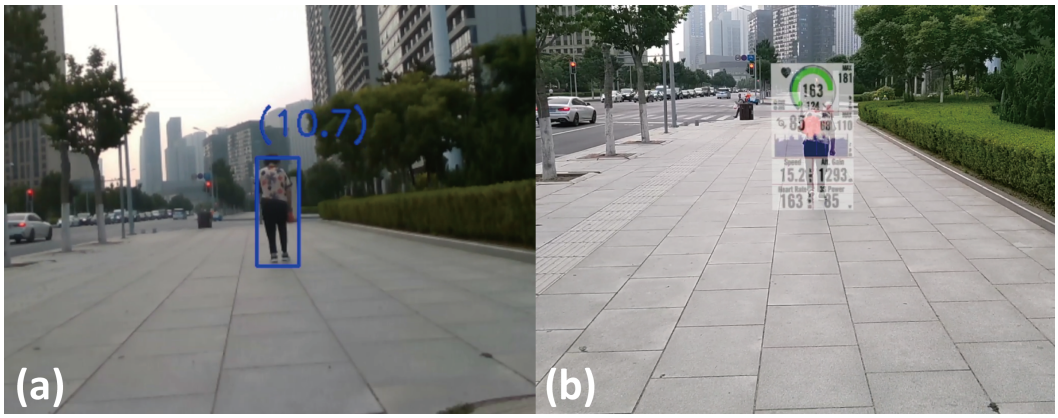

Figure 4.4: Images showing (a) the detection result from the visual recognition system with the depth value marked and (b) the UI snapped to an in-focus pedestrian (the red dot is a gaze indicator).

4.3.3 Interface Delivery Subsystem

This subsystem is developed in Unity and deployed to an HMD to afford an AR environment (Fig. 4.4). The data from the visual recognition subsystem is received via Socket client as messages and are cut based on a header in order to further avoid packets getting stuck. To compute the intersection of the HoloLens Gaze vector with hazards, transparent quads are generated in the AR environment based on the received world coordinates of detected pedestrians and potholes. An 8-meter long, 0.7-meter wide (corresponding to the bicycle width) transparent bicycle heading indicator is implemented in order to detect obstructions that stand in the way within a range of 4-12 meters.

I have also developed a method to compensate for the drift phenomenon that occurs in gyroscopes and other direction errors that result from the tilt motions that happen often while cycling. In my implementation, the bicycle's heading direction is calculated by taking the HMD's velocity vector in the MR environment, which is further corrected using the face-forward direction of the head with a 5-degree difference range. In simple terms, the heading direction will shift only when the HMD is moving in a similar direction to the face-forward direction of the head. With this justified heading direction, I are able to rotate the quads and heading indicator such that they are forward-fixed.

The height of the quads are set relative to the HMD height plus 0.3 meters in order to match the hazards' locations in-view. Gaze data is collected from the built-in eye tracker of the HMD, which allows the interface to be snapped to the quad of the hazard that the user is focusing on. In particular, as

the gaze ray comes close enough to a detected hazard (less than a 5-degree difference on the Y-axis), the UI's position will be updated to the hazard's position, providing a frame-to-frame transition from the current UI position to the next as the user's gaze follows the hazard. Considering that erratic or instantaneous motion of the UI is likely to cause confusion or distract the cyclist, a smoothing function is applied. In particular, the UI is set to move toward the hazard with a speed of 7m/s, which provides a smooth but not slow movement. To prevent variations in the scale of the text of the interface, which can degrade readability (Chen et al. (2004)), the scale of the interface is set to change relative to the distance between the gaze-hit position in order to keep the same scale in the view. In other words, the text will appear to remain the same size, but binocular cues will still give the cyclist proper depth judgement. With further logical testing, I defined the scale relationship as:

$$w = d * 0.12 \tag{4.4}$$

$$h = w * 1.7778 \tag{4.5}$$

where the d is the distance between the gaze-hit position and (w, h) are the width and height of the interface. The w-d ratio is defined by testing the minimum scale for displaying my UI content with the HMD's resolution and a ratio of 0.1 (w : d) is found as the just readable value. In order to ensure that the UI was fully readable in a variety of scenarios (especially my experiment), I defined the ratio as 0.12. The aspect ratio was fixed to 16:9 (1.7778) in order to match the aspect ratio of smartphone applications. Regarding situations in which the user is not focusing on any hazard or there is no hazard in a forward location, I provide an interface that is fixed in front of and 20 degrees below the heading direction (below the horizon from the user's viewpoint) in order to prevent blocking the user's forward vision and preserve accessibility (Orlosky et al. (2013)) at the same time.

4.4 System Evaluation

In this section, I describe two evaluations of the system's technical performance including an evaluation of detection accuracy and an evaluation of the system's process duration.

Figure 4.5: Images showing (a) the sidewalk where I conducted the user study and the three conditions in the experiment, including (b) the smartphone UI, (c) the forward-fixed UI, and (d) HazARdSnap.

4.4.1 Detection Accuracy

To set up this test, me and one of my collaborators biked in three environments for approximately 18 minutes in each environment, recording data for the duration of the trip. Then I gathered the every 10th frame from the data and set up two frame pools of 2000 frames (A.1, A.2) for testing human and pothole detection. From each frame pool, I picked the every second frame for testing in terms of true and false positives and negatives. For human detection, the result revealed 97.48% precision, 93% recall, and 2.4% false-positive rate. For pothole detection, the result revealed 96.27% precision, 82.6% recall, and 3.2% false-positive rate.

4.4.2 Process Duration

I tested the end-to-end latency of the system by printing timestamps in each step of the system process with 10 iterations and averaged the results. I found that the latency was on average 273.1ms from a hazard is detected until a quad is placed. The average latency of each step was: 82.6ms from detection until sending a socket message (with the 50ms sending interval), 186.5ms from sending a socket message until it is received, and 4ms from receiving a socket

message until a quad is placed.

4.5 Experiment

In the experiment, I compared HazARdSnap to a smartphone UI, as shown on the right of (Fig. 4.2), and a forward-fixed AR interface. Participants rode a bicycle in an open space, while attending to the UIS and avoiding virtual hazards. In addition, participants were asked to go through an MR track with assigned tasks. I examined their primary task performance (i.e., cycling and reading performance), head and gaze behavior and subjective preference with the three UI conditions. My primary hypotheses included:

- H1: HazARdSnap will result in fewer collisions with virtual objects while reading UI content.
- H2: HazARdSnap will have a significant effect on enhancing a participant's ability to read the information on the interface.
- H3: HazARdSnap will be less distracting and improve the user's feeling of safety.

4.5.1 Equipment

I used a Microsoft HoloLens 2 HMD, a Dahon Jetstream P5 bicycle, an iPhone 11 and a set of a cycling helmet with joint pads. I developed these subsystems separately, which were driven by Python 3.9 and Unity 2021.1.20f1.

4.5.2 Experimental Design

In order to preserve safety in bicycle related works, researchers have conducted experiments with virtual reality environments and fixed bicycles (Oliveira et al. (2018); Liang et al. (2021)). However, this approach of user testing may cause motion sickness and could miss some motions triggered by losing body balance. To provide a better tradeoff, my experiment was implemented by letting participants actually ride a bicycle in an open and flat space with virtual pedestrians, potholes and a track delivered by the MR HMD. Since the user study was conducted as a directly forward scenario and no direction adjustment was needed, I did not apply the RGB-D camera and laptop in order to mitigate the load on the bicycle.

The user study used a within-subjects design to test HazARdSnap (Fig. 4.5-d) against two other conditions: An interface delivered by a smartphone mounted on the left side of the handwheel (Fig. 4.5-b) and an interface which

Figure 4.6: Images showing (a) a view of the MR environment in the real world recorded through the HoloLens 2's streaming function and examples of a (b) virtual pedestrian and (c) virtual pothole from Unity.

was fixed front-below the forward direction (same as the situation that the user is not focusing on any hazard or there is no hazard located forward of my proposed approach) (Fig. 4.5-c). The order of conditions was assigned based on a balanced Latin-Square in order to minimize order effects.

Considering the possibility that participants may remember the locations of the hazards and the content on the interface, six different experimental environments and six different content on the interface were randomly assigned as each trial start. For each trial, participants were required to ride forward while avoiding the virtual hazards and boundaries in the experimental pathway. Furthermore, participants were asked to read out the content presented via UI condition and complete the UI reading task before reaching the finish line. Detailed descriptions about the hazards and content can be found in Section 4.5.3. Raw data (e.g. collisions, speed, focus, and distance) were recorded during each trial in order to help us to understand the accessibility of each method and how participants' riding behavior were affected.

4.5.3 Experimental Environment

To test HazARdSnap and the AR UI, I utilized a system that made use of virtual humans and potholes to ensure that participants would not be at risk of crashing into real hazards. Instead of detecting real objects, I adapted my method to react to the virtual objects within a Unity environment for the experiment. In the MR environment, I rendered 8 pedestrians, 4 pot-

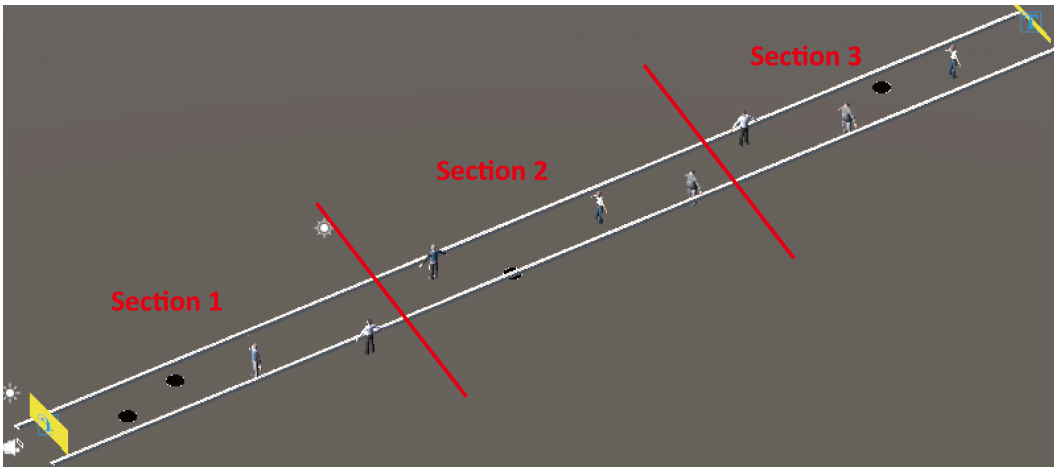

Figure 4.7: Overview of the designed virtual environment in the Unity Inspector, scaled to 50 meters long and 2.5 meters wide. This was segmented into 3 sections with 4 objects in each section.

holes and 2 boundaries (one on each side) where the size of these objects were set according to real-world sizes (Fig. 4.6). In order to simulate real pedestrian behaviours, I set one pedestrian across the track repeatedly, one pedestrian across the track suddenly while approaching and one pedestrian moves suddenly while approaching. Note that the version of HazARdSnap in the experiment did not include the smooth movement from one object to the next, which developed during further refinement.

In the virtual environment, the track was set to be 50 meters long and 2.5 meters wide to increase the likelihood that participants would finish the UI reading task and the objects would not fully block the way. To alleviate ordering effects between trials while ensuring similar difficulty, I segmented the environment into three sections, and the arrangement of these sections were randomized to six experimental environments in total (Fig. 4.7). Regarding the reading content, I randomly typed 35 letters for each UI instance (Fig. 4.8). Considering the performance of hologram displays can be highly affected by day light conditions, I conducted my user study in the same place and same time of day (5pm to 7pm) (Fig. 4.5-a).

4.5.4 Participants

24 naive individuals (9 female, mean age 32.29, SD 7.17, range 13-52), were recruited from the general population to take part in the study. All participants had normal or corrected-to-normal vision, little to no experience with AR and were able to ride bicycles. They were informed that they were al-

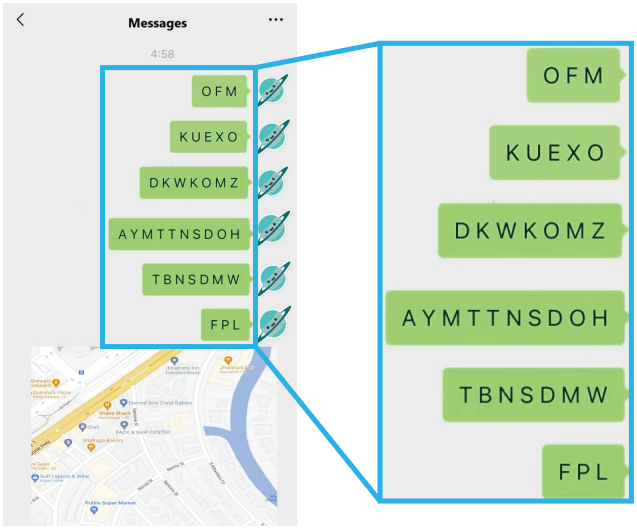

Figure 4.8: An example of the content that participants had to read while cycling during the user study.

lowed to quit the study at any time. I set the bicycle seat to a relatively low height and added provided protective gears (i.e. the helmet and joint pads as shown in Fig. 4.2) in order to protect the participants in the unlikely event that they fell down. The experiment was conducted under approval of the Osaka University Cybermedia Center Institutional Review Board (IRB), number SA2022-02.

4.5.5 Dependent Measures

During the study, I logged the data necessary to calculate a comprehensive set of dependent measures as follows:

Number of collisions (i.e. hit an obstacle or ran out of boundaries): I collected collision data using a Unity collider that moves with the HMD. With the HMD's built-in eye tracker, I gathered gaze focus data for further measurements.

Average cycling speed in m/s: I divided the MR track length by completion time.

Average cycling speed while reading (i.e. while focusing on UI) in m/s: I took the average cycling speed when the participants were focusing on the interfaces.

Reading time (total time focused on the UI): I recorded the total time that the participants were focused on the interfaces.

Reading-to-completion time ratio: I divided the reading time by trial com-

pletion time.

Average angular speed of head rotations and *average angular speed of gaze* in degrees/s: I averaged the differences in angles between frames, measured in degrees per second.

Average angle difference between head direction and face-forward (i.e., head deviation from face-forward) and *average angle differences between gaze direction and face-forward* (i.e., eye gaze deviation from face-forward) in degrees: I first collected the head facing direction and gaze direction data from the HMD and then calculated the average angle between the direction data with a forward vector (0,0,1).

Post-trial subjective ratings: participants were asked to rate the following subjective preference questions with 7-scale ratings from 1 ("Strongly disagree") to 7 ("Strongly agree"):

- I had a clear perception of the pedestrians' and potholes' positions while looking at the UI.
- I felt safe while riding.
- Reading content on the interface disturbed me.
- I could read the content easily and clearly.

In order to estimate time of focus data from the smartphone UI condition, I recorded the time periods when the gaze is pointing the left-down direction and has a degree to forward above 20 degrees.

4.5.6 Procedure

As each participant joined my experiment, I briefly introduced the experiment, and I informed the participants about the tasks and requirements. Then, the participant put on the protective gear and the HMD with help from an experimenter. Before the first trial started, the participants were asked to complete an eye calibration using the calibration tool bundled with the HoloLens 2. Before each trial was conducted, I carefully checked the clearance of the physical experimental space. After each trial, the experimental environment was reset and the participant was asked to fill in a questionnaire.

4.5.7 Results

Here, I describe the results of my experiment with respect to the three UI conditions, differences on accessibility, and subjective preference of the participants. For parametric data, I applied ANOVA tests and Tukey's HSD post-hoc tests. On the other hand, Friedman tests and Wilcoxon post-hoc tests

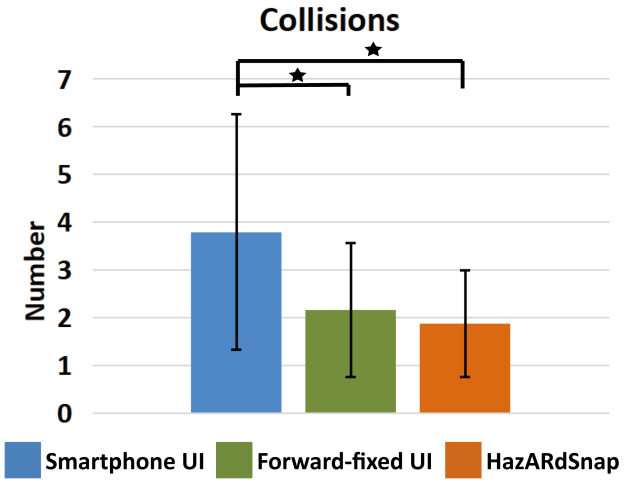

Figure 4.9: Analysis of number of collisions showed that smartphone UI was associated with significantly more collisions than forward-fixed UI and HazARdSnap.

were applied for non-parametric data. In addition, Cohen's d and Kendall's W values were included in my results to define effect sizes. I used a level of 0.05 for determining significance.

4.5.7.1 Task Performance

For the cycling task, regarding number of collisions, ANOVA test revealed significance between the three conditions ($F(2, 69) = 8.26$, $p < 0.001$) (Fig. 4.9). Further differences were found in the smartphone UI vs. the forward-fixed UI ($p < 0.01$, $d = 0.81$) and the smartphone UI vs. the HazARdSnap ($p < 0.001$, $d = 1.00$), indicating that participants were more likely to collide while reading the smartphone UI. Regarding average cycling speed, no significant difference was found ($F(2, 69) = 1.20$, $p = 0.31$) (Fig. 4.10-a). Regarding average speed while reading, significance was found ($F(2, 69) = 5.62$, $p < 0.01$) (Fig. 4.10-b) and post-hoc tests show that there was a significant effect in the smartphone UI vs. the forward-fixed UI ($p < 0.01$, $d = 0.88$), which showing that participants rode faster with the forward-fixed UI.

For the UI reading task, regarding reading time, the data revealed no significance ($F(2, 69) = 2.39$, $p = 0.10$) (Fig. 4.11-a). Regarding reading-clear time ratio, no significant difference appeared ($F(2, 69) = 1.99$, $p = 0.15$) (Fig. 4.11-b).

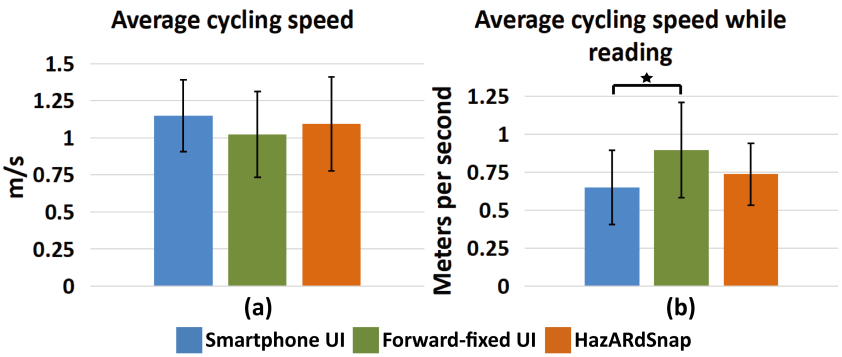

Figure 4.10: Analysis of (a) speed and (b) average speed while reading. Although there was no significant difference within cycling speed, participants tend to ride faster while reading via forward-fixed UI.

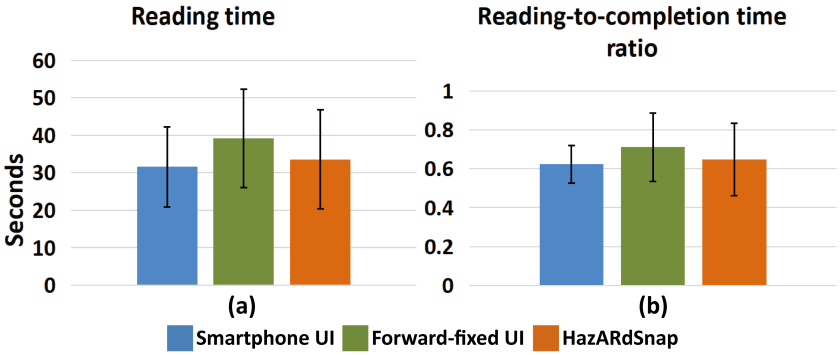

Figure 4.11: Analysis of (a) reading time and (b) reading-clear time ratio. Regarding the UI reading task, these two metrics revealed that reading speed was not affected by conditions.

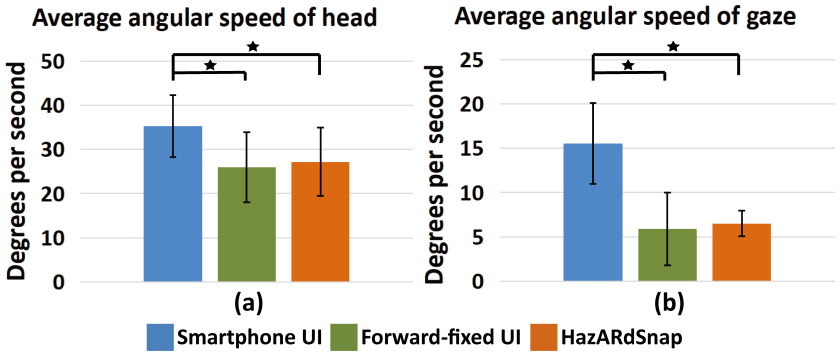

Figure 4.12: Analysis of (a) head's average angular speed and (b) gaze's average angular speed. These showed that there are significant differences in the rotating frequency of both head and gaze.

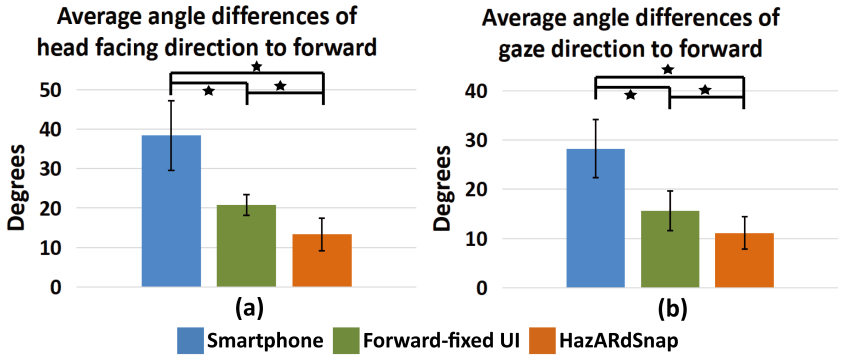

Figure 4.13: Analysis of (a) the average angle difference between head direction and face-forward and (b) average angle differences between gaze direction and face-forward. These two metrics roughly represent the average time spent viewing the road.

4.5.7.2 Head and Gaze Behavior

To determine whether the augmented UI delivering methods affected participants' head and gaze rotation, ANOVA tests were applied to reveal that there were significant differences in both of the head's average angular speed data ($F(2, 69) = 9.92$, $p < 0.001$) and the gaze's average angular speed data ($F(2, 69) = 52.81$, $p < 0.001$). Further Tukey's HSD tests showed that for head's average angular speed (Fig. 4.12-a), there were significant differences in the smartphone UI vs. the forward-fixed UI ($p < 0.001$, $d = 1.24$) and the smartphone UI vs. the HazARdSnap ($p < 0.01$, $d = 1.10$) and for gaze's average angular speed (Fig. 4.12-b), there were significant differences in the smartphone UI vs. the forward-fixed UI ($p < 0.001$, $d = 2.23$) and the smartphone UI vs. the HazARdSnap ($p < 0.001$, $d = 2.67$), indicating participants, on average, moved their heads and gazes more frequently when using the smartphone UI as compared to both the Forward-fixed and HazARsnap UI. In addition, significant differences were also found in average angle of head facing direction to forward ($F(2, 69) = 117$, $p < 0.001$) and average angle of gaze direction to forward ($F(2, 69) = 91.74$, $p < 0.001$). For average angle of head facing direction to forward (Fig. 4.13-a), significant differences were located in the smartphone UI vs. the forward-fixed UI ($p < 0.001$, $d = 2.69$), the smartphone UI vs. the HazARdSnap ($p < 0.001$, $d = 3.64$) and the forward-fixed UI vs. the HazARdSnap ($p < 0.001$, $d = 2.17$). For average angle of gaze direction to forward (Fig. 4.13-b), significant differences were located in the smartphone UI vs. the forward-fixed UI ($p < 0.001$, $d = 2.50$), the smartphone UI vs. the HazARdSnap ($p < 0.001$, $d = 3.59$) and the forward-fixed

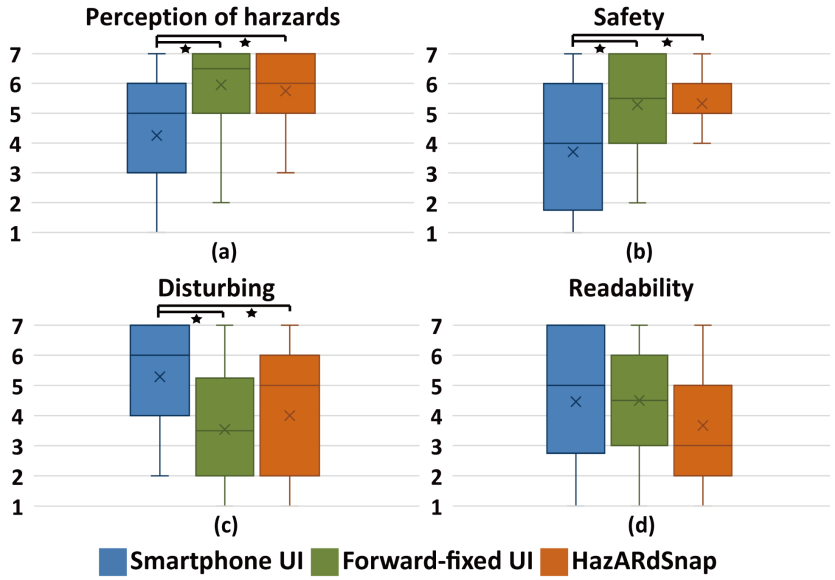

Figure 4.14: Analysis of subjective ratings for (a) perception of hazards, (b) safety, (c) disturbing, and (d) readability. Participants were more likely to feel safer and have clearer views as forward-fixed UI and HazARdSnap were applied.

UI vs. the HazARdSnap ($p < 0.01$, $d = 1.22$). These results revealed that participants averagely focused more on the forward direction when using the HazARdSnap.

4.5.7.3 Subjective Preference

I analyzed data from post-trial questionnaires with Friedman tests, where significant differences were found between the three UI conditions for "perception of hazards" ($\chi^2 = 12.61$, $p < 0.01$), "safety" ($\chi^2 = 16.22$, $p < 0.001$) and "disturbing" ($\chi^2 = 11.63$, $p < 0.01$). Secondly, Wilcoxon tests were applied as paired tests: For "perception of hazards" (Fig. 4.14-a), significant differences were found in the smartphone UI vs. the forward-fixed UI ($V = 5$, $p < 0.01$, $W = 0.56$) and the smartphone UI vs. the HazARdSnap ($V = 5$, $p < 0.01$, $W = 0.64$), which means that participants subjectively perceived less hazards with the smartphone UI applied. For "safety" (Fig. 4.14-b), significant differences were located in the smartphone UI vs. the forward-fixed UI ($V = 30$, $p < 0.01$, $W = 0.57$) and the smartphone UI vs. the HazARdSnap ($V = 5$, $p < 0.001$, $W = 0.73$), indicating participants felt less safe when reading the smartphone UI. For "disturbing" (Fig. 4.14-c), significant differences were located in the smartphone UI vs. the forward-fixed UI ($V = 157$, $p < 0.01$, W

= 0.65) and the smartphone UI vs. the HazARdSnap ($V = 140.5$, $p < 0.05$, $W = 0.67$), showing that the smartphone UI was considered more disturbing as compared with the forward-fixed UI and HazARdSnap. Besides, no significance was observed in the result of "readability" ($\chi^2 = 0.94$, $p = 0.63$) (Fig. 4.14-d).

4.6 Discussion

Based on the collisions data, participants encountered fewer collisions when using the forward-fixed UI (43%) and HazARdSnap (51%) as compared to the smartphone UI, suggesting that delivering UIs in a head-up fashion (e.g., via MR HMDs) is safer than using handlebar mounted smartphones. In addition, the head and gaze behavior revealed how AR based methods helped preventing dangers. Participants performed lower values on the head's and gaze's average angular speed data when using the forward-fixed UI and HazARdSnap. Which suggested that with the AR based methods, the participants were able to rotate their head and gaze in a less frequency and gain a longer period of stable view. Furthermore, the average angle of head and gaze direction to forward data showed significant lower values when using the forward-fixed UI and HazARdSnap. This indicated that while reading UI content via an MR HMD, the participants were allowed to look forward more often and gain a higher chance to notice front hazards. On the other hand, the average speed of riding did not show any difference between the three conditions which indicates that interface delivery methods are not likely to effect the user's preferred speed of riding. However, if the hazards were real objects, there may be differences since real crashes may cause the participants to be nervous and slow down. In conclusion, my H1 was supported by the results from my user study.

Although significant differences did not exist for most metrics between the two AR-based UIs, HazARdSnap did help participants to outperform in both average angle of head facing direction and gaze direction to forward. Fixed UIs have a limitation that they can not be placed straight forward in the middle of the view, since they will block the view. According to the demonstration by Orlosky et al. (2013), they was also preferred to be placed at a forward-down position by users. Therefore, dangers may happen with this kind of methods when cycling toward to upper located hazards such as hang-up wires or lift-up arms of pedestrians. On the other hand, keeping the view straight forward allows cyclists to better gather information about signboards, traffic signals and routes, as shown by Orlosky et al. (2014).

Regarding H2, my analysis of average cycling speed while reading data revealed that while reading the UI content, participants were able to ride faster using forward-fixed UI than the smartphone UI. This result was reasonable since looking down at the smartphone causes a huge loss of front view, therefore participants are likely to slow down in order to prevent a crash. On the other hand, the HazARdSnap did not show any significance with this data. I believe this was caused by the tendency of the UI to snap or "jump" from one object to another since several participants orally reported that they thought the snapping behavior made it difficult to find the position of the UI while cycling with HazARdSnap. Therefore when the UI was snapped, they were more likely to slow down and read. As this situation was noticed during the experiment phase, I applied a smoothing algorithm to the real implementation. Besides, the reading-to-completion time ratio data did not reveal any significance between the three conditions. However, this result demonstrated that content delivered by MR HMDs would not have a negative effect on the user's reading speed, as long as the UI was delivered in a constant and readable scale. Moreover, the reading time data did not reveal any significance between the conditions, therefore I believe that different interface delivery approaches will not effect the participants' reading performance as long as the content are with similar resolution. In addition, the subjective ratings of "Readability" also revealed no significance between the three conditions which aligned with the finding by the objective measure.

In H3, I hypothesized that my HazARdSnap may positively affect the participants' subjective feelings of safety since my approach allows them to reading while looking at a hazard at the same time. However, although both of the AR based methods were more preferred than the smartphone UI in the subjective rating results of "Perception of hazards" and "Safety", there was no significant difference founded between them. I believe that this was because both the AR based methods did not require high frequency head rotations and the stability of their views were preserved. Furthermore, looking at the smartphone was considered as more disturbing by the participants. The AR based methods were more effective on helping the participants to gain a stronger feeling of stability which leads to a feeling of safety.

In the experiment, similar to the finding by Berge et al. (2022), I have noticed that most of the participants hesitated to use the MR HMD. This situation may slightly effect the results, especially that they have to read and ride at the same time. Therefore I suggest that future works should focus more on counter balancing this situation when conducting experiments.

4.7 Chapter Conclusion and Future Work

In this work, I proposed a novel augmented reality-based information delivery system designed to help cyclists read UI content in a safe and easy manner. To accomplish this, I integrated real-time forward hazard detection and gaze tracking so that information can be snapped to a hazard that is in focus, which allows the cyclist to both read UI content and perceive the locations of hazards. I finally conducted an MR based within-subjects user study with 24 participants to study how HazARdSnap and a forward-fixed UI would perform in comparison to a traditional bike-mounted phone interface. Random text were delivered to participants via each UI, including a handlebar-mounted smartphone, a forward-fixed UI and my proposed HazARdSnap approach. Results from this study show that both the forward-fixed UI and my proposed approach effectively reduced the number of collisions with hazards in comparison to the smartphone UI. Moreover, the AR methods enabled users to maintain higher stability of head and gaze without a significant reduction of reading speed.

In addition, since the participants were not familiar with MR HMDs, the systems may prove to be more effective when the MR devices are more common and broadly accepted by users in the future. In contrast to the popular virtual reality based user tests on driving and cycling, my experimental setup is also a novel way to safely test AR based interfaces and provide more realistic but safe user tests.

One current limitation of my work is that multiple elements of information cannot yet be distributed amongst multiple hazards in the environment. One potential extension of this work is to allocate various spaces, such as the backs of multiple cyclists in a cycling team, for different types of information such as speed or biometrics. This might allow for easier readability without sacrificing the amount of information available. my interface may also benefit from eye-based engagement or control to satisfy individual user preferences for information display.

In future studies, I plan to further explore wearable UI delivery approaches that can better help cyclists handle multiple tasks at the same time while preserving safety. I also plan to conduct experiments testing user interactions with UIs while cycling. Moreover, I would like to improve the proposed approach and investigate its effectiveness with more types of hazards such as vehicles, trees and/or roadblocks.

I hope that this work will encourage the development of other AR based UI delivery techniques for cyclists that can improve riding safety and maintain

ideal readability as well as the exploration of augmented reality-based experimental designs that allow for safer testing of collision avoidance systems.

Chapter 5

Peripheral Indicators for Rear-Approaching Vehicles

5.1 Introduction

In the last chapter, a gaze-based real-time rendering approach is presented. In order to explore other visual rendering for cyclists, I developed a method that take advantage of the periphery of human eyes. Cyclists often ride through traffic or with a group of other cyclists, so they must pay attention to both the front and rear depending on the situation. Collisions are a big problem, especially those that come from behind when merging lanes. To detect hazards, cyclists typically make use of rear-view mirrors or monitors, which can be mounted to a bicycle's stem, handlebar, or bar-ends. Although current rear-view methods are able to help the cyclists to avoid rear-approaching vehicles, the act of looking at the mirror/monitor itself can be a distraction, which is similar to the distractions caused by bike-mounted cell phones (Ren et al. (2021); De Waard et al. (2014); Terzano (2013)). In addition, cyclists must rotate their heads frequently to gain information from their forward view and the rear-view to reduce their speed as necessary and avoid collisions (Chong et al. (2010); Vansteenkiste et al. (2013)). To alleviate the danger of rear collisions and distractions that come from looking down at a mirror, monitor, or other rear-view interface, I designed ReAR Indicators as an information delivery approach for improved access to forward and rear information that takes advantage of augmented peripheral visualization, as outlined in Fig. 5.1.

Head mounted displays have commonly been used as AR devices to provide more accessible information with augmented vision. Considering that cyclists always wear helmets while cycling, HMDs can be an excellent candidate for cycling since they can easily be integrated into helmets. In addition, AR HMDs have the potential to alleviate the problem which cyclists have to look down frequently and AR approaches are widely implemented as driving and hazard detecting assistance (Langlois and Soualmi (2016); Kimura et al. (2017)). However, information delivered by AR HMDs may also cause distractions since it is always overlaid in the cyclist's field of view (FoV) (Gabbard

et al. (2014)). Peripheral cues have the potential to alleviate distractions since information can be delivered without blocking the foveal vision thus they are widely applied in driving and rear collision avoidance AR approaches (Niforatos et al. (2017); Kunze et al. (2019); Hecht et al. (2022)).

In this work, I introduce a novel approach called "ReAR Indicator" that not only provides peripheral indications of rear-located vehicles (e.g. bicycles, motorbikes and cars) but also preserves attention to the front as shown in Fig. 5.1. In addition, unlike other existing peripheral indicators, this approach provides indications across the visual areas and can reach to the focal area to draw attention when danger is going to happen. First, to determine the location and width of vehicles, I gather scene data with an RGB-D camera and process scenery behind the bike through a set of computer vision based algorithms. This data is sent to a HoloLens 2 via TCP/IP socket communication. Then, scene data, motion, facing direction, and head position are gathered from the HMD. Using this data as input, I designed an algorithm that detects rear-located vehicles and projects relevant information about them to the periphery. This allows simultaneous visual access to the information and awareness of the vehicle's location and likelihood of a collision. Simply put, I provide better rear information access to cyclists without compromising their attention to the forward. In an experiment using real bicycles with virtual scenes (for participant safety), I evaluated ReAR Indicator, a front-placed virtual monitor, and a 3D arrow interface delivered by an HMD, in order to investigate the following hypotheses:

- H1: Cyclists using the ReAR Indicator will react faster and have fewer collisions with virtual vehicles.
- H2: Reactions to forward information will be faster with the ReAR Indicator.
- H3: ReAR Indicator will be less distracting and improve the user's feeling of safety.

In the following sections, I discuss related work, detail the design process and implementation of the ReAR Indicator, describe my user study, and provide an analysis of the results. I conclude with a discussion of my findings and observations.

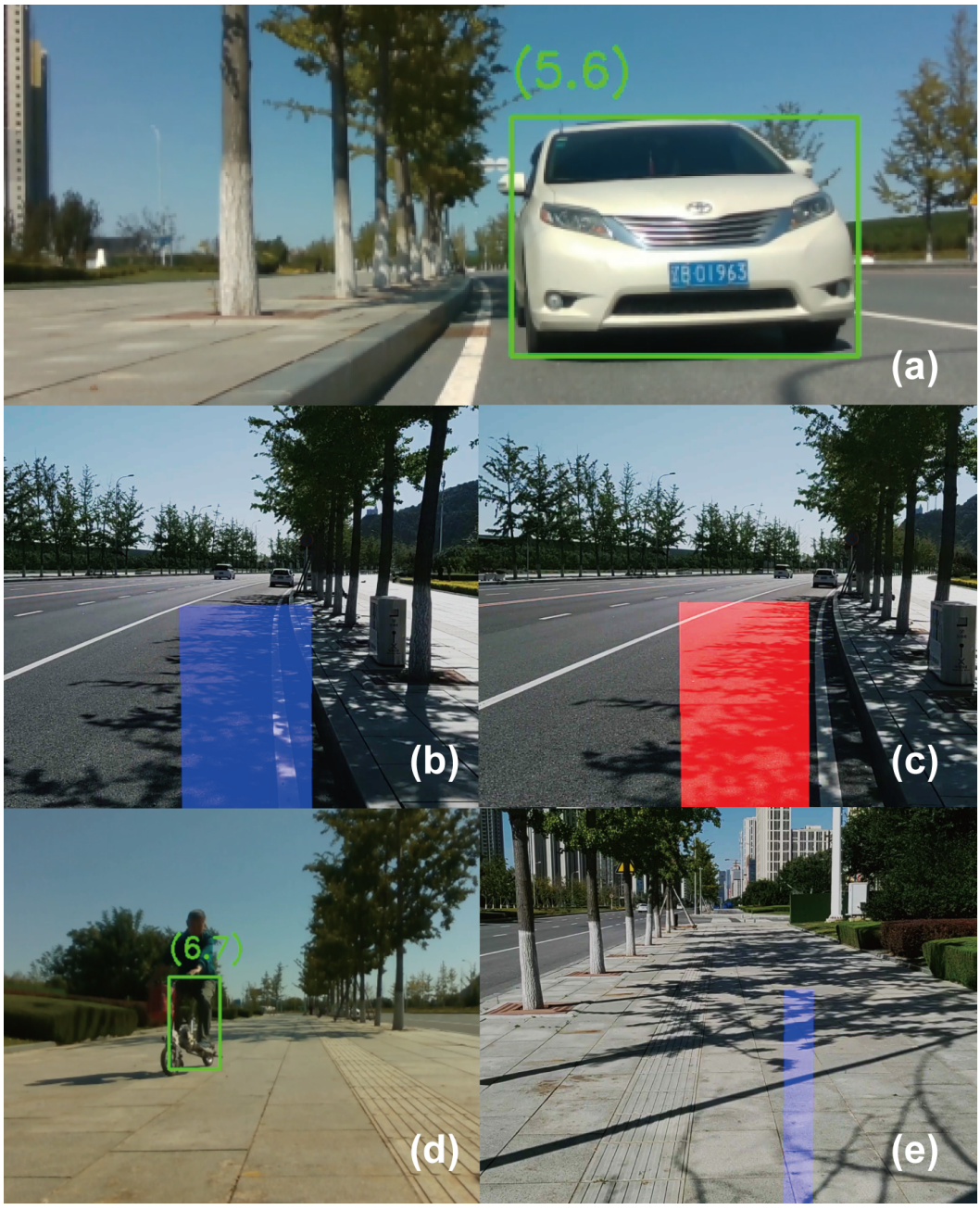

Figure 5.1: Images showing (a) real-time detection of the vehicle using a bike-mounted rear RGB-D camera, (b) the ReAR Indicator, which uses a blue bar to indicate a vehicle approaching from the rear, (c) a red bar that indicates a potential crash, (d) real-time detection of a cyclist following the user, and (e) ReAR when used to visualize other cyclists using a smaller blue indicator.

5.2 Related Work

While cycling on urban roads, cyclists can often have the potential of getting hit on their side or back, especially when focusing on their forward view or checking their bike-mounted interfaces (e.g. smartphones or cyclocomputers). To help address these problems, many system exist to help deliver easily accessible information and hazard alerts to cyclists. In addition, peripheral indications have been proposed by researchers as a non-distracting approach, which is likely useful for my cycling use case. In this section, I discuss previously proposed AR approaches related to cycling information delivery and peripheral indicators.

5.2.1 Information Delivery for Cycling Safety

Cycling through traffic is a challenging situation in which cyclists must often navigate through narrow spaces between cars or other bicycles. This often requires high attention to the forward direction, frequent steering, and low speed (Vansteenkiste et al. (2013)). Moreover, using supporting applications such as navigation assistants on a mobile phone or other type of display often causes distractions (Ren et al. (2021)) to the cyclist. To solve this problem, Dancu et al. (2015) evaluated their projection-based interface for urban cycling with a head-up display and von Sawitzky et al. (2020) investigated the effectiveness of their HUD (Head-Up Display) based augmentation concepts for cycling safely. With further considerations on usability, HMDs were proposed as a viable solution: Berge et al. (2022) found that HMDs may fulfill cyclists' need for recognition, however the cyclists may hesitate to use such devices. van Lopik et al. (2020) demonstrated that handheld devices provide higher perceived usability while AR glasses improve awareness of hazards. Ginters (2019) developed an AR glasses based interface for providing navigation and other information.

Besides visual cues, audio cues are also considered as an efficient solution. Kitagawa and Kondo proposed and evaluated a 3D audio based navigation system with wind noise reduction (Kitagawa and Kondo (2018, 2019)). Similar to Kitagawa and Kondo's studies, a bone conduction based navigation system for cyclists was developed by He and Zhao (2020). For detecting approaching vehicles and informing cyclists with sounds, Jeon and Rajamani (2017) proposed a low-cost rear-vehicle detecting system using a single beam laser sensor mounted on a rotationally controlled platform. Moreover, in the work by Schoop et al. (2018), they detect vehicles using machine learning algorithms and provide directional audios as cues.

Current approaches on augmented visions are mostly using fixed UIs delivered by HUDs or HMDs. As UI textures are fixed in the field of view, they are inevitably blocking the user's vision and requiring the gaze to move away from the forward. On the other hand, implementing auditory alerts may cause distraction and confusion since cyclists often judge the approaching vehicles' direction and distance by hearing.

5.2.2 Peripheral Indication

Augmenting peripheral vision has the advantage that the foveal vision will not be affected and distraction to the user will be lessened, so prior studies have been conducted on improving awareness and perception. Gruenefeld et al. (2018) implemented a front collision warning system for pedestrians using LEDs that are attached to ordinary eyeglasses. Regarding backward collisions, Niforatos et al. (2017) applied LEDs and sensors to a helmet in order to warn the user about approaching skiers. By applying LEDs to HMDs, Qian et al. (2018) proposed an approach for enhancing perception of peripheral vision which is likely to be blocked by the HMD. Moreover, by combining HMDs and LED light bars, Jones et al. (2013) proposed a displaying method that improves spatial perception in virtual environments.

In addition to LED based approaches, AR displays are also used as a method for delivering peripheral indications: Fan et al. (2014) developed a system for extending back view by indicating backward motions on the peripheral area of an HMD's FoV. Janaka et al. (2022) designed an interface that is located on the near-peripheral area and allows the user to access information while keeping eye contact with another person. However, considering the limited FoV that current HMDs have, the range of peripheral area is restricted. Therefore, applying peripheral indications is challenging as they have to counterbalance accessibility and rendering range. In order to provide a proper solution, Chaturvedi et al. (2019) came up with a peripheral navigator using blue color which was inspired by Murch's findings (Murch (1984)).

5.3 Design and Implementation

In order to detect vehicles behind the bicycle and provide a peripheral indicator that simultaneously represents the width, distance and likelihood of a collision of each vehicle, I first programmed a visual detection algorithm with Python and deployed it on an RGB-D camera and bike-mounted laptop (Fig. 5.2). Then, I designed and developed the peripheral indicator using

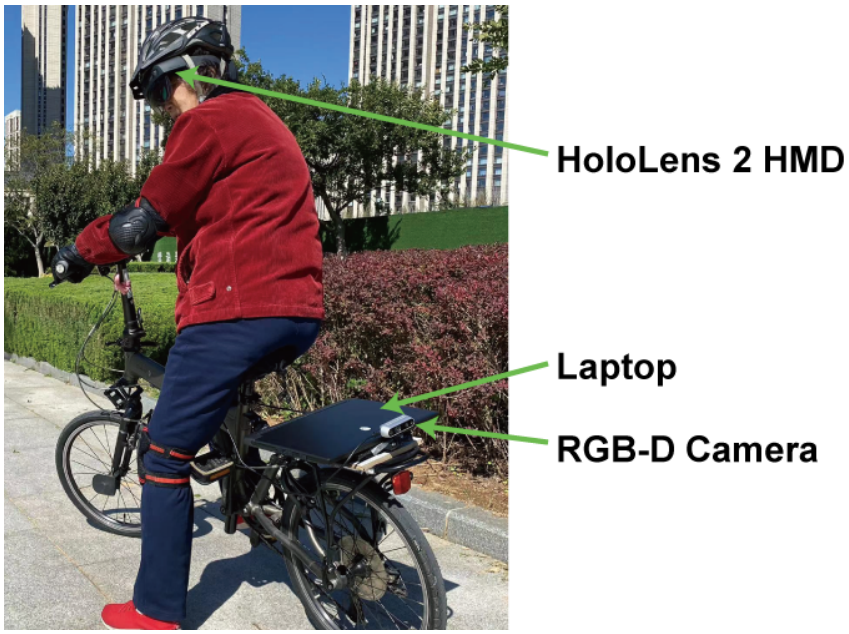

Figure 5.2: Image of the bicycle setup, including the HMD, RGB-D camera, and laptop from a rear view.

C# and HLSL and deployed it on the HoloLens 2. Asynchronous TCP/IP socket communication was applied for data exchange using a smartphone WiFi hotspot. Specific details regarding the recognition algorithm, indicator parameters, and hardware are detailed as follows.

5.3.1 Equipment

In the implementation (Fig. 4.2), I applied a Microsoft HoloLens 2 HMD, an Alienware m15 Laptop, an Intel Realsense D435i RGB-D camera, an iPhone 11 smartphone and a Dahon Jetstream P5 bicycle. The algorithms were driven by Unity 2021.1.20f1 and Python 3.9. The HMD's FoV is 43∗29 degrees. The RGB-D camera has minimum 28cm depth distance, 30 fps frame rate, and 69∗42 degrees FoV.

5.3.2 Visual Recognition

For the recognition system, I needed to recognize when a car was behind the bike so that I could provide details about the car via an indicator that would be useful for the rider. To do so, I capture both RGB and depth frames using an RGB-D camera fixed to the back seat of the bicycle, which faces directly backward with respect to the bike frame. I first align these captured RGB

frames with the depth frames via the camera's application program interface (API). Next, I apply a Caffe based MobileNet-SSD DNN model [1] for car and bicycle detection. From this DNN, I can obtain the ROIs, including vertex coordinates, and I calculate the depth of each ROI (Fig. 5.3-a). The available depth range is 0.3 meters to 20 meters due to the RGB-D camera, which covers a relatively wide range of vehicles and reaction times. For example, for an urban speed limit of 60km/h with a cyclist riding at an average speed of 20km/h, the reaction time is approximately two seconds, a similar allowance to other state-of-the-art notification systems (Schoop et al. (2018)). After obtaining the vertex coordinates and depth, the world coordinates (x, y, z) of each vertex are defined as:

$$z = depth \tag{5.1}$$

$$x = (u - cx) * z/fx \tag{5.2}$$

$$y = -(v - cy) * z/fy \tag{5.3}$$

where (u, v) is the pixel coordinate of each vertex, and cx, cy, fx, fy are from the intrinsic matrix of the RGB-D camera. To convert the v value on the top down v axis in pixel coordinates to the y axis in projection coordinates, I calculate $-(v - cy)$. This set of equations are originally from the 2D-3D coordinate transformation models of SLAM. Since I only focus on the front view, and the camera is always considered to be the origin in this study, the rotation and transformation matrices are not applied. Further, I rotate the world coordinates to the rear by converting x to $-x$ and z to $-z$.

Once the world coordinates of the ROIs are defined, I pass them through an asynchronous TCP/IP socket running in another thread. In this socket, I set the laptop to be the host with a sending interval set to 0.05 seconds in order to mitigate jitter and packet sticking of the network.

5.3.3 Peripheral Indication

My peripheral indicator is developed in Unity and deployed on the HoloLens 2. Moreover, the indicator is driven by a shader script that is applied to a camera canvas in the environment. The visual recognition data is received via

[1] https://github.com/chuanqi305/MobileNet-SSD/

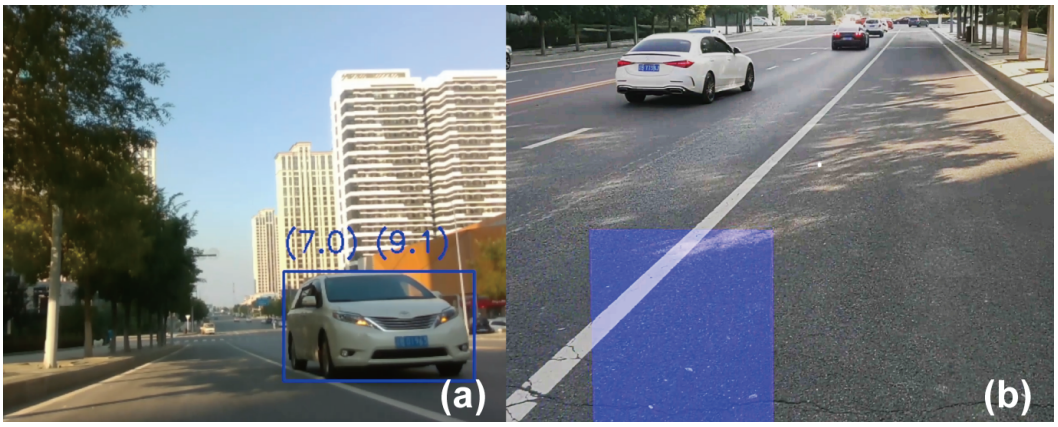

Figure 5.3: Images showing (a) detection results from the visual recognition algorithm (b) and the corresponding ReAR peripheral indicator that represents the approaching car's position, width, and speed.

Socket client as messages and are cut based on a header in order to further avoid packets getting stuck. From the data, world coordinates were projected to a u-v coordinate in a customized Unity shader. The scale relation was defined as:

$$ul = -xl * 0.1 + 0.5 \tag{5.4}$$

$$ur = -xr * 0.1 + 0.5 \tag{5.5}$$

$$vmax = 0.5 - z * 0.0125 \tag{5.6}$$

where the xl and xr are the X-axis values of the left and right vertex and the ul and ur are the projected u-axis values of xl and xr. In addition, the xl and xr are flipped to make ul and ur match the real direction and the indicator is rendered between ul and ur. The z is the depth but also represents the distance between the detected vehicle and the bicycle and the $vmax$ is the maximum height for rendering the indicator. Regarding the scale ratio of z, since the shader is designed to match the HMD's aspect ratio, I defined it as: 0.5/2/20.

With the computed values, I render a bar-shape indicator (Fig. 5.3-b), where width represents the width of the backward-located vehicle and height represents the distance between the bicycle and the vehicle. In addition, if a vehicle is just behind the bicycle, the bar will be positioned just below the

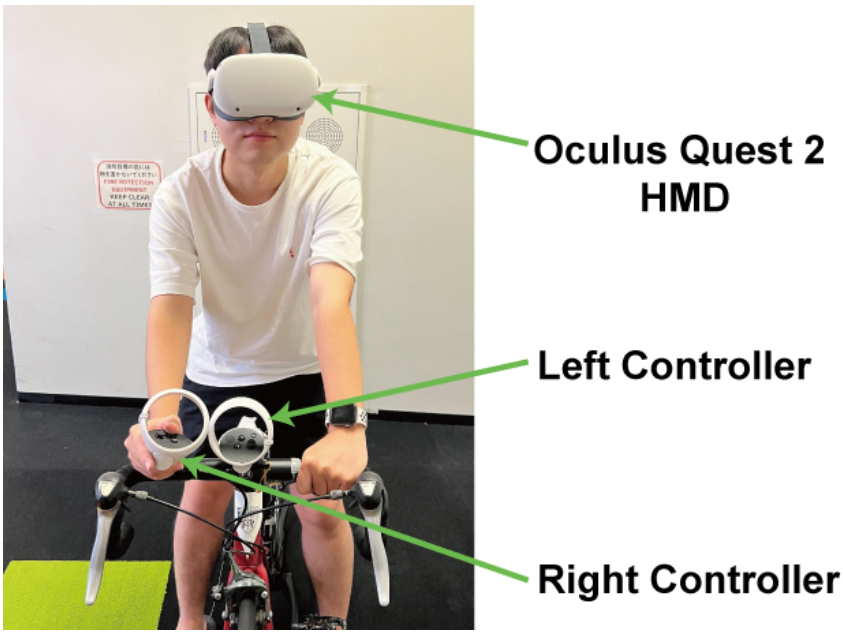

Figure 5.4: A participant riding on the bicycle and wearing the HMD for the experiment. The left controller was mounted on the middle of the handlebar and the right controller was was mounted on the right side of the handlebar.

foveal area, and as it is approaching, the bar's height will increase towards central vision in order to strongly notify the user of the oncoming hazard. Considering the distraction it may cause when implemented to HMDs with limited FoVs, the shape was designed to be simple but easy to understand (Luyten et al. (2016)), and the color was set to blue since the peripheral area is more sensitive to blue compared to the foveal area (Chaturvedi et al. (2019); Janaka et al. (2022)). Furthermore, since the DNN model is not able to calculate speed between frames, I record depth values of the vehicle (if there is a vehicle) which is just behind the bicycle between frames and compare them to detect approach. If there is an approaching vehicle, the indicator will flicker with a high frequency in order to draw attention without inducing too much distraction to the foveal region (Waldin et al. (2017)).

5.4 System Evaluation

In this section, I describe two evaluations of the system's technical performance including an evaluation of detection accuracy and an evaluation of the system's process duration.

Figure 5.5: Some representative frames in the detection accuracy test.

5.4.1 Detection Accuracy

To set up this test, one of my collaborators and I biked in two environments for approximately 16 minutes in each environment, recording data for the duration of the trip. Then I gathered the every 10th frame from the data and set up a frame pool of 2000 frames (Fig. 5.5). From the frame pool, I picked every second frame for testing true and false positives and negatives. The algorithm resulted in 97.67% precision, 92.2% recall, and a 2.2% false-positive rate.

5.4.2 Process Duration

I tested the end-to-end latency of the system by printing timestamps in each step of the system process with 10 iterations and averaged the results. I found that the latency was on average 371ms from the time a vehicle is detected until a bar is rendered. The average latency of each step was: 100.6ms from detection until sending a socket message (with the 50ms sending interval), 262.8ms from sending a socket message until it is received, and 7.6ms from receiving a socket message until a bar is rendered.

5.5 Experiment

In the experiment, I and my collaborators compared the ReAR Indicator to a front-placed virtual monitor and 3D arrow indicators in a VR environment.

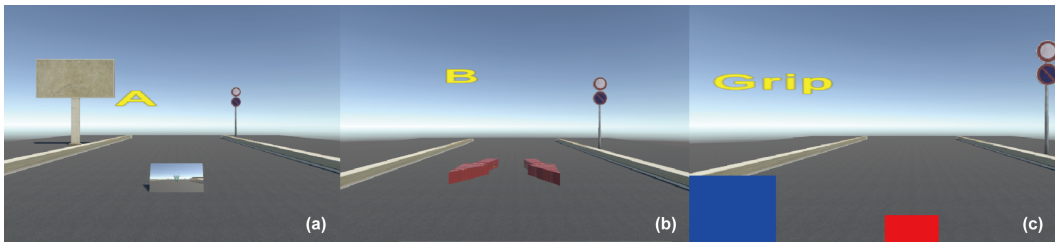

Figure 5.6: Images showing the three UI conditions in the experiment, including (a) the virtual rear-view monitor, (b) 3D arrows, and (c) ReAR Indicator. Letters or words that specified a controller button were displayed in front of the user to simulate cycling tasks such as looking for directions or watching the road.

Participants sat on a stationary bicycle, and interactions in the VR environment were driven by controller, one of which was attached to the bicycle's handlebar (Fig. 5.4). The other was used to record responses to experiment tasks. In addition, participants were asked to avoid approaching vehicles from behind while maintaining a forward focus. We examined their task performance, riding behavior and subjective preferences with the three UI conditions.

5.5.1 Equipment

For the experiment, we used an Oculus Quest 2 HMD, a PREC Ovation 2014 bicycle and a Windows PC. The VR program was developed in and driven by Unity 2021.1.20f1.

5.5.2 Experiment Design

In order to preserve safety in bicycle-related work, a widely accepted study design is to conduct experiments with virtual reality environments and a stationary bicycle (Dancu et al. (2015); von Sawitzky et al. (2020)). Therefore, we decided to implement my experiment with a similar setup. In addition, a controller was mounted on the middle of the bicycle's handlebar to detect steer actions, and the other controller was mounted on the right side of the handlebar for task-related inputs.

Two primary tasks were designed to test participants' performance with the three conditions and included forward focus and backward collision avoidance. For the forward focus task, participants had to quickly press the button that is shown in front of them within two seconds, which is a common interaction in video games, also known as a quick time event (QTE) (Fig. 5.7-d).

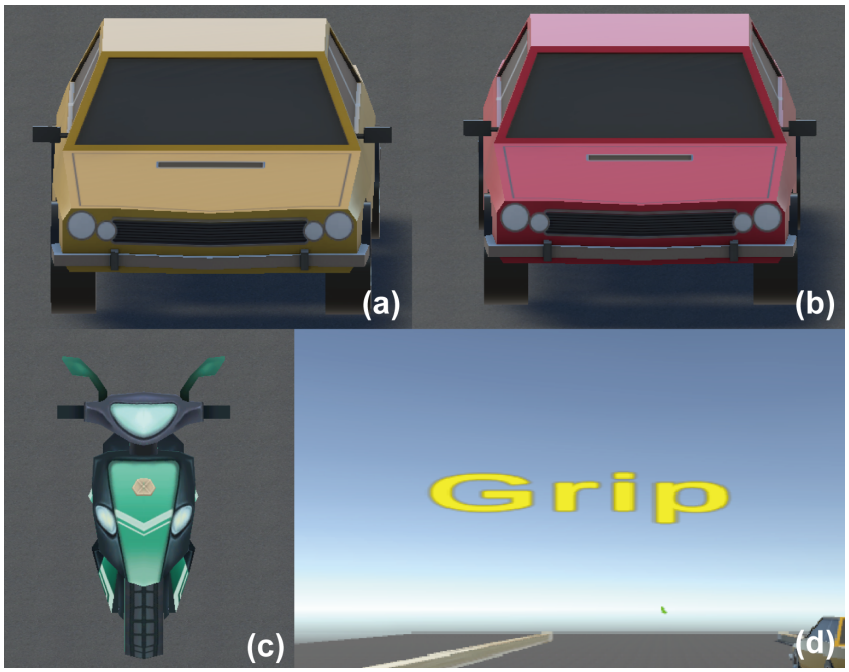

Figure 5.7: Images showing the (a, b) cars, (c) the bike, and (d) an example of the QTEs from Unity.

In order to prevent the participant from randomly pressing buttons to pass the QTEs, we applied an input interval of 0.5s. As the backward collision avoidance task, participants were asked to avoid incoming vehicles from the behind by steering left or right as necessary.

The user study used a within-subjects design that tested the ReAR Indicator against two other UI conditions, a front-placed virtual monitor and 3D arrow interface, which were both fixed in the forward direction just below the horizon (a preferred position for easy access and low distraction during mobile tasks (Orlosky et al. (2013))) (Fig. 5.6). In addition, the front-placed virtual monitor was set at 1.2 meters width and 0.72 meters height in size, six meters forward, and matched the average aspect ratio of mobile phone screen. The order of conditions was assigned based on a balanced Latin-Square in order to minimize order effects.

5.5.3 Experimental Environment

To test the three conditions, we utilized a VR environment that made use of virtual cars and bikes to ensure that participants would not be at risk of crashing into real hazards. Instead of detecting real vehicles, we adapted my method to react to the virtual vehicles within the VR environment for the

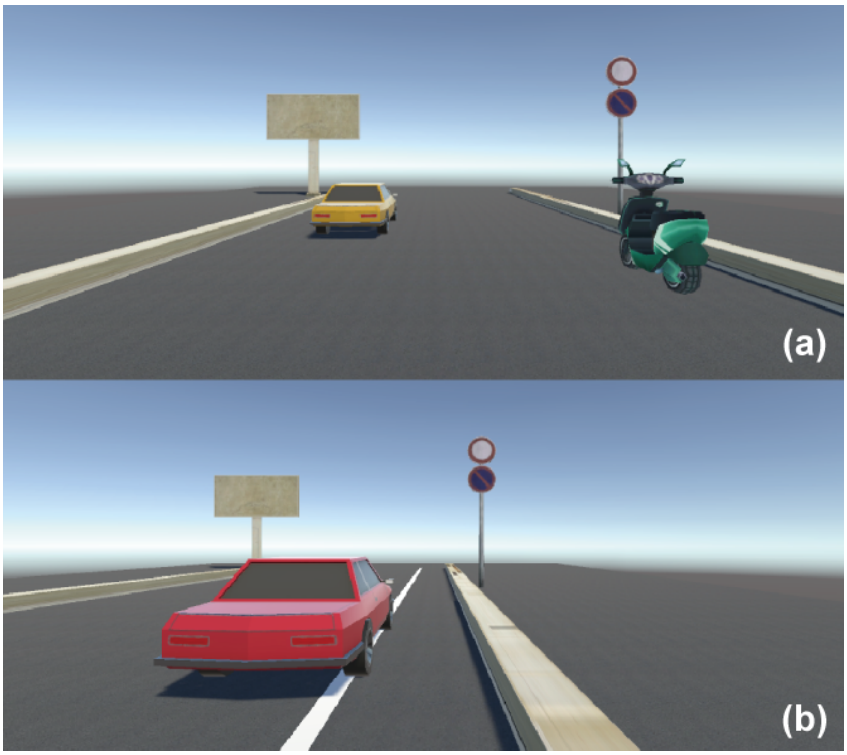

Figure 5.8: Images showing the (a) traffic and (b) roadside scenes.

experiment (Fig. 5.7). In the VR environment, we provided two scenes: a traffic scene and a roadside scene in order to simulate the situations of cycling through traffic and on a roadside cycling lane (Fig. 5.8). These scenes were delivered to the participant with randomized order in each trial.

In the traffic scene, we rendered a road with three lanes, a bike and a car. In addition, the vehicles were set to move on two random lanes at the same time and go through the road repeatedly. Therefore, there was always an empty lane for the participant to avoid the vehicles. In the roadside scene, we rendered a cycling lane and a car which repeatedly and slightly move into the lane and participants had to avoid the car accurately without going too far away from the lane. In both scenes, the lanes were set to 60 meters long and the vehicles were set to move with a random speed from 1m/s to 16m/s (the maximum speed was defined approximately to common urban speed limit 60km/h) when they restarted from the beginning of the lanes. In addition, the participant's position was set to 30 meters ahead of the start line and only passive forward movements of the virtual bicycle were provided in the virtual environments. Since the ReAR Indicator's detection range was set to align with the RGB-D camera's depth range (0.3-20 m) in the VR experiment, newly spawned virtual vehicles in the distance would not immediately be rendered

with the indicator until they came within range.

For improving presence, we also implemented traffic signs and billboards. These were set to move in the opposite direction at a speed of 6m/s (approximately to average cycling speed 20km/h). Moreover, the size of each object in the VR environment resembled the size of its real-world counterpart.

5.5.4 Participants

20 naive individuals (5 female, mean age 26.1, SD 2.73, range 23-31) were recruited for the experiment. All participants had normal or corrected-to-normal vision and at least some experience with bicycles. They were informed that they were allowed to discontinue the experiment at any time. The experiment was conducted under approval of a university ethics committee.

5.5.5 Measurements

As performance measurements, we gathered data and information from the participants with the following approaches: Regarding forward focus task performance we recorded, correct QTEs, missed QTEs (i.e. failed to press the correct button within two seconds for a QTE) and total QTEs. In addition, if a participant is able to correctly react to the QTEs, he/she will get a higher number of total QTEs. Regarding backward collision avoidance task performance, number of collisions (i.e. hit by a vehicle) and time of staying in front of vehicles were gathered. Besides, total angle of head rotation and subjective ratings were gathered. For subjective ratings, post-trial questionnaires were implemented with the following subjective preference questions in 7-scale ratings from 1 ("Strongly disagree") to 7 ("Strongly agree"):

- I had a clear perception of the backward vehicles' positions and speeds while looking forward.
- I could react to the approaching vehicles quickly.
- I felt safe while riding.
- The interface disturbed me.

5.5.6 Procedure

As each participant joined my experiment, we briefly introduced the experiment, and we informed the participants about the tasks and requirements. Then, the participant rides on the bicycle and puts on the HMD with help from an experimenter. Before the first trial started, the participants were

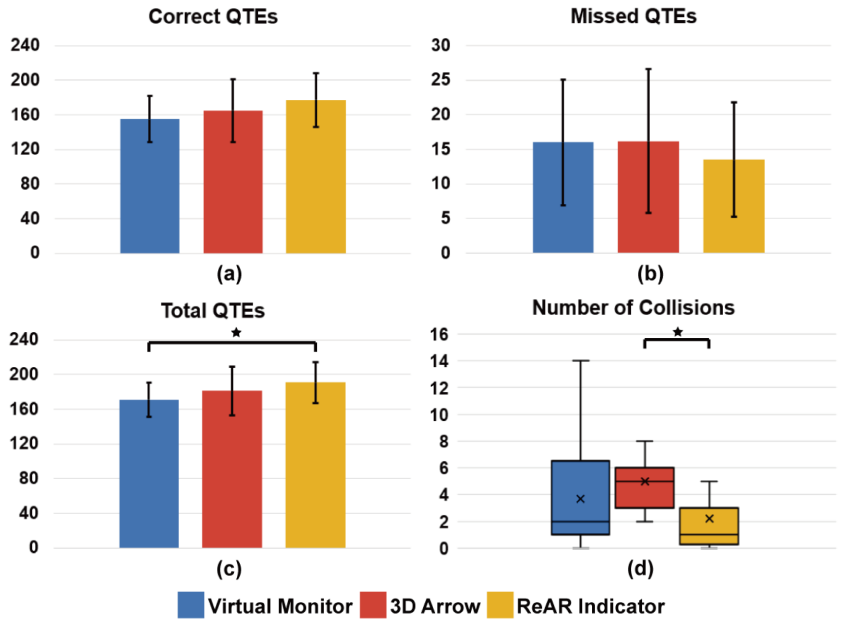

Figure 5.9: Analysis of the data collected from the **traffic scene** including (a) correct quick time events (QTEs), (b) missed QTEs, (c) total QTEs, and (d) number of collisions.

asked to complete a quick tutorial for practicing inputs of QTEs. After each trial, the experimental environment was reset and the participant was asked to fill in a questionnaire.

5.6 Results

Here, I describe the results of my experiment with respect to the three conditions and two scenes, differences in performance, and subjective preference of the participants. I firstly tested the data of each measurement with Q-Q plots and Shapiro tests to determine whether the data is normally deviated. Then, ANOVA tests with Tukey's HSD post-hoc tests and Friedman tests with Wilcoxon post-hoc tests were applied. In addition, I used a level of 0.05 for determining significance.

5.6.1 Task Performance

In the traffic scene, regarding correct QTEs, ANOVA test revealed no significance between the three conditions ($F(2, 57) = 2.298, p = 0.11$) (Fig. 5.9-a). Regarding missed QTEs, no significance was found ($F(2, 57) = 0.487, p = 0.617$) (Fig. 5.9-b). Regarding total QTEs, significance was found ($F(2, 69) =$

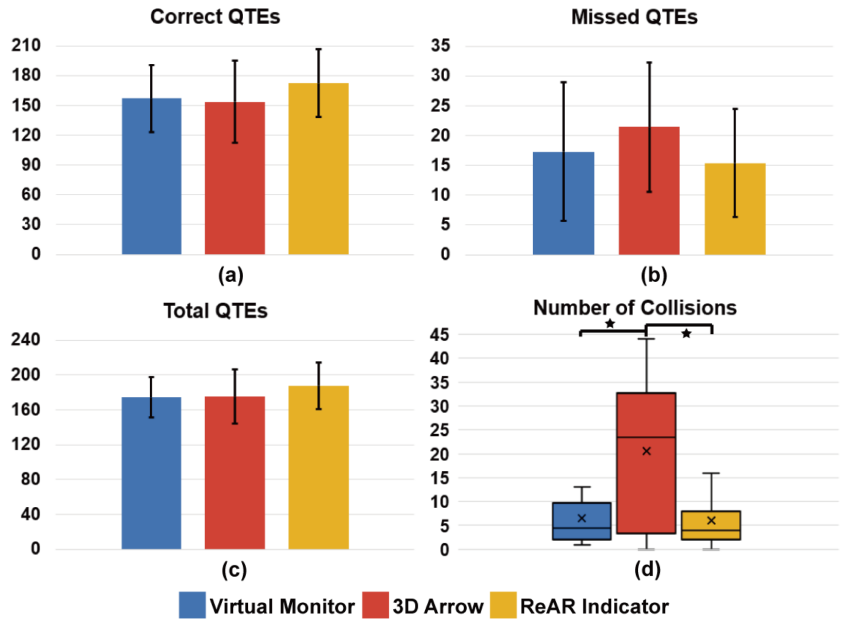

Figure 5.10: Analysis of the data collected from the **roadside scene** including (a) correct quick time events (QTEs), (b) missed QTEs, (c) total QTEs, and (d) number of collisions.

$5.62, p < 0.01$) (Fig. 5.9-c) and post hoc tests showed that there was a significant effect in the ReAR Indicator vs. the virtual monitor ($d = 0.88, p < 0.05$), indicating that when using the ReAR Indicator, participants were more likely to react to the QTEs faster than the virtual monitor. Regarding number of collisions, a significance was found by Friedman test ($\chi^2 = 15.39, p < 0.001$) (Fig. 5.9-d) and further Wilcoxon post-hoc tests revealed a significant effect in the ReAR Indicator vs. the 3D arrow ($V = 136, p < 0.01$), which showed that the ReAR Indicator were able to help participants to collide less than using the 3D arrow. Regarding time of staying in front of vehicles, significance was found ($F(2, 57) = 9.33, p < 0.001$) (Fig. 5.11-a) and post-hoc tests showed that there were significant effects in the ReAR Indicator vs. the 3D arrow ($d = 0.93, p < 0.01$) and the virtual monitor vs. the 3D arrow ($d = 1.51, p < 0.001$). These results indicated that participants averagely reacted to the approaching vehicles from right behind quicker with the ReAR Indicator and the virtual monitor.

In the roadside scene, no significance between the three conditions was found by ANOVA regarding correct QTEs ($F(2, 57) = 1.42, p = 0.25$) (Fig. 5.10-a), missed QTEs ($F(2, 57) = 1.63, p = 0.21$) (Fig. 5.10-b), and total QTEs ($F(2, 57) = 1.44, p = 0.25$) (Fig. 5.10-c). Regarding number of collisions, Friedman test revealed a significance ($\chi^2 = 16.92, p < 0.001$) (Fig. 5.10-

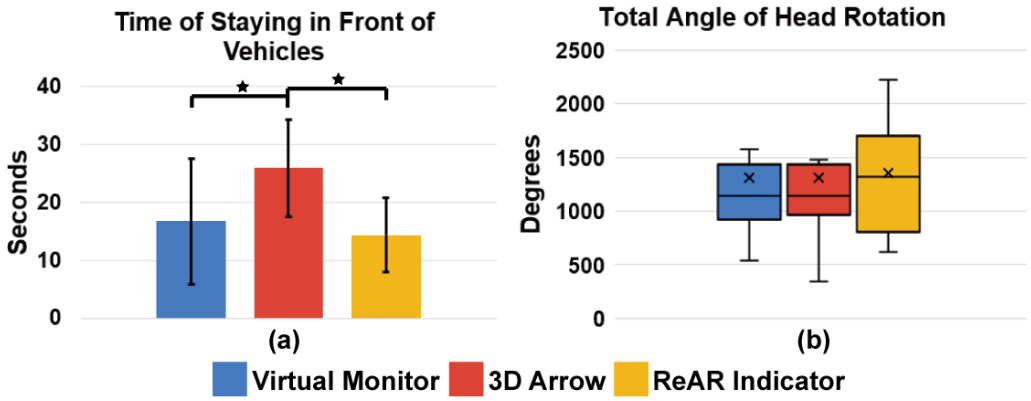

Figure 5.11: Analysis of the data collected from the **traffic scene** including (a) time of staying in front of vehicles and (b) total angle of head rotation.

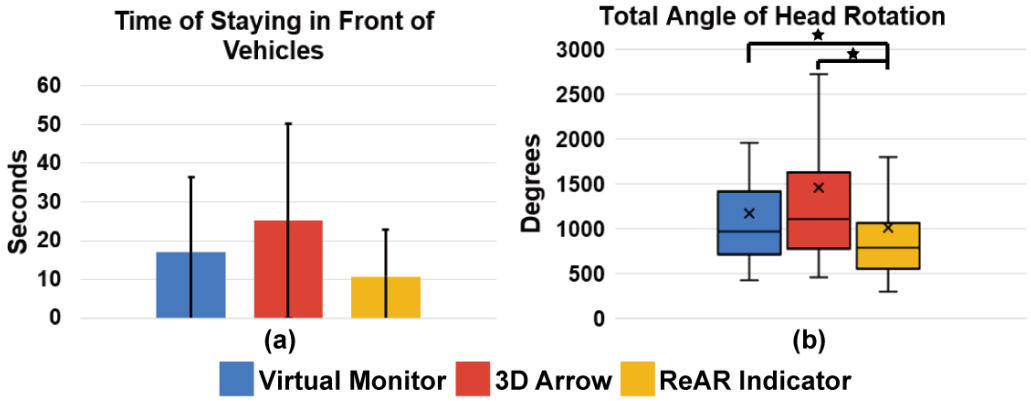

Figure 5.12: Analysis of the data collected from the **roadside scene** including (a) time of staying in front of vehicles and (b) total angle of head rotation.

d), further significant effects were observed in the virtual monitor vs. the 3D arrow ($V = 193, p < 0.01$) and the ReAR Indicator vs. the 3D arrow ($V = 181.5, p < 0.001$), indicating that participants were more likely to collide with approaching vehicles when using the 3D arrow. Regarding time of staying in front of vehicles, no significance was found ($F(2, 57) = 2.547, p = 0.09$) (Fig. 5.12-a).

5.6.2 Head Behavior

To determine whether the methods affected participants' head rotation, Friedman tests were applied to the non-normally distributed total angle of head rotation data. In the traffic scene, no significance was found ($\chi^2 = 0.38, p = 0.83$) (Fig. 5.11-b). In the roadside scene, a significance was found between the three conditions ($\chi^2 = 7.3, p < 0.05$) (Fig. 5.12-b) and post-hoc tests

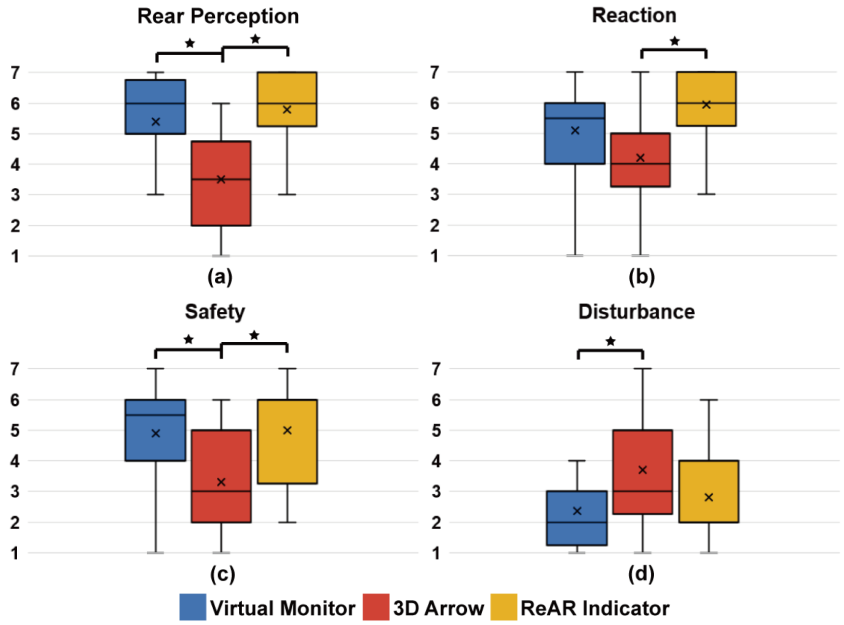

Figure 5.13: Analysis of the subjective ratings collected from the **traffic scene** including (a) rear perception, (b) reaction, (c) safety, and (d) disturbance.

showed that there were significant effects in the ReAR Indicator vs. the virtual monitor ($V = 158, p < 0.05$) and the ReAR Indicator vs. the 3D arrow ($V = 182, p < 0.01$). These results indicated that the ReAR Indicator helped participants to averagely rotate their head less compared to the virtual monitor and the 3D arrow.

5.6.3 Subjective Preference

In the traffic scene, by analyzing the data from post-trial questionnaires, significant differences were found between the three conditions for "rear perception" ($\chi^2 = 25.65, p < 0.001$) (Fig. 5.13-a), "reaction" ($\chi^2 = 12.69, p < 0.01$) (Fig. 5.13-b), "safety" ($\chi^2 = 16.09, p < 0.001$) (Fig. 5.13-c) and "disturbance" ($\chi^2 = 10.03, p < 0.01$) (Fig. 5.13-d). Secondly, post-hoc paired tests revealed significant effects: For "rear perception", significant differences were located in the virtual monitor vs. the 3D arrow ($V = 176.5, p < 0.01$) and the ReAR Indicator vs. the 3D arrow ($V = 190.5, p < 0.01$). For "reaction", a significant difference was found in the ReAR Indicator vs. the 3D arrow ($V = 143, p < 0.01$). For "safety", significant differences were located in the virtual monitor vs. the 3D arrow ($V = 150, p < 0.01$) and the ReAR Indicator vs. the 3D arrow ($V = 138, p < 0.01$). For "disturbance", a significant difference was found in the virtual monitor vs. the 3D arrow ($V = 124, p < 0.01$).

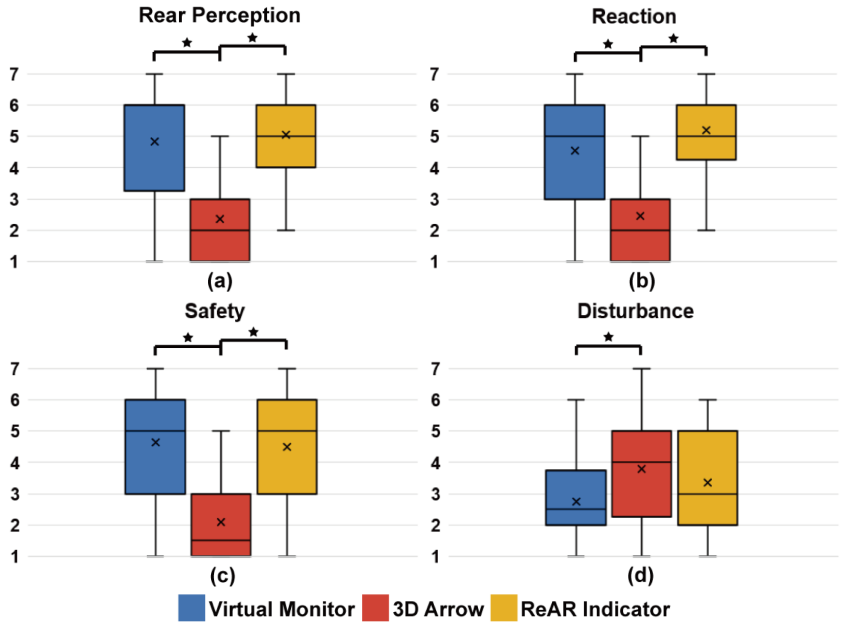

Figure 5.14: Analysis of the subjective ratings collected from the **roadside scene** including (a) rear perception, (b) reaction, (c) safety, and (d) disturbance.

In the roadside scene, significant differences were found between the three conditions for "rear perception" ($\chi^2 = 20.51, p < 0.001$) (Fig. 5.14-a), "reaction" ($\chi^2 = 20, p < 0.001$) (Fig. 5.14-b), "safety" ($\chi^2 = 20.63, p < 0.001$) (Fig. 5.14-c) and "disturbance" ($\chi^2 = 8.03, p < 0.001$) (Fig. 5.14-d). Further post-hoc paired tests revealed significant effects: For "rear perception", significant differences were located in the virtual monitor vs. the 3D arrow ($V = 166, p < 0.001$) and the ReAR Indicator vs. the 3D arrow ($V = 207, p < 0.001$). For "reaction", significant differences were located in the virtual monitor vs. the 3D arrow ($V = 149, p < 0.001$) and the ReAR Indicator vs. the 3D arrow ($V = 153, p < 0.001$). For "safety", significant differences were located in the virtual monitor vs. the 3D arrow ($V = 150.5, p < 0.001$) and the ReAR Indicator vs. the 3D arrow ($V = 147.5, p < 0.001$). For "disturbance", a significant difference was found in the virtual monitor vs. the 3D arrow ($V = 110, p < 0.01$).

5.7 Discussion

Regarding H1, my analysis of the number of collisions data revealed that participants encountered fewer collisions from behind when using the ReAR

Indicator as compared to the 3D arrow in both scenes. Moreover, the virtual monitor was able to help participants to encounter fewer collisions as compared to the 3D arrow in the roadside scene. Besides, there was no significant difference between the ReAR Indicator and virtual monitor in both scenes. These results suggested that the 3D arrow or similar pointers can not help users to perceive approaching vehicles as efficiently as virtual monitors or peripheral indicators. In addition, the ReAR Indicator is able to deliver enough information for avoiding vehicles approaching from backward as well as a large monitor which is in sight. Furthermore, by analyzing the time of staying in front of vehicles data, I found that participants using the ReAR Indicator and virtual monitor outperformed the 3D arrow when cycling through traffic, indicating that participants were able to react to approaching vehicles faster with more detailed information of the vehicles' location when there were multiple vehicles approaching from behind. For the roadside scene, no significant difference in reaction was demonstrated. I believe that while cycling on a narrow cycling lane, participants paid much attention to accurate control of the bicycle and keeping themselves in the lane.

In terms of subjective rating results, participants felt both the ReAR Indicator and virtual monitor could better enhance their perception of approaching vehicles in both scenes as compared to the 3D arrow. Moreover, participants felt that they were able to react faster to approaching vehicles with both the ReAR Indicator and virtual monitor in the roadside scene as compared to the 3D arrow. Both the ReAR Indicator and virtual monitor were able to straightly deliver location information of approaching vehicles without requirements of understanding, this feature is likely to positively affect participants' feeling of perception.

In H2, I hypothesized that my ReAR Indicator may positively affect the participants' reaction to forward information. Therefore, I investigated correct QTEs, missed QTEs, and total QTEs. However, the statistical results did not reveal any significance in correct QTEs and missed QTEs for both scenes, indicating that the three conditions did not affect participants on the number of correctly inputted QTEs or missed QTEs. Besides, analysis of total QTEs revealed that in the traffic scene, participants were able to react to the QTEs faster with the ReAR Indicator as compared to the virtual monitor. In addition, compared with the ReAR Indicator, there was a larger information flow of multiple approaching cars being displayed on the monitor, therefore, participants had to access and process the location and speed of each vehicle with longer duration, thus they had less time to react to the forward located QTEs.

Regarding H3, based on the total angle of head rotation data, participants rotated their head less when using the ReAR Indicator as compared to the virtual monitor and 3D arrow in the roadside scene, suggesting that peripheral indications are able to help users to rotate their head less frequently and keep their view more stable when cycling on roadside. On the other hand, when riding through traffic, the three methods will likely affect head behavior similarly. With further analysis on the subjective ratings of "disturbance", I found that participants on average felt the virtual monitor was less disturbing than the 3D arrow in both scenes. Although both of the methods were placed in the same location of the view, this result is still reasonable since the virtual monitor provides a realistic rearview and no learning cost for them. Besides, the subjective ratings of "safety" indicated that participants considered using the ReAR Indicator and virtual monitor to be safer than using the 3D arrow. I believe that in this user study, feeling of safety is correlated with the feeling of rear perception vehicles, as participants rated "rear perception" low, they would feel less safe.

During the user study, several participants reported that they have driving experience and felt the virtual monitor is similar to rear monitors in cars but larger, therefore they were able to take advantage of the experience without hesitation. This situation may slightly affect the results, especially that they have to pay attention to two tasks at the same time.

5.8 Chapter Conclusion and Future Work

In this work, I proposed a novel augmented reality-based peripheral indicator designed to help cyclists access information about vehicles approaching from behind in a safe and easy manner. To accomplish this, I integrated real-time vehicle detection from a rear-facing camera and peripheral rendering techniques. Vehicle information can then be converted into a bar format in the cyclist's peripheral area, which allows the cyclist to both pay attention to forward obstacles and perceive information about rear vehicles. I then conducted a VR based within-subjects user study with 20 participants to determine how participants would perform with the ReAR Indicator, a virtual monitor, and 3D arrows. In addition, the participants were asked to avoid vehicles that were approaching from the rear while inputting forward-located quick time events in traffic and roadside virtual cycling scenes. Results from this study show that both my proposed approach and a virtual rear view mirror effectively reduced the number of collisions with vehicles in comparison to the 3D arrow. Moreover, the ReAR Indicator enabled users to maintain

higher stability of the head without a significant decrease in forward reaction time. In addition, since the participants were not familiar with peripheral rendering methods, the indicator may prove to be more effective when the peripheral rendering techniques along with AR HMDs are more common and broadly accepted by users in the future.

One current limitation of my work is that vehicles approaching from afar can not be detected until they reach the 20-meters limit of the RGB-D camera. The user may not have enough time to react when a vehicle is approaching quickly. However, longer-range RGB-D cameras may solve this problem in the future. In addition, the implementations of my approach in AR and VR were slightly different regarding the HMDs' FoVs, display resolution, and render transparency. One potential extension of this work is to use a partner drone to help collect better information about obstacles or vehicles that may be out of range of the camera. This might allow for easier vehicle detection and additional time to react to a potential collision. In future studies, I plan to further explore wearable hazard indicating approaches that can better help cyclists to preserve safety in complicated environments. I also plan to conduct experiments testing combinations of forward and backward hazard indicators when cycling.

We hope that this work will encourage the development of other peripheral information delivery techniques for cyclists that can improve riding safety and maintain attention to the forward view as well as the exploration of augmented reality-based peripheral indicator designs that allow for high accessibility and low disturbance.

Chapter 6

Auditory-Vibrotactile Rendering in VEs

6.1 Introduction

In the previous chapters, several visual rendering approaches are introduced. However, rendering beyond the visual domain are also effective on enhancing locomotive experiences, especially immersion. In addition, immersion is a determinant of how much the environment can affect the user's emotion which is correlated to locomotion (Barliya et al. (2013); Uchiyama et al. (2008)) and can affect a user's movement. Therefore, immersion can be considered an important element in providing improved locomotive experiences. Up to now, studies that have aimed to improve immersion in virtual environments have mainly focused on visual and tactile input. In contrast, using audio to influence presence in virtual environments has received much less attention by researchers, despite having demonstrable effects on immersion (Davis et al. (1999)). In particular, sounds that are produced by one's body or avatar (Sound of heartbeat, breath, etc.) can have an impact on perception but have been largely overlooked. Throughout this work, I call this type of sound a "self-representing" sound, which describes a category of ecological sounds that come from a person's own body and activities. A good example is the human heartbeat, which can influence a user's subjective experience of VR as shown by Chen et al. (2017). Another example from the literature by Nordahl et al. showed that footsteps can have an effect on the perception of motion in VR (Nordahl (2005)). Other sounds falling into this category such as breathing, movement of clothing, and the friction between the hands and other objects could also potentially have an effect on presence and immersion.

In this research, my goal is to test whether certain types of sounds delivered through various auditory-vibrotactile feedback devices (Fig. 6.1) can affect factors such as immersion, presence, and emotional state in virtual environments. In addition, these psychological condition can effect locomotive behaviours and experiences (Barliya et al. (2013)). The devices I sought to test included a pair of bone conduction headphones (bonephones), common

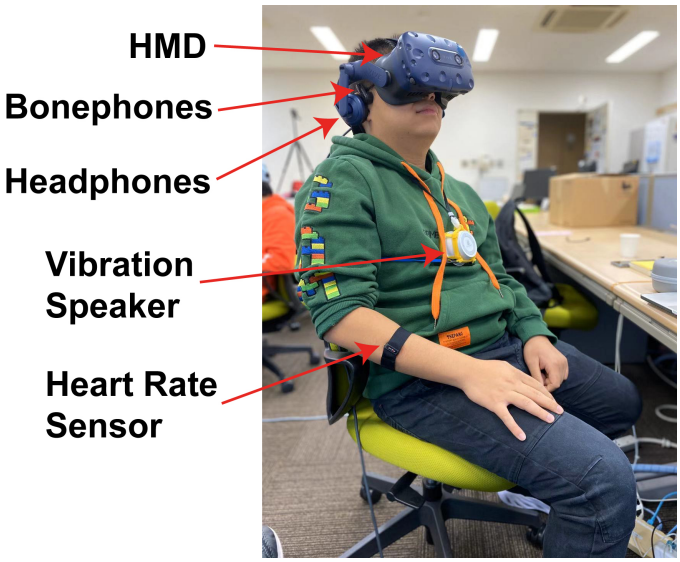

Figure 6.1: A participant equipped with auditory-vibrotactile interaction devices, including built-in headphones on the HMD, bone conduction headphones attached on his cheek and a chest-mounted vibration speaker.

headphones (headphones), a chest-mounted vibration speaker, and auditory-vibrotactile combinations thereof. My hypothesis is that some of the sounds and combinations of hardware will result in an increase in presence or emotional response(fear, relax, pleasure, arousal, and dominance) in situations that elicit certain physiological responses.

In particular, I investigated whether the auditory-vibrotactile self-representing sounds provided by these device sets can enhance immersion and the feeling of presence while immersed in intense (Fig. 6.2) and relaxing (Fig. 6.3) virtual environments. Environmental sounds were also applied to investigate the performance of input methods providing different types of sound. In addition, I observed if there was a significant difference between the bonephones and headphones independently of other factors on influencing immersion and emotion.

Many studies already exist that seek to evoke human emotions, improve immersion, or enhance presence in virtual environments. However, improving immersion with audio based stimuli in virtual reality, especially those that have to do with one's own body and avatar, still has much room for exploration. Therefore, I set out to investigate the effectiveness of self-representing sounds in virtual environments, and to do so I evaluated various audio stimuli with different physical input methods. Examples of the self-representing sounds in my experiments included a fast heartbeat and soft breathing with

Figure 6.2: The VR simulation showing an abandoned hospital, which was used as the intense scene during experiments.

light snoring.

The goal of this research is to investigate:

- whether simulating self-representing sounds by auditory-vibrotactile methods has a larger effect on immersion and emotion than traditional audio methods.
- which audio type, input method, or combination thereof has a larger effect on stimulating immersion and emotion under intense and relaxing scenes.
- the best match of audio type with input method under intense and relaxing virtual environments.
- Whether sounds produced by bonephones are more effective on immersion and emotion than headphones.

In general I want to improve the immersion and presence of virtual environments through various modes of audio. Immersive audio input should always play an important role in designing a convincing VR experience, since it's an indispensable part of the human senses and can potentially enhance a user's perception of his or her self-avatar.

Understanding how to make VR experiences more intense or immersive through sound can enables VR developers to create more effective virtual

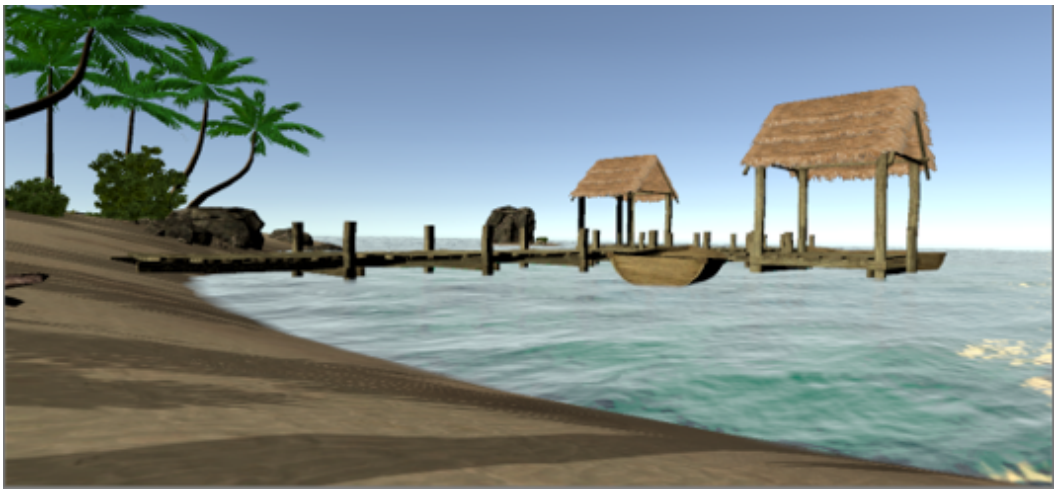

Figure 6.3: The VR simulation showing a tropical beach, which was used as the relaxing scene in experiments.

environments in areas like entertainment and training simulation. Similarly, relaxation via audio can potentially be used for psychotherapy of anxiety or insomnia.

In this research, I sought to test three hypotheses:

- **Hypothesis 1:** self-representing sound simulated by auditory-vibrotactile methods will be more immersive and emotion-eliciting than audio only or vibrotactile only under both intense and relaxing scenes.
- **Hypothesis 2:** Both auditory-vibrotactile methods and self-representing sounds will be more effective for the intense scene than the relaxing scene.
- **Hypothesis 3:** Headphones and bonephones will result in higher immersion when playing environmental sounds than auditory-vibrotactile or vibrotactile methods.

6.2 Prior Work

Many researchers have already explored techniques for evoking human emotions in virtual environments, especially fear. For example, Robillard et al. (2003) evoked anxiety in therapeutic virtual environments derived from computer games. Based on this work, Shiban et al. (2013) observed fear renewal

with multiple context exposure in virtual environments. There is a work related to self-representing sound in VR: Chen et al. (2017) tried to display users' real time heart rate by visual/audio/haptic with different emotional virtual experiences include: happiness, anxiety, fear, disgust, and sadness. They concluded that audio and haptic feedback are most suitable and natural modalities for heart rate representation in a VR environment. Based on Chen's research, I started wondering if audio and haptic inputs are immersive with a simulated heartbeat sound effect and other self-representing sounds. Considering the elements of inducing fear and intensity in VR which were investigated by Lynch and Martins (2015) and Lin (2017), the experimental virtual environment should include unknown darkness, loneliness, and personal relevant or experienced dangerous places. Therefore, I chose an abandoned hospital's dark corridor as the intense scene in my experiment.

In addition to studies on intense virtual experiences, much research has been done on VR to promote relaxation. For example, Annerstedt et al. (2013) demonstrated that stress recovery can be facilitated by the addition of nature sounds to a virtual environment. Although the audio they used in their research was not self-representing sound, they showed the possibility of enhancing positive emotions and stress relief by audio stimuli in virtual environments. Similarly, Anderson et al. (2017) evaluated relaxation in different VR natural scenes. From these studies I drew from the evidence that open, natural virtual environments are effective at relaxing people. Therefore, I chose a tropical island beach as the relaxing scene for my experiment.

The combination of audio and vibration stimuli (auditory-vibrotactile interaction) is not completely new for the research field of VR. It has been applied for several different use cases. For example, Väljamäe et al. (2008) found that auditory-vibrotactile interactions enhance linear vection in virtual environments. To simulate the feeling of walking on different materials in virtual environments, Nordahl et al. (2010) developed a pair of shoes which have both auditory and haptic feedback.

A few examples of works that focus on self-representing sounds also exist, which helped motivate my study. Tajadura-Jiménez et al. (2008) studied how to affect emotions with self-representing sound effects. In their research, they delivered arousal and valence pictures with self-representing sound effects and/or vibrations. The results suggested that self-representing sounds have the ability to evoke emotions, thus, they may play an important role in the design of virtual environments.

6.3 Methodology

Considering the potential for audio and vibrotactile input to affect immersion, I set out to design a series of experiments in which to further test different audio and device input pairings. These came in the form of a pilot test and a main experiment, which were conducted separately.

6.3.1 Pilot Experiment

I first conducted a pilot experiment to investigate whether the auditory-vibrotactile input methods might be useful in stimulating emotions in VR. Four naive (i.e. no knowledge of the experiment) participants (1 female), took part in this pilot experiment. A heartbeat sound was presented both via headphones and a vibration speaker without any visual content. Participants reported that their feelings of intensity were strongly evoked and that they felt very immersed, which were considered as positive feedback. In addition, half of them reported that bonephones were more effective than headphones because of the slight vibration it provided. Since these initial results seemed promising, I moved forward with a primary experiment.

6.3.2 Primary Experiment

In the primary experiment, I compared six different input methods to investigate which combinations of auditory and vibrotactile mechanisms would be more effective at enhancing immersion for a given self-representing sound. Additionally, I explored which kind of sound would be the best match for each input to more effectively influence emotion.

Twenty-two naive participants (9 female, mean age 25.3, SD 6.1) took part in the experiment. All participants had normal hearing and tactile sensing. They were informed that they would experience vibrations on their chests and that their heart rates would be collected during the experiment. Furthermore, they were asked if they have heart diseases or any built-in devices in their bodies. The current study was conducted under approval of a university ethics committee.

6.3.3 Apparatus

I presented two different VR scenes (intense and relaxing as shown in Fig. 6.2 and 6.3, respectively) with an HTC Vive Pro head-mounted device (Fig. 6.4a). For audio stimuli, I included two self-representing sound effects and

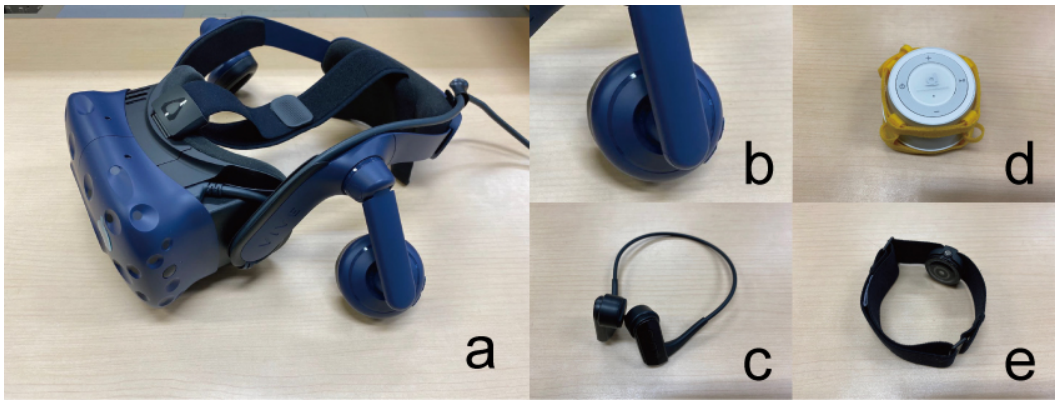

Figure 6.4: Devices used in the experiment, including a) the HTC Vive Pro, b) the speaker from the Vive (used as the regular headphone condition), c) bone conduction headphones, d) the vibration speaker, and e) the polar heart rate monitor.

two environmental sound effects in each scene. The audio stimuli were delivered with six different methods including: headphones (the headphones on the HTC Vive Pro HMD, Fig. 6.4b), bonephones (SUTOMO bone conduction headphones, Fig. 6.4c), a vibration speaker (DOCODEMO, Fig. 6.4d), headphones combined with a vibration speaker, bonephones combined with a vibration speaker, and headphones combined with bonephones. The vibration speaker was always attached to participants' chest. A heart rate sensor (Polar OH1+, Fig. 6.4e) was used to record participants' heart rates. To deliver audio using different devices at the same time, I applied audio virtualization software called "Virtual Audio Cable" [1].

6.3.4 Stimuli and Design

The intense and relaxing scenes were of a dark, abandoned hospital and a tropical beach, which were imported and modified into Unity3D, version 2019.2.8f1. Stimuli were chosen in accordance with the two virtual environments: Under the intense environment, I chose a relatively fast heartbeat (120 bpm) and evil laughing sound effects as the two stimuli. In the relaxing environment, I selected sleeping (soft breathing with light snoring) and ocean wave sound effects as stimuli. These audio clips were presented using all the six input methods described in 3.2. The volume of audio and vibrations were controlled to be the same by windows volume mixer in this experiment. Unity3D was used to view the virtual environments and trigger the sound

[1] https://vac.muzychenko.net/en/

effects and vibrations.

Auditory and/or vibrotactile stimuli lasted 1 minute for each scene. The trials for the intense and relaxing scenes were alternated, with each intense trial being followed with a relaxing trial. The relaxing scene was designed to be delivered after each intense scene so that subsequent changes between intensity and relaxation would be larger. The presentation method for each pair of intense and relaxing scenes was the same. The order of these pairs, i.e. the presentation methods, was randomized. In total, 12 pairs (6 methods times 2 kinds of audio for both scenes), a total of 24 trials, were tested for this experiment.

6.3.5 Procedure

Participants were asked if they had any heart diseases or internal devices (like pacemakers) in their bodies before the experiment started. The procedure of the experiment was then explained to the participants and they were informed that they could discontinue the experiment at any time. The participants were also asked to read a document with detailed instructions and sign a consent form. The participant was then asked to put on the HMD, heart rate sensor and audio input devices, all of which were worn throughout the experiment but only active during their respective conditions. During the experiment, participants were allowed to sit still, look around, turn around and/or freely interact with the environment. I randomized the order of the pairs to minimize ordering conditions. After each trial, participants were required to fill out a questionnaire regarding feelings of immersion, emotion, and subjective preference. Finally, participants were debriefed, thanked, and paid for their participation.

6.3.6 Measures

Participants' heart rates were recorded as an objective measurement. For subjective measures, I used a Self-Assessment Manikin (SAM) (Bradley and Lang (1994)) to measure participants' emotional feelings. In addition, participants' feelings of immersion were measured by subjective scoring on 7-point Likert scales that followed the SAM on the questionnaire. These included questions such as:

For the intense scene:

"I felt immersed when I was in the virtual environment."

"I felt fear when I was in the virtual environment."

For the relaxing scene:

"I felt immersed when I was in the virtual environment."

"I felt relaxed when I was in the virtual environment."

At the end of the experiment there was a final questionnaire with images of the two scenes that measured the participants' overall feelings about each scene.

6.4 Results

Heart rate data was analyzed with ANOVA. SAM ratings (self-reported pleasure, arousal and dominance), and subjective ratings of immersion and fear/relaxation were analyzed with a non-parametric Friedman test and Penalized quasi-likelihood method using the Generalized Linear Mixed Model (GLMMPQL), which was written in R. I tested for differences between the 6 sound delivery methods, including: vibration (V), headphones (H), bonephones (B), vibration + headphones (V+H), vibration + bonephones (V+B) and headphones + bonephones (H+B); and 2 different sound types: self-representing sound (SRS) and environmental sound (ES) under both intense virtual scene and relaxing virtual scene.

To better test my hypothesis, I analyzed data using two methods. The first was single variable, a comparison between each sound delivery method played with each sound type (12 conditions in total), which was analyzed with a Friedman test and post-hoc wilcoxon signed-rank test. The second was multiple variable, a comparison between the 6 sound delivery methods and 2 sound types, which was analyzed with GLMMPQL.

6.4.1 Effects on Heart Rate

In each trial, only the last 10 seconds of time spent in each environment were used for data analysis to allow the participant's heart rate to stabilize and to prevent the activity associated with adjustment of the HMD from affecting the results. Hence, the measure of average heart rate includes data from the last 10 seconds of each trial.

Most of the heart rate data I collected tended to be visibly stable, and changes in participants' heart rates were relatively small. An ANOVA test on the heart rate data revealed no significant effect on participant heart rate, which supports this observable tendency.

Table 6.1: Summary Output from GLMMPQL Models of Immersion Ratings in the Intense Scene.

Factor	Value	Std. Error	DF	t-value	p-value
(Intercept)	1.4575	0.0622	236	23.4447	<0.0001
MethodH	-0.0379	0.0405	236	-0.9355	0.3505
MethodHB	0.0169	0.0394	236	0.4295	0.6680
MethodV	0.0045	0.0396	236	0.1145	0.9090
MethodVB	0.0862	0.0382	236	2.2592	0.0248*
MethodVH	0.1017	0.0379	236	2.6859	0.0077*
SoundSRS	0.0340	0.0223	236	1.5269	0.1281

Correlation

Factor	(Intercept)	MethodH	MethodHB	MethodV	MethodVB	MethodVH
MethodH	-0.314					
MethodHB	-0.322	0.495				
MethodV	-0.320	0.492	0.505			
MethodVB	-0.333	0.511	0.525	0.522		
MethodVH	-0.335	0.515	0.529	0.526	0.547	
SoundSRS	-0.185	0.000	0.000	0.000	0.000	0.000

6.4.2 Subjective Ratings of Immersion, Fear, and Relaxation

6.4.2.1 Single Variable

I analyzed data regarding subjective feelings with a Friedman test since participants' feelings of immersion and fear had significant variations upon observation of the data. However, after a post-hoc wilcoxon signed-rank test with Benjamini-Hochberg correction, no significant effects emerged.

6.4.2.2 Multiple Variable

Results from the GLMMPQL analysis of immersion ratings of the intense scene (Table 6.1) suggest that the method of sound delivery had an effect on participants' feelings of immersion. The vibration + bonephones (estimate = 0.086, $p < 0.05$) and vibration + headphones (estimate = 0.102, $p < 0.01$) were higher rated for immersion than the other methods. There was no significant difference between the two sound types in this condition.

Table 6.2: Summary Output from GLMMPQL Models of Immersion Ratings in the Relaxing Scene.

Factor	Value	Std. Error	DF	t-value	p-value
(Intercept)	1.5637	0.0662	236	23.6074	<0.0001
MethodH	0.0053	0.0488	236	0.1077	0.9143
MethodHB	0.0119	0.0486	236	0.2456	0.8062
MethodV	-0.0636	0.0505	236	-1.2594	0.2091
MethodVB	0.0532	0.0476	236	1.1161	0.2655
MethodVH	0.0496	0.0477	236	1.0389	0.2999
SoundSRS	-0.1529	0.0282	236	-5.4141	<0.0001***

Correlation

Factor	(Intercept)	MethodH	MethodHB	MethodV	MethodVB	MethodVH
MethodH	-0.370					
MethodHB	-0.371	0.504				
MethodV	-0.357	0.485	0.487			
MethodVB	-0.379	0.514	0.516	0.497		
MethodVH	-0.378	0.514	0.515	0.496	0.526	
SoundSRS	-0.181	0.000	0.000	0.000	0.000	0.000

Regarding subjective ratings of fear under the intense scene, no significant interactions were revealed between sound delivery methods or sound types and feeling of fear.

The GLMMPQL analysis of immersion ratings for the relaxing scene (Table 6.2) suggested that participants' feelings of immersion was affected by sound types: the self-representing sound (estimate = -0.153, p <0.0001) was less effective on enhancing immersion than the environmental sound. On the other hand, no significant differences between the six sound delivery methods was found in this condition.

As the result of applying GLMMPQL to subjective ratings of relaxation under the relaxing scene (Table 6.3), a significant effect on feeling of relaxation was found between sound types. self-representing sound (estimate = -0.214, p <0.0001) was less effective on enhancing immersion than environmental sound. Similar to the results of immersion ratings for the relaxing scene, no significant interaction was found between sound delivery methods.

6.4.3 Effects on Self-reported Emotional Experience: SAM Ratings

6.4.3.1 Single Variable

In the intense scene with environmental sound stimulus (Fig. 6.5), a significant effect was found for pleasure rating via a non-parametric Friedman test (Chi-square = 19.227, df = 5, $p < 0.01$) and paired Wilcoxon signed-rank tests with Benjamini-Hochberg correction. These showed significant differences for vibration vs. bonephones ($V = 21$, $p < 0.05$), headphones vs. vibration + bonephones ($V = 62$, $p < 0.05$) and bonephones vs. vibration + headphones ($V = 148.5$, $p < 0.05$).

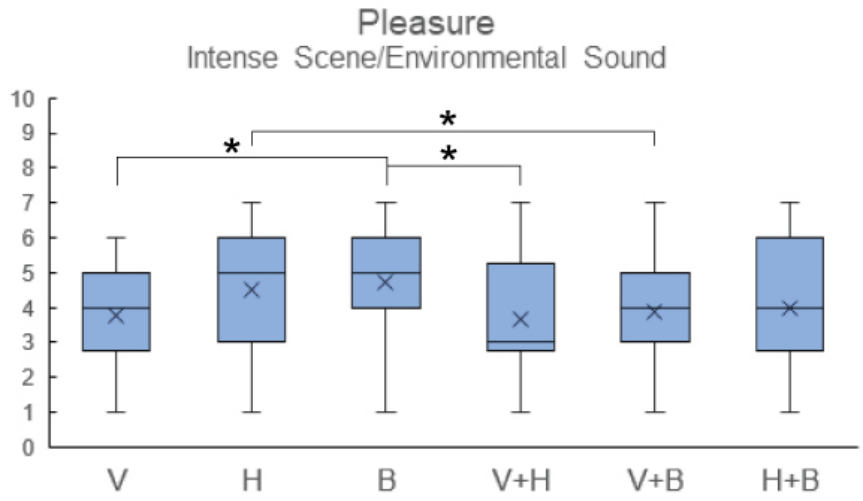

Figure 6.5: Self reported pleasure in the intense scene with the environmental sound.

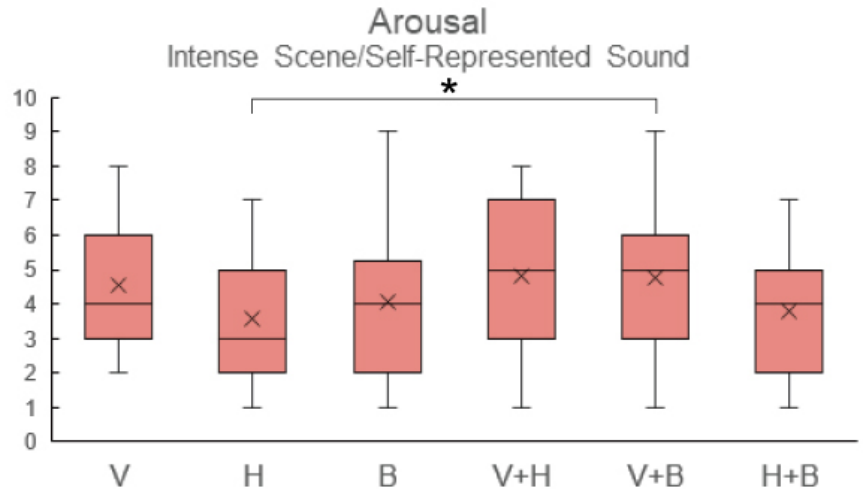

Figure 6.6: Self-reported arousal for the self-representing sound in the intense scene.

Table 6.3: Summary Output from GLMMPQL Models of Relaxation Ratings in the Relaxing Scene.

Factor	Value	Std. Error	DF	t-value	p-value
(Intercept)	1.7198	0.0549	236	31.3118	<0.0001
MethodH	0.0033	0.0515	236	0.0649	0.9483
MethodHB	0.0065	0.0514	236	0.1271	0.8990
MethodV	0.0012	0.0516	236	0.0224	0.9821
MethodVB	-0.0292	0.0524	236	-0.5567	0.5783
MethodVH	0.0148	0.0512	236	0.2879	0.7737
SoundSRS	-0.2143	0.0305	236	-7.0260	<0.0001***

Correlation

Factor	(Intercept)	MethodH	MethodHB	MethodV	MethodVB	MethodVH
MethodH	-0.471					
MethodHB	-0.471	0.503				
MethodV	-0.470	0.501	0.502			
MethodVB	-0.463	0.493	0.494	0.492		
MethodVH	-0.473	0.504	0.505	0.504	0.496	
SoundSRS	-0.219	0.000	0.000	0.000	0.000	0.000

I also observed some pairs that had p values lower than 0.1: vibration vs. headphones, headphones vs. vibration + headphones, bonephones vs. vibration + bonephones, bonephones vs. headphones + bonephones.

For the intense scene with self-representing sound stimulus (Fig. 6.6), a significant effect was found for arousal rating via a non-parametric Friedman test (Chi-square = 18.313, df = 5, p <0.01) and paired Wilcoxon signed-rank tests with Benjamini-Hochberg correction, which showed significant differences for headphones vs. vibration + bonephones (V = 0.001539, p <0.05). Some pairs also had p values lower than 0.1, including vibration vs. headphones, vibration vs. headphones + bonephones, headphones vs. vibration + headphones, headphones vs. vibration + bonephones, vibration + headphones vs. vibration + bonephones, vibration + headphones vs. headphones + bonephones.

In the relaxing scene with the self-representing sound stimulus (Fig. 6.7), significance was found for pleasure rating via a non-parametric Friedman test (Chi-square = 16.438, df = 5, p <0.01) and paired Wilcoxon signed-rank tests with Benjamini-Hochberg correction, which showed significant differences for vibration + bonephones vs. headphones + bonephones (V = 153, p <0.05).

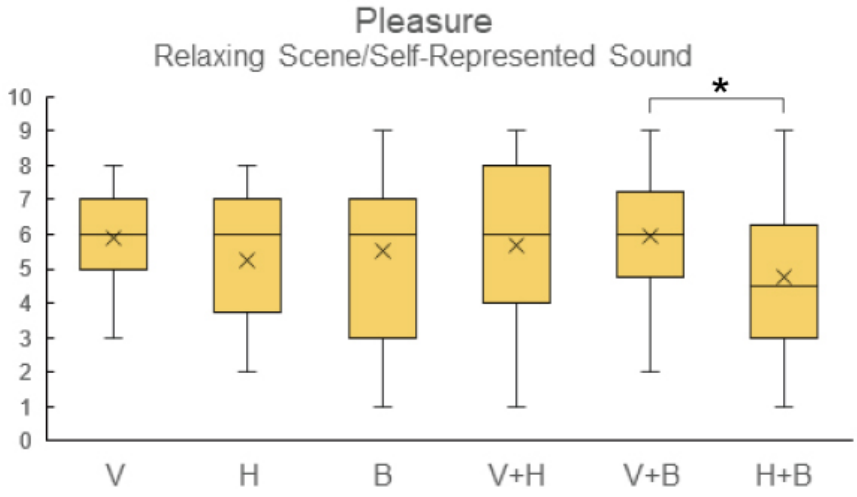

Figure 6.7: Self-reported pleasure in the relaxing scene with the self-representing sound.

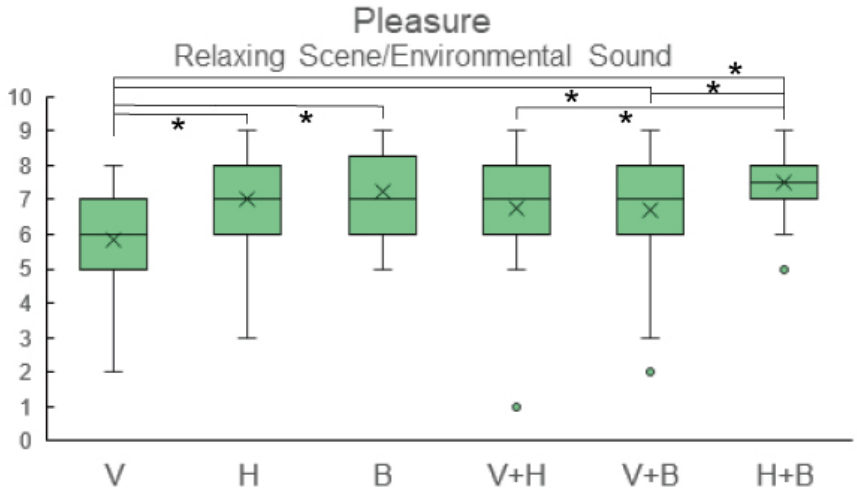

Figure 6.8: Self-reported pleasure in the relaxing scene with the environmental sound.

Table 6.4: Summary Output from GLMMPQL Models of Pleasure Ratings in the Intense Scene.

Factor	Value	Std. Error	DF	t-value	p-value
(Intercept)	1.4430	0.0842	236	17.1410	<0.0001
MethodH	-0.0074	0.0524	236	-0.1417	0.8875
MethodHB	-0.0755	0.0543	236	-1.3908	0.1656
MethodV	-0.1616	0.0570	236	-2.8368	0.0050*
MethodVB	-0.0847	0.0545	236	-1.5532	0.1217
MethodVH	-0.1092	0.0553	236	-1.9754	0.0494*
SoundSRS	0.0212	0.0323	236	0.6548	0.5132

Correlation

Factor	(Intercept)	MethodH	MethodHB	MethodV	MethodVB	MethodVH
MethodH	-0.309					
MethodHB	-0.298	0.479				
MethodV	-0.284	0.457	0.441			
MethodVB	-0.297	0.477	0.460	0.438		
MethodVH	-0.293	0.470	0.454	0.433	0.452	
SoundSRS	-0.196	0.000	0.000	0.000	0.000	0.000

At the same time, some pairs having p values lower than 0.1 were observed, including vibration vs. headphones + bonephones, headphones vs. vibration + bonephones, vibration + headphones vs. headphones + bonephones.

For the relaxing scene with environmental sound stimuli (Fig. 6.8), a significant effect was found for pleasure rating via A non-parametric Friedman test (Chi-square = 25.736, df = 5, p <0.001) and paired Wilcoxon signed-rank tests with Benjamini-Hochberg correction show significant differences for vibration vs. headphones (V = 13, p <0.01), vibration vs. bonephones (V = 4.5, p <0.01), vibration vs. vibration + bonephones (V = 22, p <0.05), vibration vs. headphones + bonephones (V = 10, p <0.01), vibration + headphones vs. headphones + bonephones (V = 4, p <0.05), vibration + bonephones vs. headphones + bonephones (V = 10, p <0.05). I also observed the condition of vibration vs. vibration + headphones, which has a p value that is lower than 0.1.

Table 6.5: Summary output from GLMMPQL models of arousal ratings in the intense scene.

Factor	Value	Std. Error	DF	t-value	p-value
(Intercept)	1.3270	0.0901	236	14.7213	<0.0001
MethodH	-0.1679	0.0793	236	-2.1184	0.0352*
MethodHB	-0.1090	0.0767	236	-1.4223	0.1562
MethodV	0.0457	0.0708	236	0.6449	0.5196
MethodVB	0.0697	0.0700	236	0.9964	0.3201
MethodVH	0.0936	0.0692	236	1.3527	0.1775
SoundSRS	0.0944	0.0421	236	2.2425	0.0259*

Correlation

Factor	(Intercept)	MethodH	MethodHB	MethodV	MethodVB	MethodVH
MethodH	-0.367					
MethodHB	-0.379	0.431				
MethodV	-0.411	0.467	0.483			
MethodVB	-0.415	0.472	0.488	0.529		
MethodVH	-0.420	0.477	0.494	0.535	0.541	
SoundSRS	-0.255	0.000	0.000	0.000	0.000	0.000

6.4.3.2 Multiple Variable

I analyzed the data I obtained from self-reported feelings of dominance in the SAM ratings. In both intense and relaxing scenes, no significant effect was found. In addition, based on oral feedback, most of my participants reported that they had difficulties in understanding the meaning of the images of dominance ratings in SAM. I believe that may be the reason why no significant results were revealed for this condition.

I applied GLMMPQL to the self-reported SAM ratings for the intense scene, and results suggested that method of sound delivery had an effect on a participant's feeling of pleasure in virtual environment (Table 6.4). I.e., the participant's feeling of pleasure was weakened when experiencing the virtual environment with vibration (estimate = -0.162, $p < 0.01$) and vibration + headphones (estimate = -0.109, $p < 0.05$), suggesting that these two methods were more effective. In the dimension of arousal (Table 6.5), the result from the GLMMPQL test showed that headphones (estimate = -0.168, $p < 0.05$) has less of an effect on a participant's feeling of arousal between sound delivery methods. In addition, self-representing sounds (estimate = 0.094, $p < 0.05$)

Table 6.6: Summary output from GLMMPQL models of pleasure ratings in the relaxing scene.

Factor	Value	Std. Error	DF	t-value	p-value
(Intercept)	1.9397	0.0510	236	37.9987	<0.0001
MethodH	-0.0317	0.0482	236	-0.6565	0.5122
MethodHB	-0.0248	0.0481	236	-0.5170	0.6056
MethodV	-0.0997	0.0500	236	-1.9920	0.0475*
MethodVB	0.0008	0.0475	236	0.0159	0.9873
MethodVH	-0.0241	0.0481	236	-0.5007	0.6170
SoundSRS	-0.2028	0.0288	236	-7.0432	<0.0001***

Correlation

Factor	(Intercept)	MethodH	MethodHB	MethodV	MethodVB	MethodVH
MethodH	-0.457					
MethodHB	-0.459	0.486				
MethodV	-0.441	0.467	0.468			
MethodVB	-0.465	0.492	0.494	0.474		
MethodVH	-0.459	0.486	0.488	0.468	0.494	
SoundSRS	-0.226	0.000	0.000	0.000	0.000	0.000

were more effective compared to environmental sound.

For the relaxing scene, a GLMMPQL test showed that there were significant differences between sound delivery methods and sound types in self-reported feelings of pleasure (Table 6.6). Vibration (estimate = -0.1, $p < 0.05$) was less effective than other methods, while self-representing sound (estimate = -0.203, $p < 0.0001$) was less effective than environmental sound.

For SAM ratings of arousal under the relaxing scene, no significant interaction between sound delivery methods, sound types, or feeling of arousal was revealed.

6.5 Discussion

In this research, I evaluated the effectiveness of audio input methods, vibration and sound type on participants' feelings of immersion and emotional changes(fear, relax, pleasure, arousal, and dominance) under two different VR scenes. In addition, I investigated the interaction between auditory-vibrotactile self-representing sound stimuli and participants' effective state in

intense and relaxing scenes. This interaction and it's effectiveness was assessed by measuring the participants' heart rate, self-reported emotional ratings and their subjective rating of the virtual environment that they experienced.

In the intense scene (the abandoned hospital), delivering sounds using vibration+bonephones and vibration+headphones resulted in a significant enhancement of immersion. In addition, methods which included vibration input weakened participants' feelings of pleasure and enhanced their feelings of arousal. With respect to sounds, the self-representing sound (fast heartbeat) was not observed to be more influential than the environmental sound (evil laughing). Although input methods with vibration significantly influenced emotional responses in the intense scene, presenting the fast heartbeat sound did not result in a significant increase in participants' heart rates for any of the six input methods. These results do not fully corroborate the results found by Tajadura-Jiménez et al. (2008), and I believe this situation was caused by the difference in presentation methodology and duration of the study. In addition, an immersive virtual environment may weaken the awareness of vibrotactile sensations on participants' real bodies, thereby reducing the affect of audio heart rate on actual heart rate. For example, studies of self-consciousness and pain thresholds with virtual bodies by Hänsell et al. (2011) are a good demonstration of this effect.

In the relaxing scene (tropical beach), the environmental sound (ocean waves) stimulus better enhanced immersion, relaxation and participants' feelings of pleasure than the self-representing sound (soft breathing with light snoring). Moreover, the vibration+bonephones and other two input methods for delivering self-representing sounds enhanced feelings of pleasure while headphones and bonephones significantly influenced the feeling of pleasure by delivering environmental sounds. These results demonstrated that in the relaxing virtual environment, the environmental sound in general improved user experience (immersion and emotion) than the self-representing sound. On the other hand, the coupling of vibration based input methods with self-representing sounds and coupling of input methods without vibration with environmental sounds more significantly evoked pleasure.

In general, self-representing sounds played with the auditory-vibrotactile input methods did not significantly enhance immersion in VR as I hypothesized. However, the combinations of these inputs with various sounds significantly evoked emotion in each scene. In addition, auditory-vibrotactile input methods themselves, regardless of sound type, were found to enhance immersion. Moreover, I found no significant differences between vibration + headphones and vibration + bonephones and between the performance of

headphones and bonephones. Originally, since bone conduction headphones deliver light vibrations on the face of the participant, I expected that the vibration to produce different results for bonephones and common headphones. However, the data suggests that this light vibration did not influence any outcomes, so the headphones and bonephones performed similarly in my experiment.

Since sounds are born from vibrations, self-representing sounds always come with vibrations inside human body. In the auditory-vibrotactile input methods, vibration speakers was chosen to simulate the vibration represented by human body while making self-representing sounds. The reason why I applied vibration speakers instead of common vibration motors, was to meet the sound wave and frequency of those sounds. In addition, I decided to place the vibration speakers on participants' chest because the self-representing sounds I investigated in this research (heartbeat and breath) are all generated around human chest. The results indicated that the application of vibration speakers met my expectations. Besides, I also concerned about place the vibration speaker directly above participants' heart but this idea was given up since it may cause resonance with the real heartbeat.

Based on the oral feedback I collected from the participants, most of them reported that hearing and/or feeling sound from the human body reduced their feeling of relaxation to some extent. This suggests that self-representing sounds may have been a distraction in the relaxing virtual environment, especially when the self-representing sounds were not synchronized with the participants' actual body movements.

6.6 Chapter Conclusion and Future Work

Based on my experimental results, certain self-representing sounds and auditory-vibrotactile inputs appear to have the ability to enhance immersion and evoke emotions in VR. Future research should test these self-representing sounds with auditory-vibrotactile inputs in more dynamic virtual environments by allowing participants to walk around or experience interactive scenarios. In addition, the actual tasks and interactions that participants complete in an experiment may have a cumulative effect when combined with various sounds, further affecting participants' feelings of presence. As such, self-representing sounds deserve further consideration as a tool to enhance immersion, especially for studies involving a self-avatar (Slater and Usoh (1994)). Avatars that include synchronized visual cues and self-representing audio feedback in particular are good targets for future studies.

In this research, I investigated the effects of self-representing sounds and auditory-vibrotactile inputs in two different virtual environments: an abandoned hospital (intention) and tropical beach (relaxation). Though I logically selected the most appropriate self-representing and environmental sounds for each scene, experiences that evoke other emotions such as inquisitiveness, wonder, fun, or even learning could be tested next.

As intense virtual experiences with auditory-vibrotactile inputs tend to be more emotionally evoking and engaging, training simulators for sports enthusiasts, stage performers and rescuers (e.g., firefighters) could benefit from enhanced sound to provide more realistic feelings of stress, nervousness, or fear. Exciting facilities in theme parks could likely make use of auditory-vibrotactile feedback, and could even be applied to 4D movies and VR roller coasters. On the other hand, since self-representing sounds could also weaken immersion during relaxation, they might be helpful for improving situational awareness, for example when students fall asleep when they get too relaxed during e-learning in VR.

Chapter 7
Conclusion

This dissertation presents theories of developing and applying real-time adaptive rendering for locomotion in virtual and augmented realities. This theory postulates that low-contrast vision modulators, gaze-based augmentation delivery, and peripheral indicators can outperform traditional methods for locomotive interaction. Three methods are deployed for the attempt, including: 1) a method governed by visual angular velocities for mitigating VR sickness, 2) a method that enables users to focus on both a hazard and interface at the same time, and 3) a method that makes use of peripheral vision to alert the user of oncoming dangers. A follow-up study also explores the usability of different real-time adaptive rendering methods that are suitable for locomotive scenarios.

7.1 Summary

A summary of the contributions and findings made by this work is shown as follows.

Motion-Based Dynamic Vision Modulator. This work proposes a novel rendering approach for visual modification that allows the user to suffer less VR sickness symptoms and stay longer in locomotive virtual environments. This approach calculates visual angular velocities of the pixels passing through the user's vision and samples the average color of the pixels in real-time. This work conducted two user studies involving short and long washout period in two different virtual scenes in order to evaluate the effectiveness of the vision modulator along with the reliability of a just-noticeable experimental design for VR sickness. In summary, the main contributions of this work are:

- The design, testing, and refinement of a vision modulator designed to reduce VR sickness and without becoming distracting by reducing optical flow, engaging only during potentially sickness-causing motions, and preserving scene coloration.

- Two experiments with different washout periods testing this modulator against a state-of-the-art FoV-modulation technique and a baseline condition, as well as analysis and discussion of the results.

Gaze-based Augmented Information Delivery. This work proposes a novel system called HazARdSnap that allows the user to simultaneously read augmented UIs while still focusing on forwardly located hazards. This system makes use of DNN based object detection and eye-tracking in two subsystems, where the subsystems are able to not only gather and calculate position data, but also exchange data via socket communication. Furthermore, this work presented the design of a novel MR based experiment, which revealed the efficiency of augmented information delivery methods. In summary, the contributions of this work are:

- This work introduces a UI snapping technique, HazARdSnap, designed to decrease the chance of collisions when accessing digital information during cycling.
- A bike-mounted system that makes use of computer vision, depth information, and environment tracking to detect oncoming hazards such as potholes and pedestrians.
- An MR-based user study that compares cyclist performance while using HazARdSnap, a forward-fixed UI, and a bike-mounted smartphone UI as well as analysis and discussion of metrics like collisions with virtual objects, time spent viewing the road, and subjective ratings.

Peripheral Indicators for Rear-Approaching Vehicles. In this work, a novel method called ReAR Indicator is proposed to properly indicate rear-approaching vehicles on the periphery while cycling. The system first takes data composed of real-time vehicle detection and then compute to the width, length and position of bar indicators on the periphery. A user study in two different virtual scenarios indicates that ReAR Indicator can help the participants to perform as well as a large rearview monitor in both scenarios. In summary, the main contributions of this work include:

- This work proposes a novel peripheral rendering which enhances awareness of rear-approaching vehicles while moving to and focusing on the forward direction.

- The experimental results reveal that that the ReAR Indicator can effectively help cyclists maintain focus on their forward view while still avoiding collisions with virtual vehicles approaching from the rear.

In general, this work presents some implications regarding applying real-time adaptive rendering techniques for improving interaction as well as user experience in locomotive VR and AR. It shows that rendering following certain designs can fulfill the need of users in both virtual or real locomotive scenarios. The experimental results demonstrated that the proposed approaches are at least able to help users to perform equally and even better in some perspectives as compared with state-of-the-art methods.

I hope that the findings of this work can contribute to the field of rendering for locomotion and inspire future designs and developments of locomotive systems in VR and AR. Furthermore, I believe that it is worthwhile to dig deeper into this field for more novel insights that could not only help evolve real-time adaptive rendering techniques, but also lead to innovations of the next-generation human-VR/AR interaction.

7.2 Suggestions for Future Work

This work with its findings has taken a step forward towards achieving high-efficient interaction for locomotion in virtual and real worlds. However, there are still various tasks left regarding rendering for locomotion. This section outlines two concrete ideas accordingly for inspiring further research in the same domain.

Gaze-Aware Dynamic Vision Modulator. In this work, Chapter 3 made one attempt towards mitigate VR sickness without immersion loss, however there still remain valuable tasks unresolved or undiscovered for the same problem. In my point of view, taking advantage of eye-trackers and monitoring gaze behaviours can be a potential direction for future works.

Although the investigation by Adhanom et al. (2020) has indicated that a foveated FoV restrictor can not significantly mitigate VR sickness symptoms, there still remains spaces for putting further efforts. For example, future studies can focus on investigating strafing eye movements since passive eye rotation is a well-known trigger of dizziness. In addition, for rendering gaze-aware vision modulators or fixed frames, effects which are relatively complicated and customizable (e.g. radial blur, bilateral filter, etc.) can be useful. The author

believes that combining real-time rendering with eye-tracking will contribute to solving the problem of locomotion induced VR sickness.

Augmented Information Exchange in Groups. To date, it is promising that AR HMDs are capable of becoming a reliable tool for gathering information and interacting with systems in different industries. However while various studies, including the works introduced in this dissertation, are focusing on promoting user experience and efficiency for single user, there still lacks novel approaches allowing for multiple collaborating users in a group to exchange information during locomotion. Moreover, studying such methods will definitely help users in collaborative scenarios such as firefighting, restaurant serving, or group cycling.

One proper direction for this topic might be sharing location, motion and detection data between HMDs in real-time and rendering the shared information in the view of each user. In addition, by predicting the interactions between user-user, user-environment, or user-object, interfaces can be rendered with less distraction. The author suggests future studies to explore this direction and contribute to the evolution of real-time rendering techniques for locomotion in groups.

Appendix A

Additional Images

A.1 Some representative frames from the frame pool of human detection in the detection accuracy test, described in Chapter 4

A.2 Some representative frames from the frame pool of pothole detection in the detection accuracy test, described in Chapter 4

Bibliography

Abdi, L. and Meddeb, A. (2018). In-vehicle augmented reality system to provide driving safety information. *Journal of Visualization*, 21(1):163–184.

Adhanom, I. B., Griffin, N. N., MacNeilage, P., and Folmer, E. (2020). The effect of a foveated field-of-view restrictor on vr sickness. In *2020 IEEE conference on virtual reality and 3D user interfaces (VR)*, pages 645–652. IEEE.

Alexander, S., Cotzin, M., Hill, C., Ricciuti Jr, E., and Wendt, G. (1945). Wesleyan university studies of motion sickness: Iii. the effects of various accelerations upon sickness rates. *The Journal of Psychology*, 20(1):3–8.

Anderson, A. P., Mayer, M. D., Fellows, A. M., Cowan, D. R., Hegel, M. T., and Buckey, J. C. (2017). Relaxation with immersive natural scenes presented using virtual reality. *Aerospace medicine and human performance*, 88(6):520–526.

Annerstedt, M., Jönsson, P., Wallergård, M., Johansson, G., Karlson, B., Grahn, P., Hansen, Å. M., and Währborg, P. (2013). Inducing physiological stress recovery with sounds of nature in a virtual reality forestâresults from a pilot study. *Physiology & behavior*, 118:240–250.

Bakker, N. H., Passenier, P. O., and Werkhoven, P. J. (2003). Effects of head-slaved navigation and the use of teleports on spatial orientation in virtual environments. *Human factors*, 45(1):160–169.

Banton, T., Stefanucci, J., Durgin, F., Fass, A., and Proffitt, D. (2005). The perception of walking speed in a virtual environment. *Presence: Teleoperators & Virtual Environments*, 14(4):394–406.

Barliya, A., Omlor, L., Giese, M. A., Berthoz, A., and Flash, T. (2013). Expression of emotion in the kinematics of locomotion. *Experimental brain research*, 225(2):159–176.

Berge, S. H., Hagenzieker, M., Farah, H., and de Winter, J. (2022). Do cyclists need hmis in future automated traffic? an interview study. *Transportation Research Part F: Traffic Psychology and Behaviour*, 84:33–52.

Bhandari, J., MacNeilage, P. R., and Folmer, E. (2018). Teleportation without spatial disorientation using optical flow cues. In *Graphics interface*, pages 162–167.

Bolas, M., Jones, J. A., Mcdowall, I., and Suma, E. (2017). Dynamic field of view throttling as a means of improving user experience in head mounted virtual environments. US Patent 9,645,395.

Bonato, F., Bubka, A., and Story, M. (2005). Rotation direction change hastens motion sickness onset in an optokinetic drum. *Aviation, Space, and Environmental Medicine*, 76(9):823–827.

Bos, J. E., de Vries, S. C., van Emmerik, M. L., and Groen, E. L. (2010). The effect of internal and external fields of view on visually induced motion sickness. *Applied ergonomics*, 41(4):516–521.

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.

Bradley, M. M. and Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1):49–59.

Brooks, K. (2001). Stereomotion speed perception is contrast dependent. *Perception*, 30(6):725–731.

Brown, J. F. (1931). The visual perception of velocity. *Psychologische Forschung*, 14(1):199–232.

Brument, H., Marchal, M., Olivier, A.-H., and Argelaguet, F. (2020). Influence of dynamic field of view restrictions on rotation gain perception in virtual environments. In *International Conference on Virtual Reality and Augmented Reality*, pages 20–40. Springer.

Bubka, A., Bonato, F., Urmey, S., and Mycewicz, D. (2006). Rotation velocity change and motion sickness in an optokinetic drum. *Aviation, Space, and Environmental Medicine*, 77(8):811–815.

Budhiraja, P., Miller, M. R., Modi, A. K., and Forsyth, D. (2017). Rotation blurring: use of artificial blurring to reduce cybersickness in virtual reality first person shooters. *arXiv preprint arXiv:1710.02599*.

Calvi, A., DâAmico, F., Ferrante, C., and Ciampoli, L. B. (2020). Effectiveness of augmented reality warnings on driving behaviour whilst approaching pedestrian crossings: A driving simulator study. *Accident Analysis & Prevention*, 147:105760.

Cao, Z., Jerald, J., and Kopper, R. (2018). Visually-induced motion sickness reduction via static and dynamic rest frames. In *2018 IEEE conference on virtual reality and 3D user interfaces (VR)*, pages 105–112. IEEE.

Carnegie, K. and Rhee, T. (2015). Reducing visual discomfort with hmds using dynamic depth of field. *IEEE computer graphics and applications*, 35(5):34–41.

Chardonnet, J.-R., Mirzaei, M. A., and Merienne, F. (2020). Influence of navigation parameters on cybersickness in virtual reality. *Virtual Reality*, pages 1–10.

Chaturvedi, I., Bijarbooneh, F. H., Braud, T., and Hui, P. (2019). Peripheral vision: a new killer app for smart glasses. In *Proceedings of the 24th International Conference on Intelligent User Interfaces*, pages 625–636.

Chen, C.-Y., Chuang, C.-H., Tsai, T.-L., Chen, H.-W., and Wu, P.-J. (2022). Reducing cybersickness by implementing texture blur in the virtual reality content. *Virtual Reality*, 26(2):789–800.

Chen, H., Dey, A., Billinghurst, M., and Lindeman, R. W. (2017). Exploring the design space for multi-sensory heart rate feedback in immersive virtual reality. In *Proceedings of the 29th Australian conference on computer-human interaction*, pages 108–116.

Chen, J., Pyla, P. S., and Bowman, D. A. (2004). Testbed evaluation of navigation and text display techniques in an information-rich virtual environment. In *IEEE Virtual Reality 2004*, pages 181–289. IEEE.

Chidsin, W., Gu, Y., and Goncharenko, I. (2021). Ar-based navigation using rgb-d camera and hybrid map. *Sustainability*, 13(10):5585.

Chong, S., Poulos, R., Olivier, J., Watson, W. L., and Grzebieta, R. (2010). Relative injury severity among vulnerable non-motorised road users: comparative analysis of injury arising from bicycle–motor vehicle and bicycle–pedestrian collisions. *Accident Analysis & Prevention*, 42(1):290–296.

Dahlman, J., Sjörs, A., Ledin, T., and Falkmer, T. (2008). Could sound be used as a strategy for reducing symptoms of perceived motion sickness? *Journal of NeuroEngineering and Rehabilitation*, 5(1):1–9.

Dancu, A., Vechev, V., Ünlüer, A. A., Nilson, S., Nygren, O., Eliasson, S., Barjonet, J.-E., Marshall, J., and Fjeld, M. (2015). Gesture bike: examining projection surfaces and turn signal systems for urban cycling. In *Proceedings of the 2015 international conference on interactive tabletops & surfaces*, pages 151–159.

Darken, R. P. and Sibert, J. L. (1993). A toolset for navigation in virtual environments. In *Proceedings of the 6th annual ACM symposium on User interface software and technology*, pages 157–165.

Davis, E. T., Scott, K., Pair, J., Hodges, L. F., and Oliverio, J. (1999). Can audio enhance visual perception and performance in a virtual environment? In *Proceedings of the human factors and ergonomics society annual meeting*, volume 43, pages 1197–1201. SAGE Publications Sage CA: Los Angeles, CA.

De Waard, D., Lewis-Evans, B., Jelijs, B., Tucha, O., and Brookhuis, K. (2014). The effects of operating a touch screen smartphone and other common activities performed while bicycling on cycling behaviour. *Transportation research part F: traffic psychology and behaviour*, 22:196–206.

Diels, C. (2014). Will autonomous vehicles make us sick. *Contemporary ergonomics and human factors. Taylor & Francis*, pages 301–307.

Diels, C. and Parkes, A. (2010). Geometric field of view manipulations affect perceived speed in driving simulators. *Annual Research*, page 55.

Duan, H., Zhai, G., Min, X., Zhu, Y., Sun, W., and Yang, X. (2017). Assessment of visually induced motion sickness in immersive videos. In *Pacific Rim Conference on Multimedia*, pages 662–672. Springer.

Dużmańska, N., Strojny, P., and Strojny, A. (2018). Can simulator sickness be avoided? a review on temporal aspects of simulator sickness. *Frontiers in psychology*, 9:2132.

Ewert, J.-P. and Hock, F. (1972). Movement-sensitive neurones in the toad's retina. *Experimental Brain Research*, 16(1):41–59.

Falanga, D., Mueggler, E., Faessler, M., and Scaramuzza, D. (2017). Aggressive quadrotor flight through narrow gaps with onboard sensing and computing using active vision. In *2017 IEEE international conference on robotics and automation (ICRA)*, pages 5774–5781. IEEE.

Fan, K., Huber, J., Nanayakkara, S., and Inami, M. (2014). Spidervision: extending the human field of view for augmented awareness. In *Proceedings of the 5th augmented human international conference*, pages 1–8.

Fernandes, A. S. and Feiner, S. K. (2016). Combating vr sickness through subtle dynamic field-of-view modification. In *2016 IEEE symposium on 3D user interfaces (3DUI)*, pages 201–210. IEEE.

Fukatsu, S., Kitamura, Y., Masaki, T., and Kishino, F. (1998). Intuitive control of âbird's eyeâ overview images for navigation in an enormous virtual environment. In *Proceedings of the ACM symposium on Virtual reality software and technology*, pages 67–76.

Gabbard, J. L., Fitch, G. M., and Kim, H. (2014). Behind the glass: Driver challenges and opportunities for ar automotive applications. *Proceedings of the IEEE*, 102(2):124–136.

Gavrila, D. M. (2000). Pedestrian detection from a moving vehicle. In *European conference on computer vision*, pages 37–49. Springer.

Gerstweiler, G., Vonach, E., and Kaufmann, H. (2015). Hymotrack: A mobile ar navigation system for complex indoor environments. *Sensors*, 16(1):17.

Ghiurău, F.-T., Baytaş, M. A., and Wickman, C. (2020). Arcar: On-road driving in mixed reality by volvo cars. In *Adjunct Publication of the 33rd Annual ACM Symposium on User Interface Software and Technology*, pages 62–64.

Gibson, E. J., Gibson, J. J., Smith, O. W., and Flock, H. (1959). Motion parallax as a determinant of perceived depth. *Journal of experimental psychology*, 58(1):40.

Ginters, E. (2019). Augmented reality use for cycling quality improvement. *Procedia Computer Science*, 149:167–176.

Golding, J. F., Doolan, K., Acharya, A., Tribak, M., and Gresty, M. A. (2012). Cognitive cues and visually induced motion sickness. *Aviation, space, and environmental medicine*, 83(5):477–482.

Gómez-Jordana, L. I., Stafford, J., Peper, C. E., and Craig, C. M. (2018). Virtual footprints can improve walking performance in people with parkinson's disease. *Frontiers in neurology*, 9:681.

Gruenefeld, U., Stratmann, T. C., Jung, J., Lee, H., Choi, J., Nanda, A., and Heuten, W. (2018). Guiding smombies: Augmenting peripheral vision with low-cost glasses to shift the attention of smartphone users. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 127–131. IEEE.

Hänsell, A., Lenggenhagerl, B., von Känell, R., Curatolol, M., and Blankel, O. (2011). Seeing and identifying with a virtual body decreases pain perception. *European journal of pain*, 15(8):874–879.

Hariyono, J., Hoang, V.-D., and Jo, K.-H. (2014). Moving object localization using optical flow for pedestrian detection from a moving vehicle. *The Scientific World Journal*, 2014.

He, J. and Zhao, H. (2020). A new smart safety navigation system for cycling based on audio technology. *Safety science*, 124:104583.

Hecht, T., Weng, S., Kick, L.-F., and Bengler, K. (2022). How users of automated vehicles benefit from predictive ambient light displays. *Applied ergonomics*, 103:103762.

Held, R. and Bossom, J. (1961). Neonatal deprivation and adult rearrangement: complementary techniques for analyzing plastic sensory-motor coordinations. *Journal of comparative and physiological Psychology*, 54(1):33.

Helmholtz, H. (1925). Treatise on psychological optics. *Optical Society of America*, 3:482.

Hillaire, S., Lécuyer, A., Cozot, R., and Casiez, G. (2007). Depth-of-field blur effects for first-person navigation in virtual environments. In *Proceedings of the 2007 ACM symposium on Virtual reality software and technology*, pages 203–206.

Hillaire, S., Lécuyer, A., Cozot, R., and Casiez, G. (2008). Depth-of-field blur effects for first-person navigation in virtual environments. *IEEE computer graphics and applications*, 28(6):47–55.

Jäger, J. and Henn, V. (1981). Vestibular habituation in man and monkey during sinusoidal rotation. *Annals of the New York Academy of Sciences*, 374:330–339.

Janaka, N., Haigh, C., Kim, H., Zhang, S., and Zhao, S. (2022). Paracentral and near-peripheral visualizations: Towards attention-maintaining secondary information presentation on ohmds during in-person social interactions. In *CHI Conference on Human Factors in Computing Systems*, pages 1–14.

Jenkins, M. P. and Young, D. (2016). Barracuda: An augmented reality display for increased motorcyclist en route hazard awareness. In *2016 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, pages 68–72. IEEE.

Jeon, W. and Rajamani, R. (2017). Rear vehicle tracking on a bicycle using active sensor orientation control. *IEEE Transactions on Intelligent Transportation Systems*, 19(8):2638–2649.

Jia, J., Elezovikj, S., Fan, H., Yang, S., Liu, J., Guo, W., Tan, C. C., and Ling, H. (2021). Semantic-aware label placement for augmented reality in street view. *The Visual Computer*, 37(7):1805–1819.

Jones, J. A., Swan II, J. E., and Bolas, M. (2013). Peripheral stimulation and its effect on perceived spatial scale in virtual environments. *IEEE transactions on visualization and computer graphics*, 19(4):701–710.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., and Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220.

Keshavarz, B. and Hecht, H. (2014). Pleasant music as a countermeasure against visually induced motion sickness. *Applied ergonomics*, 45(3):521–527.

Keshavarz, B., Hecht, H., and Zschutschke, L. (2011). Intra-visual conflict in visually induced motion sickness. *Displays*, 32(4):181–188.

Kimura, R., Matsunaga, N., Okajima, H., and Koutaki, G. (2017). Driving assistance system for welfare vehicle using virtual platoon control with augmented reality. In *2017 56th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*, pages 980–985. IEEE.

Kitagawa, T. and Kondo, K. (2018). Evaluation of a wind noise reduction method using dnn for bicycle audio augmented reality systems. In *2018*

IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW), pages 1–2. IEEE.

Kitagawa, T. and Kondo, K. (2019). Detailed evaluation of a wind noise reduction method using dnn for 3d audio navigation system audio augmented reality for bicycles. In *2019 IEEE 8th Global Conference on Consumer Electronics (GCCE)*, pages 863–864. IEEE.

Kunze, A., Summerskill, S. J., Marshall, R., and Filtness, A. J. (2019). Conveying uncertainties using peripheral awareness displays in the context of automated driving. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pages 329–341.

Labayrade, R., Aubert, D., and Tarel, J.-P. (2002). Real time obstacle detection in stereovision on non flat road geometry through" v-disparity" representation. In *Intelligent Vehicle Symposium, 2002. IEEE*, volume 2, pages 646–651. IEEE.

Langbehn, E., Raupp, T., Bruder, G., Steinicke, F., Bolte, B., and Lappe, M. (2016). Visual blur in immersive virtual environments: Does depth of field or motion blur affect distance and speed estimation? In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, pages 241–250.

Langlois, S. and Soualmi, B. (2016). Augmented reality versus classical hud to take over from automated driving: An aid to smooth reactions and to anticipate maneuvers. In *2016 IEEE 19th international conference on intelligent transportation systems (ITSC)*, pages 1571–1578. IEEE.

LaViola Jr, J. J. (2000). A discussion of cybersickness in virtual environments. *ACM Sigchi Bulletin*, 32(1):47–56.

Liang, X., Zhang, T., Xie, M., and Jia, X. (2021). Analyzing bicycle level of service using virtual reality and deep learning technologies. *Transportation research part A: policy and practice*, 153:115–129.

Lim, K., Lee, J., Won, K., Kala, N., and Lee, T. (2021). A novel method for vr sickness reduction based on dynamic field of view processing. *Virtual Reality*, 25(2):331–340.

Lin, J.-H. T. (2017). Fear in virtual reality (vr): Fear elements, coping reactions, immediate and next-day fright responses toward a survival horror zombie virtual reality game. *Computers in Human Behavior*, 72:350–361.

Lin, J.-W., Duh, H. B.-L., Parker, D. E., Abi-Rached, H., and Furness, T. A. (2002). Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment. In *Proceedings ieee virtual reality 2002*, pages 164–171. IEEE.

Lin, Y.-X., Venkatakrishnan, R., Venkatakrishnan, R., Ebrahimi, E., Lin, W.-C., and Babu, S. V. (2020). How the presence and size of static peripheral blur affects cybersickness in virtual reality. *ACM Transactions on Applied Perception (TAP)*, 17(4):1–18.

Lindemann, P. and Rigoll, G. (2017). Examining the impact of see-through cockpits on driving performance in a mixed reality prototype. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications Adjunct*, pages 83–87.

Liu, C., Plopski, A., and Orlosky, J. (2020a). Orthogaze: Gaze-based three-dimensional object manipulation using orthogonal planes. *Computers & Graphics*, 89:1–10.

Liu, S.-H., Yen, P.-C., Mao, Y.-H., Lin, Y.-H., Chandra, E., and Chen, M. Y. (2020b). Headblaster: a wearable approach to simulating motion perception using head-mounted air propulsion jets. *ACM Transactions on Graphics (TOG)*, 39(4):84–1.

Longuet-Higgins, H. C. and Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 208(1173):385–397.

Lou, R. and Chardonnet, J.-R. (2019). Reducing cybersickness by geometry deformation. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 1058–1059. IEEE.

Lucero, A. and Vetek, A. (2014). Notifeye: using interactive glasses to deal with notifications while walking in public. In *Proceedings of the 11th Conference on Advances in Computer Entertainment Technology*, pages 1–10.

Luyten, K., Degraen, D., Rovelo Ruiz, G., Coppers, S., and Vanacken, D. (2016). Hidden in plain sight: an exploration of a visual language for near-eye out-of-focus displays in the peripheral view. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 487–497.

Lynch, T. and Martins, N. (2015). Nothing to fear? an analysis of college students' fear experiences with video games. *Journal of Broadcasting & Electronic Media*, 59(2):298–317.

Mantell, J., Rod, J., Kage, Y., Delmotte, F., and Leu, J. (2010). Navinko: Audio augmented reality-enabled social navigation for city cyclists. In *Programme, Workshop Pervasive 2010 Proc.*

Matsunaga, N., Kimura, R., Ishiguro, H., and Okajima, H. (2018). Driving assistance of welfare vehicle with virtual platoon control method which has collision avoidance function using mixed reality. In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 1915–1920. IEEE.

Merenda, C., Kim, H., Tanous, K., Gabbard, J. L., Feichtl, B., Misu, T., and Suga, C. (2018). Augmented reality interface design approaches for goal-directed and stimulus-driven driving tasks. *IEEE transactions on visualization and computer graphics*, 24(11):2875–2885.

Min, B.-C., Chung, S.-C., Min, Y.-K., and Sakamoto, K. (2004). Psychophysiological evaluation of simulator sickness evoked by a graphic simulator. *Applied ergonomics*, 35(6):549–556.

Murch, G. M. (1984). Physiological principles for the effective use of color. *IEEE Computer Graphics and Applications*, 4(11):48–55.

Nataprawira, J., Gu, Y., Asami, K., and Goncharenko, I. (2020). Pedestrian detection in different lighting conditions using deep neural networks. In *IICST*, pages 97–104.

Nie, G.-Y., Duh, H. B.-L., Liu, Y., and Wang, Y. (2019). Analysis on mitigation of visually induced motion sickness by applying dynamical blurring on a user's retina. *IEEE transactions on visualization and computer graphics*, 26(8):2535–2545.

Niforatos, E., Fedosov, A., Elhart, I., and Langheinrich, M. (2017). Augmenting skiers' peripheral perception. In *Proceedings of the 2017 ACM International symposium on wearable computers*, pages 114–121.

Nordahl, R. (2005). Self-induced footsteps sounds in virtual reality: Latency, recognition, quality and presence. In *The 8th Annual International Workshop on Presence, PRESENCE 2005*, pages 353–355. Department of Computer Science, Aalborg University.

Nordahl, R., Berrezag, A., Dimitrov, S., Turchet, L., Hayward, V., and Serafin, S. (2010). Preliminary experiment combining virtual reality haptic

shoes and audio synthesis. In *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*, pages 123–129. Springer.

Norouzi, N., Bruder, G., and Welch, G. (2018). Assessing vignetting as a means to reduce vr sickness during amplified head rotations. In *Proceedings of the 15th ACM Symposium on Applied Perception*, pages 1–8.

Oliveira, W., Gaisbauer, W., Tizuka, M., Clua, E., and Hlavacs, H. (2018). Virtual and real body experience comparison using mixed reality cycling environment. In *International Conference on Entertainment Computing*, pages 52–63. Springer.

Oman, C. M. (1982). A heuristic mathematical model for the dynamics of sensory conflict and motion sickness hearing in classical musicians. *Acta Oto-Laryngologica*, 94(sup392):4–44.

Ono, H. and Ujike, H. (2005). Motion parallax driven by head movements: Conditions for visual stability, perceived depth, and perceived concomitant motion. *Perception*, 34(4):477–490.

Ono, S., Ogawara, K., Kagesawa, M., Kawasaki, H., Onuki, M., Abeki, J., Yano, T., Nerio, M., Honda, K., and Ikeuchi, K. (2005). A photo-realistic driving simulation system for mixed-reality traffic experiment space. In *IEEE Proceedings. Intelligent Vehicles Symposium, 2005.*, pages 747–752. IEEE.

Orlosky, J., Kiyokawa, K., and Takemura, H. (2013). Towards intelligent view management: A study of manual text placement tendencies in mobile environments using video see-through displays. In *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 281–282. IEEE.

Orlosky, J., Kiyokawa, K., and Takemura, H. (2014). Managing mobile text in head mounted displays: studies on visual preference and text placement. *ACM SIGMOBILE Mobile Computing and Communications Review*, 18(2):20–31.

Park, H. S., Park, M. W., Won, K. H., Kim, K.-H., and Jung, S. K. (2013). In-vehicle ar-hud system to provide driving-safety information. *ETRI journal*, 35(6):1038–1047.

Peck, T. C., Fuchs, H., and Whitton, M. C. (2009). Evaluation of reorientation techniques and distractors for walking in large virtual environments. *IEEE transactions on visualization and computer graphics*, 15(3):383–394.

Peng, Y., Qu, D., Zhong, Y., Xie, S., Luo, J., and Gu, J. (2015). The obstacle detection and obstacle avoidance algorithm based on 2-d lidar. In *2015 IEEE international conference on information and automation*, pages 1648–1653. IEEE.

Peng, Y.-H., Yu, C., Liu, S.-H., Wang, C.-W., Taele, P., Yu, N.-H., and Chen, M. Y. (2020). Walkingvibe: Reducing virtual reality sickness and improving realism while walking in vr using unobtrusive head-mounted vibrotactile feedback. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–12.

Perrone, J. A. and Stone, L. S. (1994). A model of self-motion estimation within primate extrastriate visual cortex. *Vision research*, 34(21):2917–2938.

Qian, L., Plopski, A., Navab, N., and Kazanzides, P. (2018). Restoring the awareness in the occluded visual field for optical see-through head-mounted displays. *IEEE transactions on visualization and computer graphics*, 24(11):2936–2946.

Rateke, T., Justen, K. A., and Von Wangenheim, A. (2019). Road surface classification with images captured from low-cost camera-road traversing knowledge (rtk) dataset. *Revista de Informática Teórica e Aplicada*, 26(3):50–64.

Reason, J. T. (1978). Motion sickness adaptation: a neural mismatch model. *Journal of the Royal Society of Medicine*, 71(11):819–829.

Reason, J. T. and Brand, J. J. (1975). *Motion sickness.* Academic press.

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788.

Ren, J., Chen, Y., Li, F., Xue, C., Yin, X., Peng, J., Liang, J., Feng, Q., and Wang, S. (2021). Road injuries associated with cellular phone use while walking or riding a bicycle or an electric bicycle: a case-crossover study. *American Journal of Epidemiology*, 190(1):37–43.

Riecke, B. E., Schulte-Pelkum, J., Avraamides, M. N., Heyde, M. V. D., and Bülthoff, H. H. (2006). Cognitive factors can influence self-motion perception (vection) in virtual reality. *ACM Transactions on Applied Perception (TAP)*, 3(3):194–216.

Robillard, G., Bouchard, S., Fournier, T., and Renaud, P. (2003). Anxiety and presence during vr immersion: A comparative study of the reactions of phobic and non-phobic participants in therapeutic virtual environments derived from computer games. *CyberPsychology & Behavior*, 6(5):467–476.

Runeson, S. (1974). Constant velocityânot perceived as such. *Psychological Research*, 37(1):3–23.

Russell, M. E. B., Hoffman, B., Stromberg, S., and Carlson, C. R. (2014). Use of controlled diaphragmatic breathing for the management of motion sickness in a virtual reality environment. *Applied psychophysiology and biofeedback*, 39(3-4):269–277.

Schoop, E., Smith, J., and Hartmann, B. (2018). Hindsight: enhancing spatial awareness by sonifying detected objects in real-time 360-degree video. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–12.

Schweigkofler, A., Monizza, G. P., Domi, E., Popescu, A., Ratajczak, J., Marcher, C., Riedl, M., and Matt, D. (2018). Development of a digital platform based on the integration of augmented reality and bim for the management of information in construction processes. In *IFIP International Conference on Product Lifecycle Management*, pages 46–55. Springer.

Seay, A. F., Krum, D. M., Hodges, L., and Ribarsky, W. (2001). Simulator sickness and presence in a high fov virtual environment. In *Proceedings IEEE Virtual Reality 2001*, pages 299–300. IEEE.

Shashua, A., Gdalyahu, Y., and Hayun, G. (2004). Pedestrian detection for driving assistance systems: Single-frame classification and system level performance. In *IEEE Intelligent Vehicles Symposium, 2004*, pages 1–6. IEEE.

Shiban, Y., Pauli, P., and Mühlberger, A. (2013). Effect of multiple context exposure on renewal in spider phobia. *Behaviour research and therapy*, 51(2):68–74.

Shupak, A. and Gordon, C. R. (2006). Motion sickness: advances in pathogenesis, prediction, prevention, and treatment. *Aviation, space, and environmental medicine*, 77(12):1213–1223.

Slater, M. and Usoh, M. (1994). Body centred interaction in immersive virtual environments. *Artificial life and virtual reality*, 1(1994):125–148.

Stanney, K. M. and Kennedy, R. S. (1998). Aftereffects from virtual environment exposure: How long do they last? In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 42, pages 1476–1480. SAGE Publications Sage CA: Los Angeles, CA.

Stone, L. S. and Thompson, P. (1992). Human speed perception is contrast dependent. *Vision research*, 32(8):1535–1549.

Suma, E. A., Clark, S., Krum, D., Finkelstein, S., Bolas, M., and Warte, Z. (2011). Leveraging change blindness for redirection in virtual environments. In *2011 IEEE Virtual Reality Conference*, pages 159–166. IEEE.

Tajadura-Jiménez, A., Väljamäe, A., and Västfjäll, D. (2008). Self-representation in mediated environments: the experience of emotions modulated by auditory-vibrotactile heartbeat. *CyberPsychology & Behavior*, 11(1):33–38.

Terzano, K. (2013). Bicycling safety and distracted behavior in the hague, the netherlands. *Accident Analysis & Prevention*, 57:87–90.

Thompson, P. (1982). Perceived rate of movement depends on contrast. *Vision research*, 22(3):377–380.

Tran, C., Bark, K., and Ng-Thow-Hing, V. (2013). A left-turn driving aid using projected oncoming vehicle paths with augmented reality. In *Proceedings of the 5th international conference on automotive user interfaces and interactive vehicular applications*, pages 300–307.

Truong-Allié, C., Paljic, A., Roux, A., and Herbeth, M. (2021). User behavior adaptive ar guidance for wayfinding and tasks completion. *Multimodal Technologies and Interaction*, 5(11):65.

Uchiya, T. and Futamura, R. (2021). Consideration of presentation timing in bicycle navigation using smart glasses. In *International Conference on Intelligent Networking and Collaborative Systems*, pages 209–217. Springer.

Uchiyama, I., Anderson, D. I., Campos, J. J., Witherington, D., Frankel, C. B., Lejeune, L., and Barbu-Roth, M. (2008). Locomotor experience affects self and emotion. *Developmental psychology*, 44(5):1225.

Väljamäe, A., Larsson, P., Västfjäll, D., and Kleiner, M. (2008). Sound representing self-motion in virtual environments enhances linear vection. *Presence: Teleoperators and Virtual Environments*, 17(1):43–56.

Van der Horst, A. R. A., de Goede, M., de Hair-Buijssen, S., and Methorst, R. (2014). Traffic conflicts on bicycle paths: A systematic observation of behaviour from video. *Accident Analysis & Prevention*, 62:358–368.

van Lopik, K., Schnieder, M., Sharpe, R., Sinclair, M., Hinde, C., Conway, P., West, A., and Maguire, M. (2020). Comparison of in-sight and handheld navigation devices toward supporting industry 4.0 supply chains: First and last mile deliveries at the human level. *Applied ergonomics*, 82:102928.

Vansteenkiste, P., Cardon, G., DâHondt, E., Philippaerts, R., and Lenoir, M. (2013). The visual control of bicycle steering: The effects of speed and path width. *Accident Analysis & Prevention*, 51:222–227.

Vansteenkiste, P., Cardon, G., and Lenoir, M. (2015). Visual guidance during bicycle steering through narrow lanes: A study in children. *Accident Analysis & Prevention*, 78:8–13.

von Sawitzky, T., Wintersberger, P., Löcken, A., Frison, A.-K., and Riener, A. (2020). Augmentation concepts with huds for cyclists to improve road safety in shared spaces. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–9.

Waldin, N., Waldner, M., and Viola, I. (2017). Flicker observer effect: Guiding attention through high frequency flicker in images. In *Computer Graphics Forum*, volume 36, pages 467–476. Wiley Online Library.

Wang, H., Li, Y., and Wang, S. (2019). Fast pedestrian detection with attention-enhanced multi-scale rpn and soft-cascaded decision trees. *IEEE Transactions on Intelligent Transportation Systems*, 21(12):5086–5093.

Wang, S., Cheng, J., Liu, H., Wang, F., and Zhou, H. (2018). Pedestrian detection via body part semantic and contextual information with dnn. *IEEE Transactions on Multimedia*, 20(11):3148–3159.

Ware, C. and Mikaelian, H. H. (1986). An evaluation of an eye tracker as a device for computer input. *SIGCHI Bull.*, 17(SI):183–188.

Webb, N. A. and Griffin, M. J. (2003). Eye movement, vection, and motion sickness with foveal and peripheral vision. *Aviation, space, and environmental medicine*, 74(6):622–625.

Weech, S., Moon, J., and Troje, N. F. (2018). Influence of bone-conducted vibration on simulator sickness in virtual reality. *PloS one*, 13(3):e0194137.

Wist, E. R., Diener, H., Dichgans, J., and Brandt, T. (1975). Perceived distance and the perceived speed of self-motion: linear vs. angular velocity? *Perception & Psychophysics*, 17(6):549–554.

Wu, A., Zhang, W., Hu, B., and Zhang, X. (2007). Evaluation of wayfinding aids interface in virtual environment. In *International Conference on Human-Computer Interaction*, pages 700–709. Springer.

Wu, F. and Rosenberg, E. S. (2019). Combining dynamic field of view modification with physical obstacle avoidance. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 1882–1883. IEEE.

Wu, T., Sachdeva, E., Akash, K., Wu, X., Misu, T., and Ortiz, J. (2022). Toward an adaptive situational awareness support system for urban driving. In *2022 IEEE Intelligent Vehicles Symposium (IV)*, pages 1073–1080. IEEE.

Yamashita, T., Hashimoto, W., Nishiguchi, S., and Mizutani, Y. (2021). Presenting a sense of self-motion by transforming the rendering area based on the movement of the userâs viewpoint. In *International Conference on Human-Computer Interaction*, pages 410–417. Springer.

Yeo, D., Kim, G., and Kim, S. (2020). Toward immersive self-driving simulations: Reports from a user study across six platforms. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–12.

Young, L. R. (1973). On visual-vestibular interaction. In *Fifth symposium on the role of the vestibular organs in space exploration*, pages 205–210.

Zhao, Y., Kupferstein, E., Castro, B. V., Feiner, S., and Azenkot, S. (2019). Designing ar visualizations to facilitate stair navigation for people with low vision. In *Proceedings of the 32nd annual ACM symposium on user interface software and technology*, pages 387–402.

Zielasko, D., Meißner, A., Freitag, S., Weyers, B., and Kuhlen, T. W. (2018). Dynamic field of view reduction related to subjective sickness measures in an hmd-based data analysis task. In *IEEE 4th Workshop on Everyday Virtual Reality (WEVR)*.

Zielasko, D., Weyers, B., and Kuhlen, T. W. (2019). A non-stationary office desk substitution for desk-based and hmd-projected virtual reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 1884–1889. IEEE.