



Università
di **Genova**

DIPARTIMENTO DI
INFORMATICA, BIOINGEGNERIA,
ROBOTICA E INGEGNERIA DEI SISTEMI

Towards Achieving Natural Visual Experience in Immersive Virtual Environments: Addressing Cybersickness and Depth Perception

Razeen Hussain

Università di **Genova**

Dipartimento di Informatica, Bioingegneria,
Robotica ed Ingegneria dei Sistemi

Ph.D. Thesis in
Computer Science and Systems Engineering
Computer Science Curriculum

**Towards Achieving Natural Visual Experience
in Immersive Virtual Environments:
Addressing Cybersickness and Depth
Perception**

by

Razeen Hussain

July, 2022

Ph.D. Thesis in Computer Science and Systems Engineering (S.S.D. INF/01)
Dipartimento di Informatica, Bioingegneria,
Robotica ed Ingegneria dei Sistemi
Università di Genova

Candidate

Razeen Hussain
razeen.hussain@edu.unige.it

Title

Towards Achieving Natural Visual Experience in Immersive Virtual Environments: Addressing Cybersickness and Depth Perception

Advisors

Fabio Solari
DIBRIS, Università di Genova, Italy
fabio.solari@unige.it

Manuela Chessa
DIBRIS, Università di Genova, Italy
manuela.chessa@unige.it

External Reviewers

Lucile Sassatelli
Université Côte d'Azur, France
lucile.sassatelli@univ-cotedazur.fr

Bruno Herbelin
École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
bruno.herbelin@epfl.ch

Location

DIBRIS, Università di Genova
Via Dodecaneso, 35
I-16146 Genova, Italy

Submitted On

July 2022

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements.

Genova, Italy
July 2022

Razeen Hussain

Abstract

Recent years have seen the virtual reality field flourish with the introduction of low cost VR devices. However, many problematics still remain such as heavy processing requirements, hardware limitations and the need for portability. Due to these, users tend to feel discomfort after long exposure to a VR device. These issues need to be addressed to achieve the full potential VR technology possesses. To this aim, this work addresses some of the perceptual issues prevalent in the modern VR technology. Visual stimuli plays an important role in how users perceive and interact with the virtual world. However, there are many discrepancies between how humans visually perceive in the real world and how they perceive in the virtual world. In VR devices, the stimuli is presented in pin-sharp focus which is different to how the stimuli is processed in the real world where humans convergence their eyes and alter their accommodation to focus on objects in the environment.

To address perceptual discrepancies in VR systems, we have developed frameworks that alter the visual stimuli. The first system aims to introduce space-variant blurring in VR environments. The technique developed takes inspiration from the human visual system and adopts a hybrid approach to introduce multi-region foveation and depth-of-field effects into the rendering pipeline. The system can be integrated to any VR device and application. The effectiveness of the technique was validated through a user study on cybersickness. Users were exposed to a virtual environment for a fixed duration of time and the induced level of cybersickness was measured through self-assessment questionnaires and physiological signals. Overall, the system was able to reduce cybersickness levels by 66%. Other factors such as gender and age were also evaluated.

Depth perception is a crucial part of how humans behave and interact with their environment. Convergence and accommodation are two important depth cues. However, when humans are immersed into the virtual environments, they experience a mismatch between these cues. This mismatch causes users to feel a discomfort while also hindering their ability to fully perceive object distances. To address the conflict, a system was developed which encompasses inverse blurring into immersive media devices. The inverse blurring system utilizes the classical Wiener deconvolution approach. The system's ability to improve depth perception was verified through two user studies aimed at reaching and spatial awareness respectively. The two studies yielded a 36% and 48% improvement in user performance respectively.

Overall, the research aimed at bridging the gap between visual perception in the real world and the virtual one. The work done demonstrates how visual stimuli can be modified to allow users to achieve a more natural interaction with the virtual environment.

Publications

Some ideas and figures have appeared previously in the following publications:

- **Razeen Hussain**, Manuela Chessa and Fabio Solari. 'Mitigating Cyber-sickness in Virtual Reality Systems through Foveated Depth-of-Field Blur'. In: *Sensors* 21:12:4006. 2021. DOI: 10.3390/s21124006.
- Aneeq Zia, Kiran D. Bhattacharyya, Xi Liu, Ziheng Wang, Satoshi Kondo, Emanuele Colleoni, Beatrice van Amsterdam, **Razeen Hussain**, Raabid Hussain, Lena Maier-Hein, Danail Stoyanov, Stefanie Speidel and Anthony M. Jarc. 'Surgical Visual Domain Adaptation: Results from the MICCAI 2020 SurgVisDom Challenge'. In: *ArXiv:2102.13644*. 2021. DOI: 10.48550/arXiv.2102.13644.
- **Razeen Hussain**, Manuela Chessa and Fabio Solari. 'Modelling Foveated Depth-of-field Blur for Improving Depth Perception in Virtual Reality'. In: *IEEE 4th International Conference on Image Processing, Applications and Systems (IPAS)*. pp. 71–76. 2020. DOI: 10.1109/IPAS50080.2020.9334947.
- **Razeen Hussain**, Fabio Solari and Manuela Chessa. 'Simulated foveated depth-of-field blur for virtual reality systems'. In: *16th ACM SIGGRAPH European Conference on Visual Media Production (CVMP)*. London, UK. 2019. DOI: 10.13140/RG.2.2.35437.31208. URL: <https://www.cvm-conference.org/files/2019/short/24.pdf>
- **Razeen Hussain**, Fabio Solari and Manuela Chessa. 'A Method based on Inverse Blurring to Mitigate Vergence Accommodation Conflict in Immersive Media Devices'. Patent. 2022. [Under Review]

Acknowledgements

First and foremost, I would like to thank the Almighty for giving me the strength to tackle various obstacles that occurred during the project.

I would like to express my gratitude to my thesis supervisors, Fabio Solari and Manuela Chessa, for their continuous support throughout the thesis. Their guidance was essential in completing the thesis goals. They also gave me freedom to evolve and make my own decisions which helped me develop as a researcher. Thank you for all the advice, ideas and constructive feedback.

I would like to show my appreciation to my lab colleagues, Giorgio Ballestin, Chiara Bassano and Eros Viola. Thank you for all the valuable discussions we had regarding the project. This work would not have been possible without your encouragement and insightful comments.

I would also like to thank all the people who voluntarily participated in the experimental sessions for giving me their time and energy without receiving any reward. A shout out also to the Pakistani student community in Genoa for making my time in Genoa enjoyable.

Lastly, I would like to acknowledge the support shown by my family throughout the course of the project. Their motivational words helped me work harder, especially during some setbacks caused by the Covid-19 pandemic.

Contents

| | | |
|------------|---|-----------|
| I | CONTEXT | 1 |
| 1 | INTRODUCTION | 2 |
| 1.1 | Motivation | 3 |
| 1.2 | Contribution | 4 |
| 1.3 | Outline | 6 |
| II | STATE OF THE ART | 7 |
| 2 | BACKGROUND | 8 |
| 2.1 | VR Devices | 8 |
| 2.1.1 | Gaze Tracking in VR | 11 |
| 2.2 | Visual Perception in Virtual Environments | 12 |
| 2.2.1 | Vergence Accommodation Conflict | 13 |
| 2.3 | Cybersickness | 14 |
| 2.3.1 | Causes | 15 |
| 2.3.2 | Assessing Cybersickness | 16 |
| 2.3.3 | Solutions | 16 |
| 2.4 | Spatial Blurring | 18 |
| 2.4.1 | Depth-of-field | 19 |
| 2.4.2 | Foveated Rendering | 20 |
| 2.4.3 | Assessing Image Quality | 22 |
| III | RESEARCH WORK | 25 |
| 3 | ADDRESSING CYBERSICKNESS IN VIRTUAL REALITY SYSTEMS | 26 |
| 3.1 | Foveated Depth-of-field | 26 |
| 3.1.1 | Depth-of-field | 27 |
| 3.1.2 | Foveated Imaging | 29 |
| 3.1.3 | Artifact Removal | 31 |
| 3.2 | Cybersickness Study | 33 |
| 3.2.1 | Experimental Setup | 33 |
| 3.2.2 | Procedure | 34 |
| 3.2.3 | Data Analysis and Results | 37 |
| 3.2.4 | Outcome | 47 |
| 4 | DEPTH PERCEPTION IN VIRTUAL ENVIRONMENTS | 48 |
| 4.1 | Depth Perception Study | 48 |
| 4.1.1 | Experimental Setup | 48 |
| 4.1.2 | Procedure | 49 |
| 4.1.3 | Data Analysis and Results | 50 |
| 4.1.4 | Outcome | 53 |
| 4.2 | Inverse Blurring | 54 |
| 4.2.1 | Wiener Deconvolution | 54 |

| | | |
|-----------|---|-----------|
| 4.2.2 | Implementation and Parameter Tuning | 55 |
| 4.3 | Reaching Experiments | 61 |
| 4.3.1 | Experimental Setup | 61 |
| 4.3.2 | Procedure | 63 |
| 4.3.3 | Data Analysis and Results | 65 |
| 4.3.4 | Outcome | 73 |
| 4.4 | Spatial Awareness Experiments | 74 |
| 4.4.1 | Experimental Setup | 74 |
| 4.4.2 | Procedure | 74 |
| 4.4.3 | Data Analysis and Results | 75 |
| 4.4.4 | Outcome | 79 |
| IV | DISCUSSION | 82 |
| 5 | CONCLUSIONS | 83 |
| 6 | PERSPECTIVE | 87 |
| | BIBLIOGRAPHY | 89 |

List of Figures

| | | |
|-----------|--|----|
| Figure 1 | Virtual reality devices | 8 |
| Figure 2 | Eye calibration step | 12 |
| Figure 3 | Vergence accommodation conflict | 13 |
| Figure 4 | Dynamic field-of-view | 17 |
| Figure 5 | Saliency-based blur | 17 |
| Figure 6 | Visual acuity | 19 |
| Figure 7 | Depth-of-field example | 20 |
| Figure 8 | Multi-region foveation example | 21 |
| Figure 9 | Log-polar mapping | 22 |
| Figure 10 | Foveated depth-of-field process flow | 27 |
| Figure 11 | Circle of confusion | 29 |
| Figure 12 | Depth-of-field effects | 30 |
| Figure 13 | Human field-of-view | 31 |
| Figure 14 | Multi-region foveation | 32 |
| Figure 15 | Foveated depth-of-field effect example | 33 |
| Figure 16 | Roller-coaster outline | 34 |
| Figure 17 | Roller-coaster cycle | 35 |
| Figure 18 | Roller-coaster virtual environment | 36 |
| Figure 19 | SSQ score for NB session | 37 |
| Figure 20 | SSQ score for GC session | 38 |
| Figure 21 | SSQ score for FD session | 38 |
| Figure 22 | Post-Pre SSQ difference | 39 |
| Figure 23 | IPQ scores | 41 |
| Figure 24 | Heart rate observations | 42 |
| Figure 25 | Gaze heatmap | 43 |
| Figure 26 | Saccadic motion histogram | 45 |
| Figure 27 | Age-wise SSQ scores | 46 |
| Figure 28 | Gender-based SSQ scores | 46 |
| Figure 29 | Top-view of the observation table | 49 |
| Figure 30 | Selection interface | 50 |
| Figure 31 | Sketch of the test scenario | 51 |
| Figure 32 | Mean absolute error | 52 |
| Figure 33 | Mean error | 53 |
| Figure 34 | Circular PSF | 55 |
| Figure 35 | Inverse blurring process flow | 57 |
| Figure 36 | Tuning parameters | 58 |
| Figure 37 | Tuning w.r.t. PSNR | 59 |
| Figure 38 | Tuning w.r.t. mean-SSIM | 59 |
| Figure 39 | Fine tuning | 60 |
| Figure 40 | Reaching experiment overview | 61 |

| | | |
|-----------|--|----|
| Figure 41 | Kinect v2 joint hierarchy | 62 |
| Figure 42 | Kinect reference frame | 63 |
| Figure 43 | Distance verification | 63 |
| Figure 44 | Z-error in the reaching task | 66 |
| Figure 45 | X-error in the reaching task | 66 |
| Figure 46 | Y-error in the reaching task | 67 |
| Figure 47 | Euclidean distance-error in the reaching task | 67 |
| Figure 48 | XY plane heatmap during normal viewing | 69 |
| Figure 49 | XY plane heatmap during inverse blurring | 69 |
| Figure 50 | XZ plane heatmap during normal viewing | 70 |
| Figure 51 | XZ plane heatmap during inverse blurring | 70 |
| Figure 52 | YZ plane heatmap during normal viewing | 71 |
| Figure 53 | YZ plane heatmap during inverse blurring | 71 |
| Figure 54 | Symptom questionnaire scores for reaching task during normal viewing | 72 |
| Figure 55 | Symptom questionnaire scores for reaching task during inverse blurring | 72 |
| Figure 56 | Session comparison questionnaire scores for reaching task | 73 |
| Figure 57 | Spatial awareness experiment stimuli | 75 |
| Figure 58 | Accuracy plot | 76 |
| Figure 59 | Error plot | 77 |
| Figure 60 | Discrimination sensitivity plot | 78 |
| Figure 61 | Symptom questionnaire scores for spatial awareness task during normal viewing | 79 |
| Figure 62 | Symptom questionnaire scores for spatial awareness task during inverse blurring | 80 |
| Figure 63 | Session comparison questionnaire scores for spatial aware- ness task | 80 |

List of Tables

| | | |
|----------|--|----|
| Table 1 | Common VR devices | 10 |
| Table 2 | Factors effecting cybersickness in VR | 14 |
| Table 3 | Cybersickness clusters | 16 |
| Table 4 | Wilcoxon rank sum test confidence scores | 37 |
| Table 5 | SSQ cluster scores | 40 |
| Table 6 | Comparison with other techniques | 40 |
| Table 7 | Saccadic motion | 44 |
| Table 8 | Frame rate comparison | 47 |
| Table 9 | User accuracy | 52 |
| Table 10 | Group error | 52 |
| Table 11 | Optimal values of SNR and R | 60 |
| Table 12 | Mean average absolute errors | 65 |
| Table 13 | T-test results | 68 |
| Table 14 | Time to perform the reaching task | 68 |
| Table 15 | Mean performance | 76 |
| Table 16 | Discrimination sensitivity | 77 |
| Table 17 | Time to perform the spatial awareness task | 78 |

Acronyms

| | |
|-------|--|
| 2AFC | Two-Alternative Forced Choice |
| 2D | Two-Dimensional |
| 3D | Three-Dimensional |
| AR | Augmented Reality |
| CI | Confidence Interval |
| CoC | Circle of Confusion |
| DFT | Discrete Fourier Transform |
| DOF | Degrees-of-Freedom |
| DoF | Depth-of-Field |
| EEG | Electroencephalography |
| FD | Foveated Depth-of-field Blur |
| FPS | First Person Shooter |
| FoV | Field-of-View |
| GC | Unity's Post-processing Stack Blur |
| HMD | Head-Mounted Display |
| HSV | Hue-Saturation-Value |
| IDFT | Inverse Discrete Fourier Transform |
| IPD | Interpupillary Distance |
| IPQ | Igroup Presence Questionnaire |
| IQA | Image Quality Assessment |
| ISO | International Organization for Standardization |
| LPF | Low Pass Filter |
| MR | Mixed Reality |
| MS | Motion Sickness |
| MSE | Mean Squared Error |
| NB | No Blur |
| PoV | Point-of-View |
| PSF | Point Spread Function |
| PSNR | Peak Signal-to-Noise Ratio |
| RGB | Red-Green-Blue |
| RGB-D | Red-Green-Blue-Depth |

| | |
|------|---------------------------------------|
| SDK | Software Development Kit |
| SNR | Signal-to-Noise Ratio |
| SS | Simulator Sickness |
| SSIM | Structural Similarity Index Measure |
| SSQ | Simulator Sickness Questionnaire |
| VAC | Vergence Accommodation Conflict |
| VE | Virtual Environment |
| VIF | Visual Information Fidelity |
| VR | Virtual Reality |
| VRSQ | Virtual Reality Symptom Questionnaire |
| XR | Extended Reality |

PART I

Context

The following part aims to introduce the various problems with the current virtual reality technology that will be tackled in the thesis.

1

Introduction

Virtual Reality (VR) technology has seen a major boom in the past decade due to the introduction of affordable commercial Head-Mounted Displays (HMDs) such as the Oculus Rift and the HTC Vive. These HMDs are compact and light-weight allowing portability and ease of installation.

VR is not a new technology. It has existed since the 1960s when the first immersive device, Sensorama was introduced [39]. Initial applications mainly focused on military training, flight simulations and multimedia production. However, their scope has since been expanded and are now widely applied to video games, medical training, education and cinemas.

Virtual reality in essence is the use of a computer technology to generate a simulated environment. However, unlike most traditional user interfaces in which the users view a screen in front of them, VR places the user inside the virtual world. The users are able to interact with 3D objects, but also senses such as vision, touch and hearing can be simulated. Thus, allowing a fully immersive experience.

When addressing VR, it is important to distinguish between immersive and non-immersive technologies. Non-immersive media encompasses the traditional user interface where a screen is placed in front of the user usually through a monitor. The users experience the virtual content without their Field-of-View (FoV) being occluded. This generates the feeling of being involved in the virtual space but not being able to simulate actually being there. On the other case, immersive media utilizes room-filling technologies and/or head-mounted displays which occlude the real world surrounding the users and allows them to not only see the virtual content but also to explore and interact with the virtual world. This offers a greater sense of presence. Examples of immersive media include Augmented Reality (AR), Virtual Reality (VR), Mixed Reality (MR), holopresence and many more.

In case of Three-Dimensional (3D) content, the users are able to see a flat sphere of visual content around them which adapts to the user's Point-of-View (PoV) i.e., position and orientation, managed by 3 Degrees-of-Freedom (DOF). While in the case of Extended Reality (XR) (VR, AR, MR), user experience is mapped onto the physical space allowing a deeper interaction. This is known as 6 DOF. Immersive media can fall into either 3DOF or 6DOF categories.

The immersive media industry is worth billions of dollars and VR shares a significant chunk of the share [2].

The efficacy of VR setups to any area of application can be categorized through two important concepts, namely the sense of presence and sense of immersion [108]. From a technical aspect, immersion is achieved by removing as many real world sensations as possible and replacing them with sensations from the Virtual Environment (VE) [58]. Immersion is intended to make the user feel as if they have left the real world and are now present in the virtual one [114]. This notion of being present is central to VE. Whereas immersion is a technology related objective aspect of VR, presence is a psychological, perceptual and cognitive consequence of immersion. Thus, the sense of presence is a psychological perception of being in the VE [26, 110]. The degree to which the sense of presence resonates with each user will differ based on display and human factors. For example, awareness of physical objects can diminish the subjective illusion while internal factors such as immersion propensity and social factors such as interaction with virtual characters can affect the sense of presence.

The potential VR technology possesses has aroused the interest of a diverse group of researchers ranging from applications to data visualization [25, 81, 89], serious games [18, 29], machine learning datasets [73, 141], modelling and testing complex engineering systems [8, 128, 136], flight and military simulations [67, 78], computer-assisted surgical training and procedures [46, 68, 139], physical rehabilitation [45, 65, 118], sports [126], and human psychology [104, 134].

The advent of VR technology has opened up a broad set of possibilities that can address real world problems without involving the life threatening exercises that are posed in the real world. However, the current VR technology is far from perfect. Although there are various studies and corresponding guidelines found in literature on how to design virtual setups, there are still some limitations posed by the spatial and perceptual limitations of the current generations of VR HMDs.

How humans perceive the real world is different to how they perceive the virtual world. The aim of this research is to highlight some of the perceptual and interaction issues faced by the immersive media community and possible solutions to tackle them. The context is in the research setup, however, the results found can be generalized to any immersive media application ranging from training to entertainment.

1.1 Motivation

An important factor considered during the VR application design process is how users will perceive the stimuli. Stimuli can be visual, auditory or olfactory. Due to the hardware limitations of common HMDs, olfactory stimuli are rarely available in immersive media technologies. On the other hand, visual and auditory stimuli are present in most consumer devices. However, visual

stimuli dominates the user experience when it comes to XR. Current AR/VR technology presents the visual stimuli on a screen placed at a fixed distance from the user's eyes. The users converge or diverge their eyes to focus at different objects in the scene. However, the lens in the eye does not contract or expand since the accommodation remains constant. This leads to mismatching visual cues and leads to visual fatigue or cybersickness. Hardware-based solutions to such issues are present, however, these solutions are not compact enough to be adopted by the consumer industry. So, software-based solutions need to be found.

With the recent introduction of eye tracking to consumer VR technology, a new direction has been opened on how users interact with the virtual world. It is now possible to identify areas generating user attention at each given moment of time. This allows VR application developers and researchers to modify the visual stimuli in real-time by highlighting objects that require user attention or simply by rendering specific areas of the screen to where the user is looking.

VR has the advantage of offering a well-controlled experimental setup while still giving the subject freedom of movement and placing it in a relatively natural environment. It is possible for the subject to look in all directions by moving the head, just like in the real world. This opens up a whole a new avenue of research and training. Experimental sessions such as those pertaining to behavioural psychology can be conducted where the user is exposed to various scenarios and their behaviours are studied. VR allows even life threatening scenarios to be presented as an example without posing any threat to human life. Applications such as rehabilitation therapy allows patients to recover their cognitive abilities. In order for this wide range of application domains to be widely adopted, it is crucial that the problematics of current VR technology be addressed.

In this research, I am planning to explore how humans perceive visual stimuli in immersive media technologies. More specifically, how the virtual content can be adapted to provide a real world like experience. The inspiration of the work yields from the understanding of the human visual system and how visual perception works in the real world. There are various ways and tools to achieve a natural interaction in VR, however, the focus of the work in this thesis is on visual perception in VEs.

1.2 Contribution

The main contributions of this thesis consists in the development of a software-based solution for addressing cybersickness in VR systems and in the development of a system that minimizes the conflict between convergence and accommodation in immersive media technologies with the aim of improving depth perception.

The first system is inspired from the human visual system. The system aims to introduce spatial blurring effects into VR setups. Spatial resolution in human

vision is not constant. It is highest in the foveal region and degrades as it moves towards the periphery. For VR systems to one day fully compete with or replace real world experience, it should be able to provide visual cues similar to the real world. The developed system uses a hybrid approach to introduce artifact free spatial blurring for VR systems. The approach combines foveation and Depth-of-Field (DoF) effects, both processes present in nature. The system was integrated with an eye tracking setup to provide gaze-contingency.

The system was then validated through a user study on cybersickness. The aim was to understand whether the spatial blurring setup can help mitigate the onset of cybersickness. The user study was conducted on a VR headset. A virtual roller-coaster environment was used to induce cybersickness to users. Three experimental conditions were considered for comparison. The first was with the developed foveated DoF effects enabled, the second one was with the Unity in-built DoF effect and the third one was the normal viewing condition (without any effects).

Another user study was conducted on the developed foveated DoF effects. This study focused on depth perception in VE. The users were shown cluttered environments and had to identify objects at the same distance to a reference object. Two experimental conditions were considered, one with the spatial blurring enable and the other with it disabled.

The second system addressed the widely known issue of Vergence Accommodation Conflict (VAC). While the first system developed a blurring technique, this system approached the problem in the inverse configuration. Blur is naturally present in human visual system. If we know how much blur is present and apply its inverse to the visual stimuli, we can potentially remove the effects of the natural blur present. The natural blur gives mismatching cues to the users in XR systems. An inverse blurring technique was integrated into the system. The system parameters were tuned through a rigorous image quality assessment procedure.

We then developed two experimental sessions to investigate the performance and user experience of the developed solution to improve depth perception. The first experimental session was based on a reaching task. The user was asked to reach a series of target positions shown using a 3D screen. A Kinect v2 was used to measure the finger positions. Two experimental conditions were considered, one in which our solution was enabled while the other was the normal viewing experience.

The second experimental session was about spatial awareness. The setup and experimental conditions were similar to the reaching experiment. In this experiment, the user was shown a series of paired virtual objects and asked to identify which object was closer to the user. The task was based on the Two-Alternative Forced Choice (2AFC) paradigm.

The data recorded in all four experimental sessions was both quantitative and qualitative such as the finger locations, user gaze and self-assessment questionnaires. This allowed us to examine both the user performance in terms of accuracy and precision but also their user experience, in terms of immersion and fatigue symptoms.

1.3 Outline

The next chapter introduces the state of the art technologies and software frameworks currently available for VR devices along with some relevant interaction and perception modalities. How humans perceive the virtual world and some related issues will also be discussed.

Next, the research work carried out during the thesis will be described (Chapters 3 and 4). More specifically, Chapter 3 focuses on the work done on mitigating cybersickness and Chapter 4 concentrates on understanding how to improve depth perception through visual perception. Section 3.1 introduces the spatial blurring system developed for VR systems. In Section 3.2, the experimental study conducted to verify the effectiveness of the developed foveated DoF effects is explained. Section 4.1 details the pilot study conducted on depth perception. Section 4.2 introduces the inverse blurring system along with how its various parameters were tuned. Sections 4.3 and 4.4 highlights the two experimental studies conducted to verify the usefulness of the inverse blurring system to improve depth perception.

Finally, Chapter 5 will discuss the research findings with respect to the original research questions and Chapter 6 will describe some open issues and possible future developments.

PART II

State of the Art

The following part aims to introduce all the state of the art required to understand the research work described in the thesis. Since various fields are involved such as visual perception in immersive virtual environments and user experience assessment, a multidisciplinary and heterogeneous approach is used.

2

Background

2.1 VR Devices

Virtual Reality (VR) is a computer generated environment with virtual objects and scenes shown to the users making them feel as if they are immersed into their new surroundings. This environment is perceived through a device referred to as a VR headset or helmet. The most common type of device is a Head-Mounted Display (HMD). Other devices include room filling technologies.

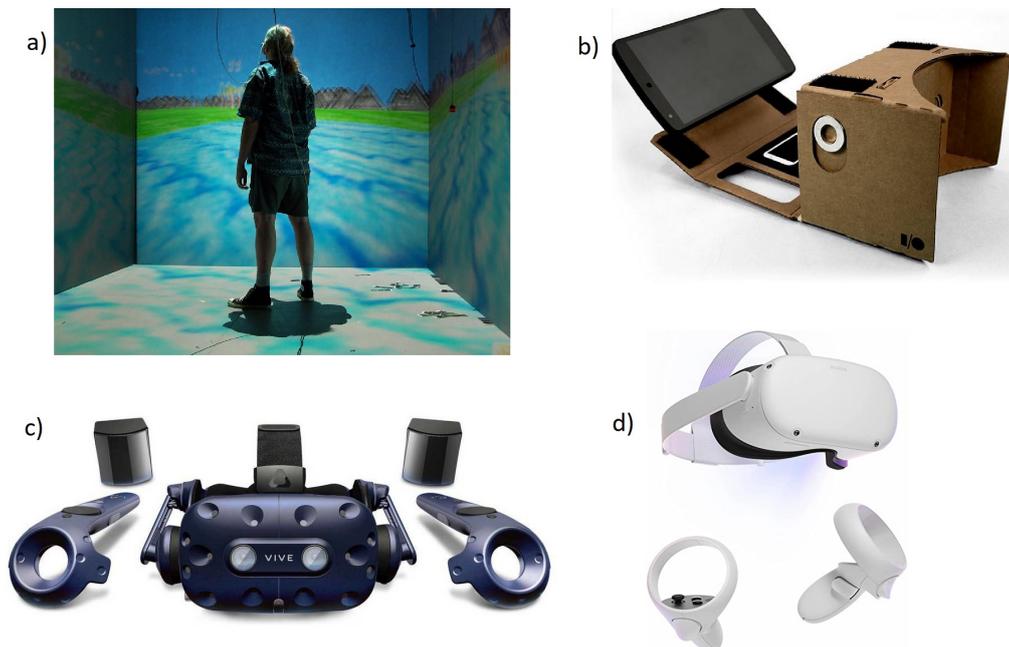


Figure 1: Examples of the VR devices. a) Room-filling technology (CAVE). b) Mobile headset (Google Cardboard). c) Tethered device (HTC Vive Pro). d) Standalone setup (Oculus Quest).

The current age of Virtual Reality (VR) began in 2010 when the first prototype of a VR headset was introduced by Palmer Luckey through a kick starter

campaign. This device would later evolve into the Oculus Rift¹ which was later bought by Meta (known as Facebook at that time). As time progressed, more competitors emerged such as the HTC Vive² and Sony Playstation VR³. Soon, smartphone based VR devices entered the market such as the Samsung Gear VR⁴ and Google Cardboard⁵.

VR headsets normally come with a tethered setup in which the HMDs are wired to a high-end computer with a dedicated graphics card. These devices require high processing power so typically a minimum of 4GB RAM is required along with a powerful processor. This tethered setup is due to the fact that performing on-board computations of such high processing power requires specialized equipment which can be heavy and potentially ruin the immersive experience.

Recently, a new generation of devices such as Oculus Quest⁶ and HTC Vive Focus⁷ have been introduced which are portable as they do not require to be tethered. These devices have embedded processor and sensors instead. These devices can act stand-alone although they are unable to provide a high resolution and fast frame rates like the tethered versions.

Mobile headsets require the insertion of smartphones into special goggles. The smartphone acts as the processor and display. These devices are able to provide a sufficient immersive experience for a limited time. However, they only offer a lower resolution and a smaller FoV. These limitations have pushed the mobile headsets to extinction with major companies having stopped development. Some common commercially available VR devices are listed in Table 1.

Room filling technologies such as the CAVE⁸, YURT⁹ and AlloSphere¹⁰ are systems that project images through a series of high resolution displays covering a 360° surface. These setups usually require the users to wear specialized stereoscopic or 3D glasses giving the perception of the objects floating in the air. The user's movements are tracked so that the projected images can be adjusted accordingly to adapt to individual perspectives. These systems are generally very expensive and require a dedicated installation so are mainly used in universities for research, in industries for large scale data visualization and in some cases also in multimedia productions and archaeological sites for reconstruction.

An important element of VR devices is the ability to interact with the virtual environment. To simulate user movements in a virtual environment, it is necessary to track and recognize the user's hand movements [69]. Also, in order to allow the user to have a natural interaction, haptic feedback needs to

1 <https://www.oculus.com/rift/>
 2 <https://www.vive.com/us/product/vive-pro-full-kit/>
 3 <https://www.playstation.com/en-us/ps-vr/>
 4 <https://www.samsung.com/global/galaxy/gear-vr/>
 5 <https://arvr.google.com/cardboard/>
 6 <https://www.oculus.com/quest-2/>
 7 <https://www.vive.com/us/product/vive-focus3/overview/>
 8 <http://www.visbox.com/products/cave/>
 9 <https://www.brown.edu/academics/early-cultures/resources-brown/yurt>
 10 <http://www.allosphere.ucsb.edu/>

| DEVICE | TYPE | RESOLUTION | FOV | REFRESH RATE |
|------------------|------------|------------|------|--------------|
| Oculus Rift | Tethered | 1080x1200 | °94 | 90 Hz |
| Oculus Rift S | Tethered | 1280x1440 | 90° | 80 Hz |
| Oculus Quest | Standalone | 1440x1600 | 94° | 72 Hz |
| HTC Vive | Tethered | 1080x1200 | 110° | 90 Hz |
| HTC Vive Pro | Tethered | 1440x1600 | 110° | 90 Hz |
| HTC Vive Focus | Standalone | 1440x1600 | 110° | 75 Hz |
| HTC Vive Cosmos | Tethered | 1440x1700 | 110° | 90 Hz |
| PlayStation VR | Tethered | 960x1080 | 100° | 120 Hz |
| FOVE | Tethered | 1280x1440 | 110° | 70 Hz |
| Samsung Gear VR | Mobile | 1280x1440 | 96° | 60 Hz |
| Google Cardboard | Mobile | 1280x1440 | 80° | 60 Hz |
| Oculus Go | Standalone | 1280x1440 | 100° | 72 Hz |
| Razer OSVR | Tethered | 1080x1200 | 110° | 90 Hz |
| HP Reverb | Tethered | 2160x2160 | 114° | 90 Hz |
| Pico Neo | Standalone | 1280x1080 | 102° | 90 Hz |

Table 1: Common commercially available VR devices.

be provided. These interaction devices are generally of two types, namely the wearable sensor based devices and computer vision based devices.

Most VR headsets come with a pair of controllers, one for each hand. These are ad-hoc joystick style wireless devices. They provide an intuitive interaction through buttons, triggers and track-pads and feedback through vibrations. They are also integrated with inertial sensors and external tracking systems to allow position and orientation of objects and users in the real world to be imitated in the virtual world. An alternate form of wearable sensors are the data gloves such as the Manus¹¹ and Gloveone¹². These devices offer a higher precision and a more natural embodiment to the user.

Computer vision based devices such as the Microsoft Kinect¹³ and Leap Motion¹⁴ use largely trained vision based algorithms to track the user's movements. These devices either use an Red-Green-Blue-Depth (RGB-D) sensor or a stereo camera to detect objects in the real world without having any markers installed on them. However, tracking accuracy and precision is slightly lower for these systems as opposed to their wearable counterparts.

¹¹ <https://www.manus-meta.com/haptic-gloves>

¹² <https://www.kickstarter.com/projects/gloveone/gloveone-feel-virtual-reality>

¹³ <https://developer.microsoft.com/en-us/windows/kinect/>

¹⁴ <https://www.ultraLeap.com/product/leap-motion-controller/>

2.1.1 *Gaze Tracking in VR*

A recent trend in the field of XR is the introduction of eye tracking technology into HMDs. Eye tracking is a fairly old field with sufficient developments and applications in psychological experimentation. The exact means to how eye movements are measured have evolved over the course of history. Initially specialized contact lenses with pointers were used. However, today the technology has moved towards video based systems using computer vision algorithms. With the introduction of lightweight compact cameras, it is now possible to incorporate such technology inside a headset. Common examples of VR HMDs using eye-tracking are HTC Vive Pro Eye¹⁵ and FOVE¹⁶.

Eye-tracking in VR has opened up new possibilities to how research in human cognition is carried out. The subject is immersed into a relatively natural environment that reacts to movements and actions, while all experimental conditions can be precisely controlled. The combination of VR and eye tracking makes it possible to calculate the user's gaze direction in 3D space and measure where the user was looking throughout the session [20]. Defining regions of interest is comparatively easier in VR as compared to traditional eye tracking as the gaze points can be traced over time.

To determine where the user was looking in VR, it is necessary to obtain the 3D gaze vector going from the user's eyes to the point where it is looking. Eye trackers have a built-in eye model which is used to determine the pupil characteristics. Subsequently, this models allows eye-trackers to compute the 3D gaze. Depth can be computed using the divergence of the two eyes by calculating the crossing point of the gazes from each eye. A limitation of such approach is that, the measurements are only precise when the calibration is perfect [52].

Usually when using eye tracking, a calibration process needs to be carried out. This is typically done by showing the user an image of target points and asking the user to fixate at them. A typical example is shown in Figure 2. It is important to show these points in screen space rather than world space as this makes them move together with the head allowing the whole Field-of-View (FoV) to be covered.

During calibration, the eye-tracking uses these points as reference to adjust its computation of the gaze vector. The accuracy of such procedure deteriorates over time due to drifts as the headset slightly slips with head movements [20]. There are some techniques available that continuously re-calibrate the eye-tracking algorithm [122], however some additional information such as the user's viewing behaviour are required making it harder to generalize the procedure. Therefore, it is recommended to re-calibrate the eye-tracking setup every few minutes.

¹⁵ <https://www.vive.com/us/product/vive-pro-eye/overview/>

¹⁶ <https://fove-inc.com/product/fove0/>

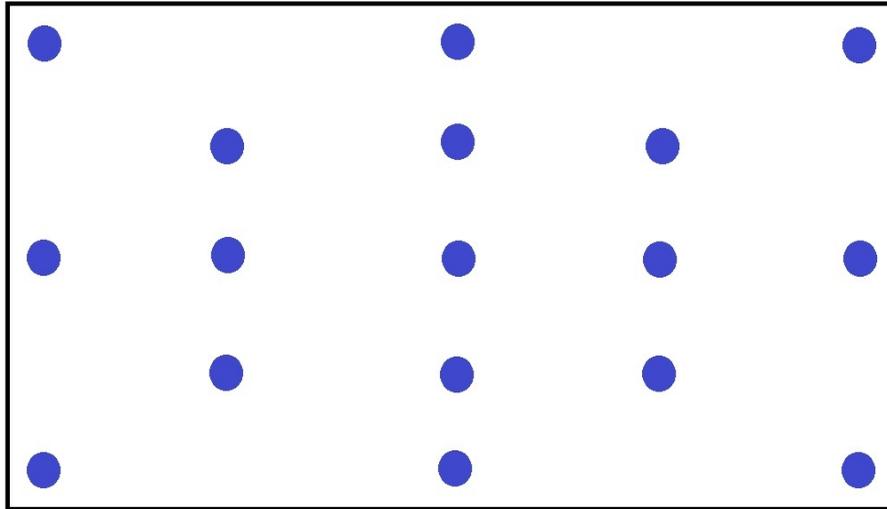


Figure 2: Typical example of the eye calibration step. The user is asked to fixate at the blue dots one by one.

2.2 Visual Perception in Virtual Environments

The idea of what a perfect VR system may look like was first introduced by Ivan Sutherland many decades ago [117]. He suggested that in a perfect VR system, the users will not be able to differentiate whether they are interacting with a real object or a virtual one. This system should be able to stimulate all human senses. Modern consumer technology offers a very realistic representation, however, some perceptual issues still remain that subsequently lower the sense of immersion in HMDs [34].

Humans use a variety of cues to determine the size and distance of objects in their environment. Typical cues include disparity, motion parallax, occlusion, convergence and accommodation [99, 102]. However, not all cues are used at all times. Which cues are being utilized are more or less determined by the distance to the objects. Human spatial reach can be divided into three circular egocentric regions, namely the personal space, action space and vista space [22]. Objects within 2m are considered to be in personal space. Disparity, accommodation and convergence are more prevalent in this case. From 2 to 30m, it is referred to as the action space and occlusion and motion parallax are more dominant. Distances beyond 30m are considered to be vista space. Only pictorial depth cues such as occlusion and relative size are used [23].

Over the years, several studies [5, 56, 59, 70, 84] have been conducted on VR which suggest that users typically underestimate the distances to objects by around 25% [102]. This is significantly higher when compared to human performance in the real world where even when blind walking only 8% underestimation occurs [135]. There are many potential reasons for this difference. The weight of the HMDs combined with limited FoV are one potential reason. In close surroundings, disparity also plays a role in this difference. Some studies

have suggested that the Interpupillary Distance (IPD) setting also plays a crucial role [11, 101]. The larger the IPD the higher the amount of underestimation.

2.2.1 Vergence Accommodation Conflict

When it comes to visual perception in VR systems, users tend to experience conflicting cues. Such contradictions not only give rise to many errors in object size and distance estimations but also affects immersion and makes the user feel uncomfortable over long exposures. The most prevalent of such mismatches in modern AR/VR devices is the Vergence Accommodation Conflict (VAC).

When humans view objects in the real world, the eyes converge inwards while the ciliary muscles deform the lens. The former is referred as convergence while the latter is called accommodation. This process occurs so that a sharp image is formed on the fovea. However, this is not the case in XR setups. The image is shown at a fixed distance while the depth of the virtual object varies with the content according to the disparity [62, 91, 95]. The basic geometry of this is shown in Figure 3.

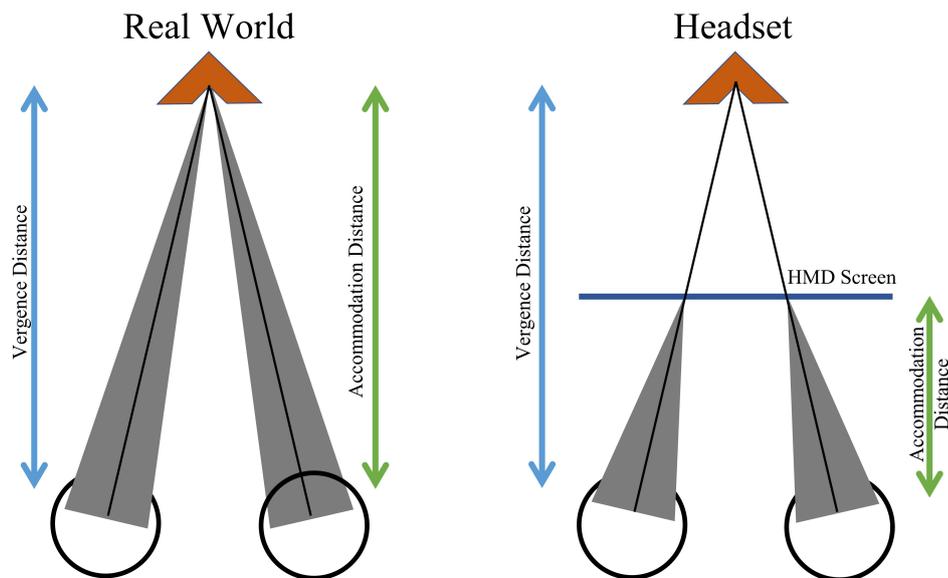


Figure 3: Vergence accommodation conflict. In natural viewing, eyes focus and converge at the same distance. Whereas, in stereo 3D viewing, the eyes focus and converge at different distances.

Convergence and accommodation are two important cues prevalent in personal space. Since most immersive media devices display stimuli to the user through a screen placed very close to the user eyes, it is important that these two cues work cohesively.

In literature, most researchers have proposed hardware solutions such as adjustable or focus tunable lenses [9, 93]. These systems are able to alter the focal length of the lens depending on where the user is looking. This way they

no longer have fixed accommodation and can potentially be used to correct hyperopia and myopia in VR systems. The major drawback of such setups is that they are hardware intensive and cannot be adapted to modern lightweight HMDs.

2.3 Cybersickness

One of the most significant hurdles in the wider uptake of VR technology is the onset of cybersickness. Users have long reported nausea like symptoms causing discomfort after being exposed to VR for long duration of time. This discomfort is often referred to as Simulator Sickness (SS) or visual fatigue or eye strain. Perceived discrepancies between how humans perceive and move in the virtual world as compared to the virtual one is the mostly accepted underlying cause.

Existence of cybersickness is not a new phenomena. It has been inherent since the start of VR. Cybersickness can lead to a wide range of symptoms such as nausea, disorientation, headaches, sweating and eye fatigue [24]. How significantly these symptoms occur vary largely on the user and on the application. Various factors contributing to cybersickness in VR are summarized in Table 2. Cybersickness can occur within a few minutes of exposure even in some trivial VR applications [21]. Users have also reported symptoms to a varying degree under the same experimental condition [53].

| INDIVIDUAL FACTORS | DEVICE FACTORS | TASK FACTORS |
|--------------------|----------------|--------------|
| Age | Lag | Control |
| Gender | Flicker | Duration |
| Illness | Calibration | |
| Posture | Ergonomics | |

Table 2: Factors effecting cybersickness in virtual reality. Adapted from [24].

Cybersickness is thought to be closely related to Motion Sickness (MS). MS is the unpleasant feeling, accompanied by nausea and vomiting that occurs when a person is travelling in a moving vehicle. Previous studies have found that younger children between the ages of 4 and 12 are more susceptible to MS. This susceptibility of experiencing MS in childhood has proven to be a good indicator of experiencing cybersickness later [35]. The susceptibility to cybersickness decreases as people get older [61]. However, some interesting research has found that experienced people are also more prone to the onset of SS [53].

Apart from age, another important factor is gender. Women have a wider Field-of-View (FoV) making them more prone to flicker perception which increases the likelihood of suffering from cybersickness [64]. Female hormones are also one potential candidate to why women suffer more from cybersick-

ness [61]. Physical factors such as fatigue, flu and hangover are also closely linked with cybersickness.

Apart from human factors, the VR device itself can contribute towards cybersickness. Device manufacturers normally consider a wide range of factors such as lag, flicker, calibration accuracy and general ergonomics when designing devices to make sure the users experience the least amount of symptoms. User movements need to be tracked immediately and the view needs to be updated as soon as possible. Typically, this tracking occurs at 50–60 Hz. Tracking inaccuracies resulting from poor calibration or faulty sensors can lead to cybersickness. Likewise, the IPD which is the distance between the two eyes needs to be set appropriately as it varies significantly from user to user. Subsequently, the offset in the stereoscopic view needs to be adjusted.

2.3.1 *Causes*

Although, cybersickness has been identified a long time ago, there are still some contradicting views to why it occurs and possible strategies to tackle it. In literature, there are three popular theories to explain why cybersickness occurs, namely the poison theory, the postural instability theory and the sensory conflict theory.

Poison berry theory or simply poison theory suggests that an evolutionary mechanism is triggered when the experienced sensory input is different from what is expected [121]. The resulting symptoms such as dizziness and vomiting are often associated with poisoning, hence the name. Although this explanation explains some of the symptoms, however, it fails to justify some of the broader range of cybersickness symptoms.

Postural instability theory was first introduced by Riccio and Stoffregen as the source of all MS [103]. The theory suggests that the main goal of humans is to maintain postural stability in the environment and sickness occurs when mechanisms for maintaining posture are compromised. Prolonged exposure to this instability results in sickness symptoms and the extent of the symptoms increases as the duration of the instability increases. In Virtual Environment (VE), abrupt or unnatural visual changes that have not been previously learned by the user lead to conflict in the postural control strategies resulting in the cybersickness symptoms.

While the previous two theories have their own merits, most researchers believe the sensory conflict to be the most dominant reason for the onset of cybersickness [57, 97]. This theory is based on the mismatch in cues posed by the human visual system and the human vestibular system. These systems provide important information about the person's orientation and perceived motion. The theory when applied to VR suggests that although in the virtual world, the user is moving, the person in reality is not. The visual system based on various cues such as optic flow suggests to the brain the person is in motion, however, in reality the person is stationary as dictated by the vestibular system. The resulting conflict causes the user to experience cybersickness.

2.3.2 *Assessing Cybersickness*

There are several methods to measure cybersickness, however, self-reported questionnaires remain the most popular approach [27, 97]. The first form of such questionnaire dates back to the 1960s when the Pensacola Motion Sickness Questionnaire was introduced. The questionnaire was based on 27 previously identified issues. Over the course of history many alterations have been proposed.

In 1990s, Robert Kennedy along with his colleagues performed an extensive analysis on simulator sickness and came up with the Simulator Sickness Questionnaire (SSQ) [54]. This questionnaire reduced the previously available questionnaires to 16 items. The questionnaire asks the user to rate each of the 16 items on a 4-point Likert scale. The SSQ introduced a system of multivariate measures related to oculomotor effects, nausea and disorientation. The symptoms associated with each of these three clusters is shown in Table 3. The three clusters are not orthogonal to each other.

| NAUSEA | OCULOMOTOR | DISORIENTATION |
|----------------------|---------------------|----------------|
| Stomach awareness | Eyestrain | Dizziness |
| Increased salivation | Difficulty focusing | Vertigo |
| Burping | Blurred vision | |
| | Headache | |

Table 3: Simulator sickness questionnaire clusters. Adapted from [24].

The SSQ has become a sort of a standard tool in research to measure cybersickness and remains the most cited tool. With the lift-off of the XR devices, there was more emphasis on coming up with standardized questionnaires that specifically target VR. For this reason, the Virtual Reality Symptom Questionnaire (VRSQ) was developed. It draws from experience from previous questionnaires and suggests a 13 item questionnaire divided into ocular and non-ocular categories. However, this questionnaire lacks validation and has not been widely adopted in the VR field.

Recently, there has been interest among researchers to move away from subjective measure to objective ones. Emphasis has been put on using physiological signals such as heart rate, respiratory rate, Electroencephalography (EEG), skin conductance and blink rate instead [10, 33].

2.3.3 *Solutions*

Improving the hardware capabilities of VR devices such as increasing the maximum supported frame rate and higher FoV may well ultimately eliminate cybersickness [1, 97]. However, in the absence of such features researchers have proposed many techniques over the years to reduce the level of induced cybersickness.

Dynamically altering the FoV is one of the proposed solution [30]. The approach tends to subtly alter the FoV of a stationary person in response to visually perceived motion as they transverse the virtual world (see Figure 4). Although reducing the FoV has shown to reduce cybersickness, it is at the expense of reduced sense of presence. Other approaches proposed in literature include incorporating spatial or defocus blurring [3, 44]. Saliency-based dynamic blurring is a method in which the virtual scene is blurred based on user movement except for the salient areas in the scene such as road signs (see Figure 5 for an example). However, it only worked for high speed scenes [86].

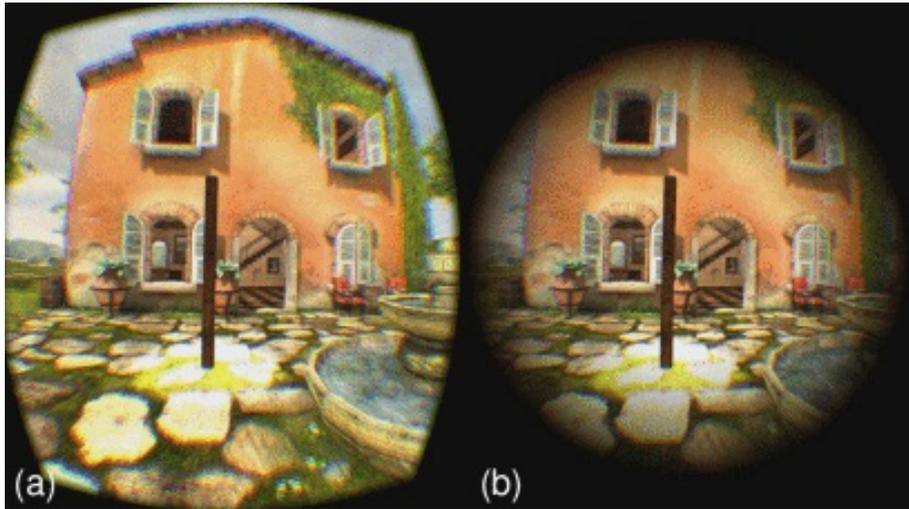


Figure 4: Dynamically altered field-of-view. a) Original scene. b) Reduced FoV scene. Image taken from [30]



Figure 5: Saliency-based blurring technique proposed by [86]. Blur is applied based on color in a forest scene. The text on the yellow colored barrel remains readable while everything else is blurred. Image taken from [3].

There are some methods to mitigate the effects ofvection which is also highlighted as a cause of cybersickness [12]. Vection is the perception of self-motion

in the absence of any physical movement, often caused by secondary moving objects in the user-view. Rotational blurring can be used for this which is essentially applying a Gaussian filter to the entire scene when peripheral objects undergo rotational movements.

Some researches tried to reduce the onset of cybersickness by addressing the optic flow in the peripheral regions. The VR scene can be divided into circular sections and the application design can be done in such a way that the object motion in the outer or peripheral regions can be minimized [13].

Use of vignetting during amplified head movements to counter cybersickness had an opposite effect [87]. Moreover, a recent study demonstrated that introducing spatial blur effects in VR systems can also help with depth perception [47].

2.4 Spatial Blurring

Blurring or smoothing an image is not a new phenomena. It is essentially the process of applying a Low Pass Filter (LPF) to the image usually by convolving the filter kernel with the image. The choice of the LPF is usually application dependent. Typical filters include moving average filter, Gaussian filter and disc filter. The main purpose of smoothing an image is to reduce noise and highlight patterns in the image. For example, digital cameras such as the ones found on most smartphones use Gaussian smoothing to reduce noise associated with International Organization for Standardization (ISO) light sensitivities. The effects of such filtering are irreversible and results in some loss of detail.

Unlike traditional blurring approaches, spatial blurring is inspired from nature. It assumes that the intensity of blurring can be space-variant, similar to how it is in the human eyes. Humans use photo receptive cells on the retina to visually perceive their environment. The spatial density of such cells varies with the foveal region having the highest density and it decrease as it moves towards the periphery [105] (see Figure 6). To resolve detail in the person's surroundings, the lens in the eye needs to project a focused image on the fovea. The projections hitting the periphery are lower in spatial resolution. This results in an image that is space-variant. Foveal vision is used for extracting detail while peripheral vision is aimed at recognizing structures and movements.

With context to the computer graphics field, spatial blurring implementations can be divided into two main categories, namely the object space and the image space methods [6]. Object space methods operate directly on the 3D scene and are built into the rendering pipeline. In contrary, image space methods are considered a post-processing operation since they operate on images and their corresponding depth maps. Images and depths are obtained from the output of the normal rendering pipeline. Each pixel is blurred using information from the camera model and depth map.

Object space methods tend to have more accurate results and suffer less from artifacts as compared to image space methods. However, image space methods

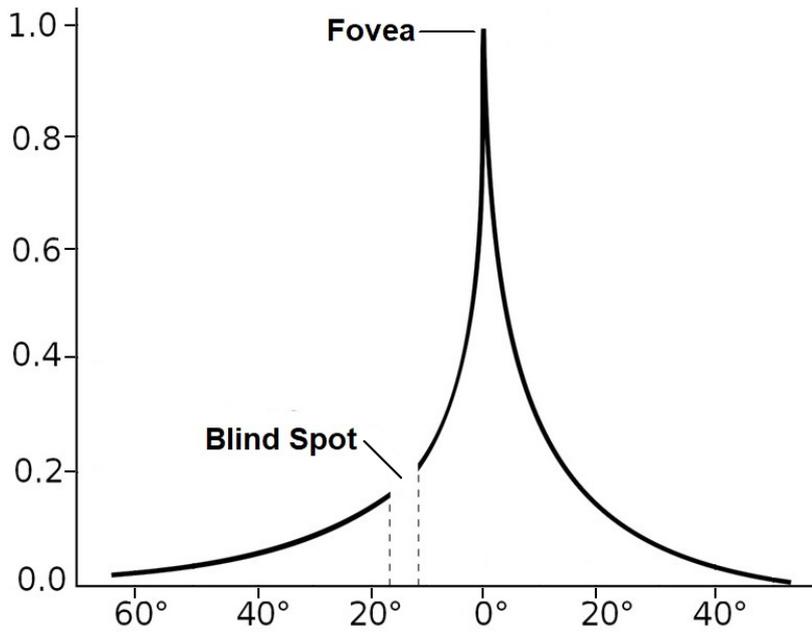


Figure 6: Visual acuity spread.

are much faster. Speed is of critical importance in virtual reality applications so image space methods are usually preferred. Image space methods need to be tuned carefully in order to avoid artifacts. Most commonly encountered artifacts include intensity leakage and depth discontinuity. Intensity leakage is when a blurred background blurs on top of an in-focus object. Depth discontinuity is when the background is in-focus but the silhouette of the foreground object appears sharp. These artifacts mainly occur when there is an abrupt change in the depth map.

There are two popular techniques as far as spatial blurring in XR is concerned. These are Depth-of-Field (DoF) effects and foveated rendering and are discussed in more detail below.

2.4.1 *Depth-of-field*

In the computer graphics field, Depth-of-Field (DoF) rendering is a popular approach to incorporate spatial blur. Images are blurred using information from the camera model and the corresponding depth maps. An example of this effect is shown in Figure 7

Several attempts have been made to introduce DoF blur effects in VR systems [16, 43]. These systems assume a focus distance and use the lens model to compute the Circle of Confusion (CoC) which is an optical spot caused by light rays directed onto the camera's focal plane by the lens. The amount of blur in the peripheral pixels is based on the depth difference between the point of fixation and that particular pixel. However, these systems are not gaze-contingent as they assume either a fixed focus distance or assume the user is always fixated at the center of the scene.



Figure 7: Illustration of depth-of-field effect. The images on the right are high acuity images while the ones on the left have the depth-of-field effect applied to them. Image taken from [16].

Alternatively, gaze contingent systems have also been proposed for near-eye displays [9, 93]. These systems use adjustable lenses and can potentially be used to correct hyperopia and myopia in VR systems. The major drawback of such setups is that they are hardware intensive and cannot be adapted to modern lightweight HMDs.

2.4.2 Foveated Rendering

Foveated imaging is a technique in which the image resolution varies across the image according to the fixation point (see Figure 8). This technique aims to simulate a drop in acuity in the visual system from fovea to periphery, as experienced by humans, by rendering peripheral content to a smaller frame buffer resolution and then resampling it using a range of temporal and spatial upscaling algorithms [7].

Recent developments in the field of foveated rendering [79, 94, 119] has helped reduce the computational load for VR devices. Such systems are able to reduce the required number of processed pixels up to 20x and can offer approximately 3x faster rendering times. A popular approach to implementing space-variant blurring is the log-polar mapping [19, 115, 120]. The image is first transformed into the cortical domain and then into the retinal domain. This results in an output that has higher resolution in the center and lower resolution away from the center of the image. The concept is illustrated in Figure 9. Such techniques were exploited by Meng et al. who proposed a kernel



Figure 8: Illustration of multi-region foveation. The scene is divided into various sections and a different amount of down-sampling is applied to each region. Image taken from [7].

based foveated rendering approach that maps well to current generation of GPUs [79].

Alternatively, a phase-aligned approach towards foveated rendering has also been developed [124]. Only the high acuity foveal region is aligned with the head movements while the peripheral region is instead aligned with the virtual world. Thus, only the high acuity regions require additional processing in each frame.

Current foveated rendering methods use fixed parameters that are often tuned manually. A recent work proposed to use a content aware prediction model based on luminance and contrast to compute the optimal parameters [125].

A common issue in most foveated rendering techniques is geometric aliasing which appears in the form of temporal flickering and can be easily noticed by users [37]. Some solutions have recently been proposed to overcome these artifacts such as temporal foveation built into the rasterization pipeline [32]. This is achieved by introducing a confidence function based on which it is decided whether to re-project the pixels from the previous frame or to redraw them. Such algorithms work relatively well on dynamic objects, which is a bottleneck for most foveated rendering algorithms. However, since this approach does not always use a freshly rendered image as input and rely on data from previous frame to achieve a high computational performance, it does not work well with reflections and transparent objects. Alternatively, DoF has been proposed

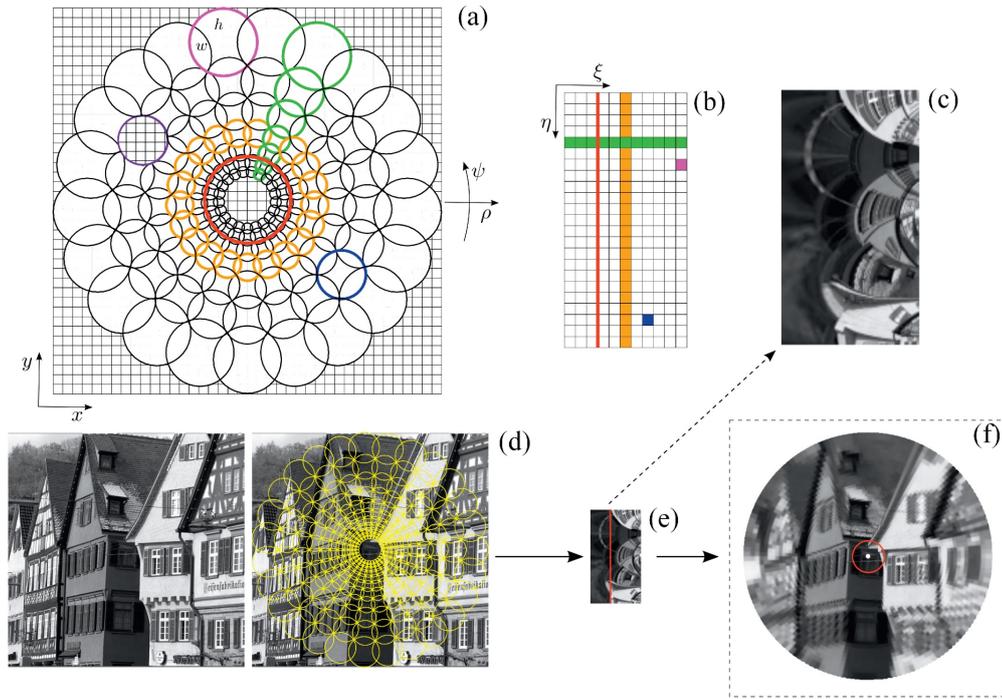


Figure 9: Illustration of the log-polar mapping. a) Cartesian domain with overlying log-polar pixels represented by circles. The area inside the red circular curve represents the fovea b) Cortical domain where orange and green strips denote the corresponding log-polar pixels from (a). c) Enlarged image in cortical domain. d) Image in Cartesian domain. e) Transformed image in cortical domain. f) Transformed image in retinal domain. Image taken from [19].

as a post-step to remove artifacts introduced by foveated rendering algorithms [131, 132].

Some researchers have investigated how the size of the foveal region or the central window influences cybersickness [71]. No correlation was found between the amount of induced sickness and the size of the central window. However, it was observed that users adapt more quickly to larger foveal regions. Although techniques based on foveated imaging can reduce visual fatigue, but they provide focus information uncoupled from depth information. By using a combination of the multi-region foveation and DoF, a more natural scene can be produced [40]. Moreover, some recent works [47, 49, 75] showed that foveation can affect human depth perception.

2.4.3 Assessing Image Quality

When filters are applied to an image, it is important to have a quantitative measure to assess how good the output is. Image quality refers to the weighted combination of all the visually significant attributes of an image [14]. It is an indicator of how aesthetically pleasant the image is to the viewer. Subjective

measures are quite expensive as they require a huge number of participants and are difficult to automate in real-time.

Image Quality Assessment (IQA) methods can be divided into three main categories:

- **Full-reference:** These metrics try to assess the quality with reference to a reference image that is assumed to have perfect quality.
- **Reduced-reference:** These metrics try to assess the quality based on features extracted from a given image and a source image.
- **No-reference:** These metrics assess image quality without any reference to an original image.

Peak Signal-to-Noise Ratio (PSNR) is the ratio between the power of a signal to the noise corrupting the signal. When applied to an image, it first computes the Mean Squared Error (MSE) for each channel. Given a monochrome source image, I and a noisy image, K of size $m \times n$, MSE and PSNR can be computed by:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2 \quad (1)$$

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right) \quad (2)$$

It is commonly used to assess the reconstruction quality of an image. A higher PSNR usually indicates a better quality image. Although PSNR is a simple tool, it often performs poorly when estimating how humans will perceive the image as compared to other methods.

Structural Similarity Index Measure (SSIM) is a perceptual model that is used to measure the similarity between two images. It perceives the noise or degradation in image as changes in the structural information contained in the image [130]. SSIM is mainly applied to videos but is also widely applied in the still photography industry. SSIM extracts three key features of the image, i.e., luminance, contrast and structure and bases its computation around these properties. Instead of applying the measure globally on an image, it is a common practice to apply it locally to regions on the image and then to compute the average. This new measure is called the mean-SSIM and is a more effective method.

The previously discussed metrics require a reference image which may not be always available. An alternate metric is the Visual Information Fidelity (VIF) which is based on the natural scene statistics [109]. The reference image is modeled as being the output of a stochastic natural source that passes through the Hue-Saturation-Value (HSV) channel and is processed later by the brain. The same measure is also computed in the distortion channel. These two are then combined to compute the VIF index.

Recently a new approach has been proposed that has been designed on the space-variant resolution found in human visual system. FovVideoVDP is

a visual difference metric that models the temporal aspects of vision and accounts for foveated viewing [77]. The measure was developed due to a need for image quality metric specifically targeting the AR/VR rendering technologies. The advantage of such metric is its ability to generalize across a diverse range of contents and types of spatio-temporal artifacts.

PART III

Research Work

The following part aims to introduce the various research work done as part of the thesis along with the experimental process that was used to validate the design concept. Texts and figures in this part have previously appeared in the candidate's authored published articles [47–49] and an under-review patent [50].

3

Addressing Cybersickness in Virtual Reality Systems

Cybersickness has been a major stumbling block in the widespread usage of VR devices. Various techniques have been proposed in literature to address this issue, however, most of them are at the expense of presence. In this chapter, a spatial blurring technique is presented. The technique can be incorporated into any VR application. Although the technique is designed focusing on VR devices, it can be adopted easily to other XR platforms. The effectiveness of the technique in mitigating cybersickness is validated through a user study.

3.1 Foveated Depth-of-field

The developed spatial blur technique incorporates DoF blur and foveation effects. For any algorithm to work seamlessly for VR devices, it is essential for it to have real-time capabilities. To ensure this, the technique is implemented at the shader level. Since image space methods are more desired for VR applications due to their superior processing times, they are exploited in the linear color space. To apply blurring to an image, a smoothing filter has to be applied. Different types of smoothing filters were considered, such as Gaussian filtering, Bokeh [80], and disc effects. However, since the system takes inspiration from the human physiological system, the Bokeh filter was preferred as it better mimics the aperture present in the human eyes and can lead to a more realistic output.

The implementation is done using a custom four-pass shader. Algorithm 1 describes the pseudocode of the foveated DoF effects, while the process flow of the developed technique is shown in Figure 10. In the first shader pass, the Circle of Confusion (CoC) diameters are computed using the raw depth values and stored in a single-channel texture object. The CoC diameters are shown as grey for objects farther from the fixation plane and as purple for objects in between the user and the fixation plane. Simultaneously, the image is divided into three circular sections by computing the distance of each pixel to the fixation pixel. Red pixels represent the pixels in the foveal area, while green

and blue pixels represent the near and mid peripheral regions, respectively. Using the source image and the CoC texture, DoF effects are computed in the second shader pass. Similarly, using the foveation mask and the source image, the foveation effects are computed in the third shader pass. In the last shader pass, the effects are combined to obtain the final output. The smoothing filters are applied at half resolution of the source image and the resultant frames are later up-sampled. Details of the individual processes involved are described in the following subsections.

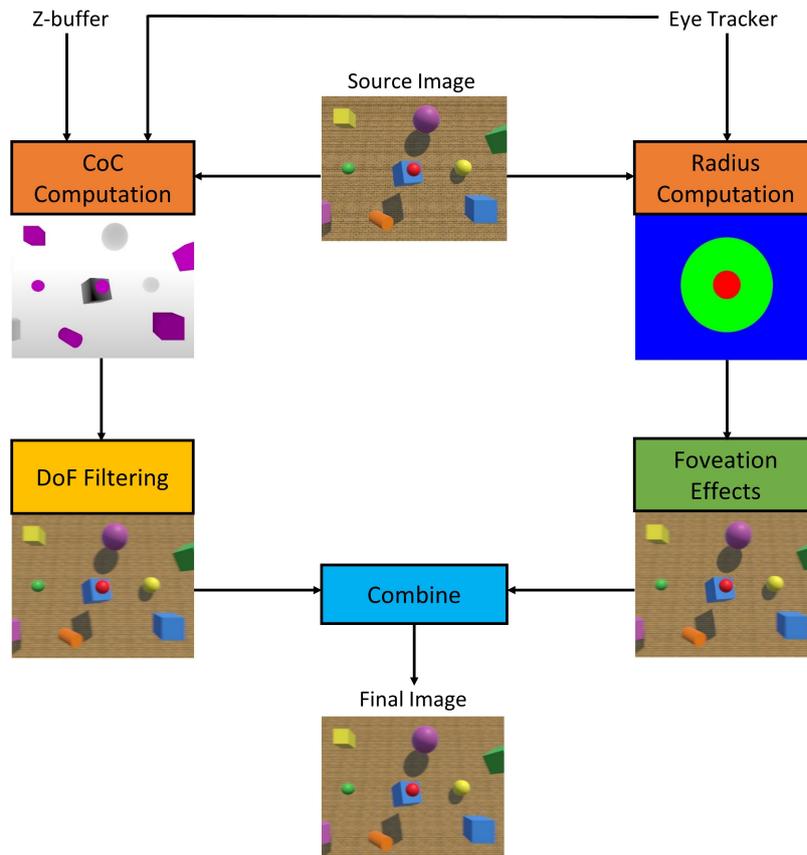


Figure 10: Process flow of the proposed foveated depth-of-field technique showing the intermediate outputs. Fixation is at the center of the red sphere.

3.1.1 *Depth-of-field*

When humans visually perceive their surroundings, the retinal images contain a space-variant resolution due to which the peripheral content appears blurred. This variation is due to the objects being placed at different depth planes and is an important cue for depth perception. In order to synthesize this blur effect in VR systems, depth texture object is used to create the depth map of the virtual scene. Depth values corresponding to each pixel on the HMD screen are computed and stored in a Z-buffer. The information inside the Z-buffer is scaled between 0.0 and 1.0 to ensure the system can be used with any

Algorithm 1: Foveated DoF effects for VR

Input: Z-buffer (B_z), eye tracker data (E_t), source image (T_S)
Output: Foveated DoF image (T_{FD})
 $T_d = \text{computeDepthMap}(B_z);$
// Shader Pass 1
for each pixel do
 $T_{CoC} = \text{computeCoC}(E_t, T_d);$
 $T_r = \text{computeRadius}(E_t, T_S);$
end
downSample $T_S;$
// Shader Pass 2
for each pixel do
 $\sigma_d = \text{computeBlurParameter}(T_{CoC});$
 $T_{DoF} = \text{applyDoF}(T_S, \sigma_d);$
end
// Shader Pass 3
for each pixel do
 $\sigma_f = \text{computeBlurParameter}(T_r);$
 $T_{Fov} = \text{applyFoveation}(T_S, \sigma_f);$
end
// Shader Pass 4
for each pixel do
 $B = \text{computeBlending};$
 $T_{FD} = \text{combine}(T_{DoF}, T_{Fov}, B);$
end
upSample T_{FD}

HMD configuration. This depth information is used to define the parameters of the smoothing filter. An eye tracker is used to identify the fixation plane and the amount of blur is varied based on the difference in pixel depths, i.e., on the difference in depth of the scene objects with respect to the fixation plane. Objects on the accommodative plane are kept as they are in the source image while a smoothing filter is applied on every other region.

The Circle of Confusion (CoC) concept from the field of optics is used to model the amount of blur associated with each pixel. An illustration of the concept can be seen in Figure 11. When the lens is focused at the object placed at distance Z_f , the light rays from the object placed at distance Z_p projects a circle with diameter C on the retina. This circle is referred to as CoC. The formulation developed by Held et al. [41] is used for computing C and this is defined by:

$$C = As \left| \frac{1}{Z_f} - \frac{1}{Z_p} \right| \quad (3)$$

where s is the distance between the retina and lens, more commonly known as the posterior nodal distance, and A is the aperture of the eye.

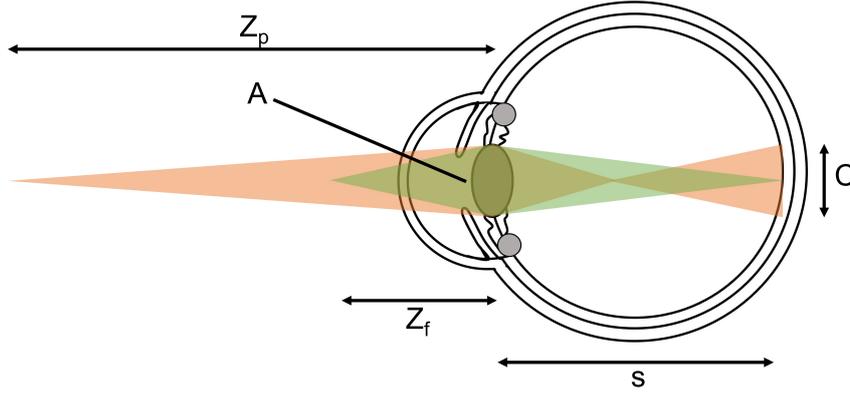


Figure 11: Illustration of the circle of confusion concept. Point of fixation is at distance Z_f . Point located at distance Z_p forms a circle on the retina with diameter C . A denotes the aperture and s is the posterior nodal distance.

CoC is used to alter the blur associated with each pixel. The bigger the size of C , the higher the amount of blur that is present. This implies that the parameter of the blur σ_d has a direct relation to the size of the circle of confusion, i.e., $\sigma_d \propto C$. Equation 3 is adapted to the developed system, so, σ_d is defined by:

$$\sigma_d = K \left| \frac{1}{D_f} - \frac{1}{D_p} \right| \quad (4)$$

where D_f is the depth of the fixation point, D_p is the depth of the rendered pixel, and parameter K is the fitting of A s and the constant relating C and σ_d . The parameter K is scene and user dependent and has to be tuned accordingly. This parameter is tuned based on the quality index of the image proposed by Wang and Bovik [129]. Image degradation such as contrast loss is often associated with blurred images [125]. The value of K is chosen which ensures a sufficient quality index.

A detailed illustration of this DoF effect can be seen in Figure 12. The first two images show the original scene along with its calculated depth map. The last two images show the output for the plane of fixation at different depths. The plane of fixation in the third image is on the vase. Pixels at the vase depth plane appear sharp. The last image shows the output when the plane of fixation is on the tree. It can be seen that the chair (only partially visible as it is occluded by the vase) also forms a sharp image as it is at the same depth as the tree.

3.1.2 Foveated Imaging

Human visual Field-of-View (FoV) is composed of foveal and peripheral regions [116]. The divisions of the human visual system can be seen in Figure 13. The central foveal region is sharp and detailed since the light rays entering the eye form a sharp image on the retinae while the peripheral region lacks fidelity and appears blurred on the retinal image due to the decrease in density of the light sensitive cells in the periphery. The peripheral region can be subdivided into three further categories, namely the near, mid, and far peripheral

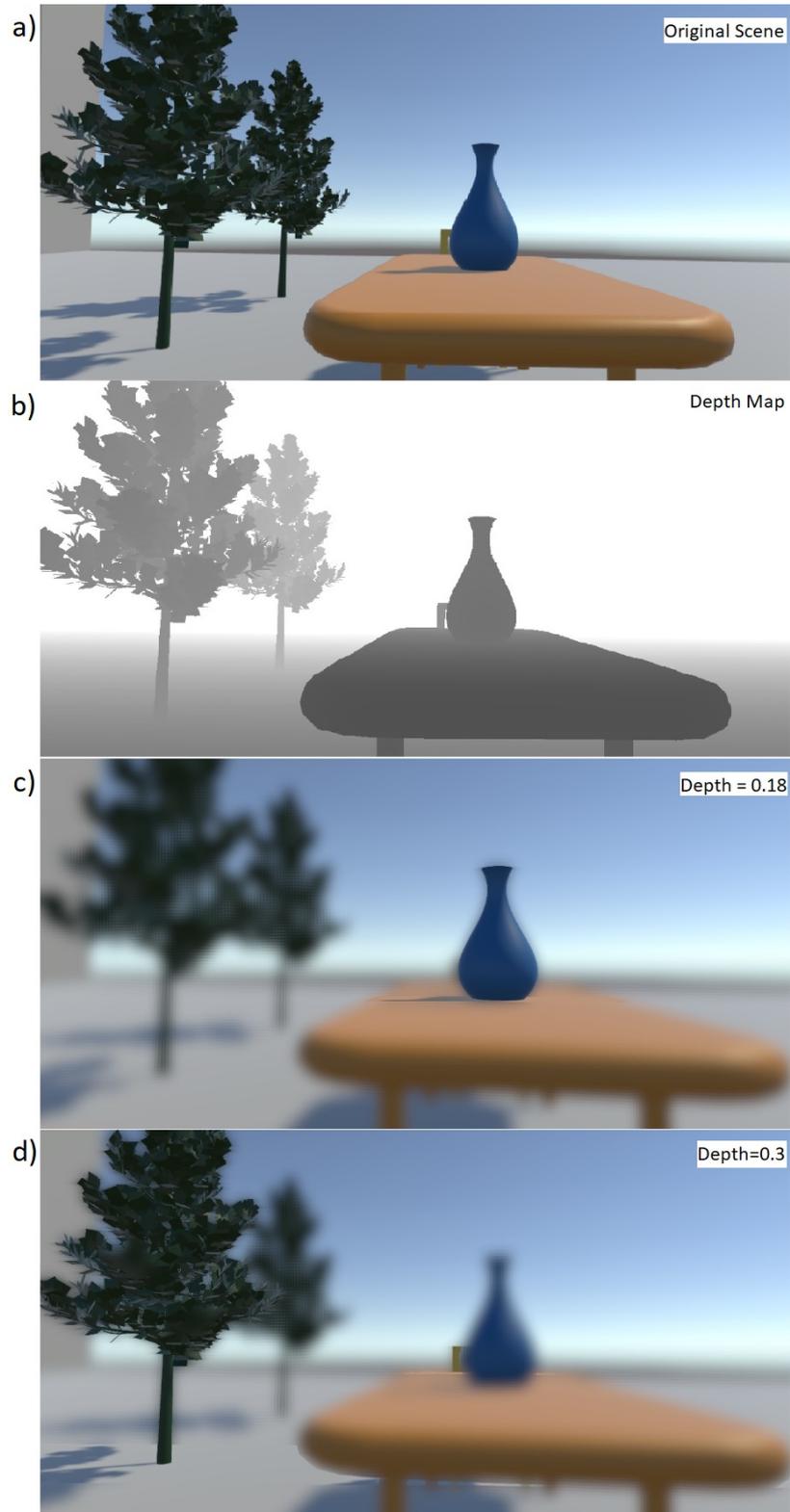


Figure 12: Depth-of-field effects for different planes of fixation. a) Original scene. b) Corresponding depth map. c) Point of fixation on the vase. d) Point of fixation on the tree.

regions. The amount of perceived detail in each region decreases as it moves further from the center. Far peripheral region is only visible to one eye and does not contribute to stereoscopic vision.

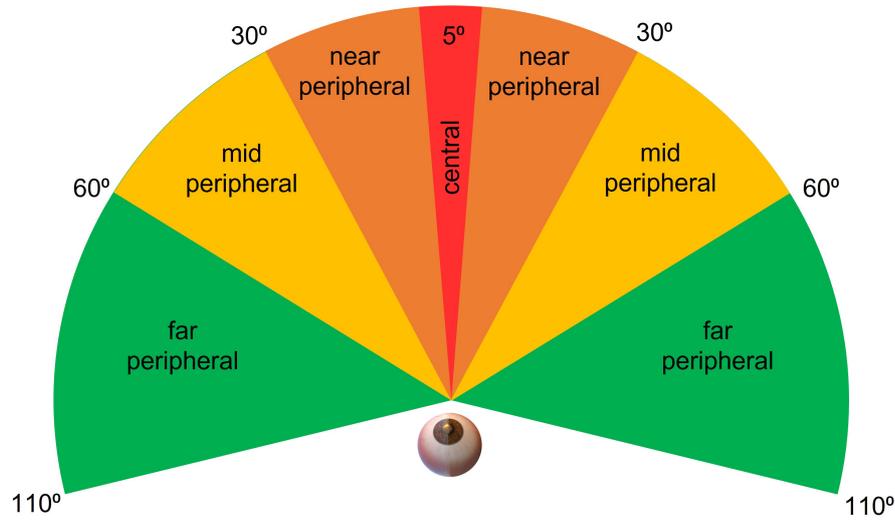


Figure 13: Human field-of-view for both eyes showing the foveal, near, mid, and far peripheral regions.

To implement foveation, the overall scene is divided into three circular sections corresponding to the foveal, near, and mid peripheral regions. The far peripheral region is not visible in modern HMDs due to their optical limitations and thus is not considered. However, the system can be adapted to include it as well by simply increasing the divisions of the rendered scene. Circular divisions were preferred over rectangular ones since it better represents the shape of the lenses present in commercially available HMDs. The fixation point is considered the reference center of the circular regions, and the regions are sketched around it. The central division defines the foveal region and is output without any further processing while the smoothing filter is applied to the other regions. The parameter of the blur σ_f associated with each pixel is determined by the location of that particular pixel in the divided scene. σ_{f_m} for the mid peripheral region is kept as double of σ_{f_n} of the near peripheral region.

An example of this effect is shown in Figure 14. For illustration purposes, the center of the image is assumed as the fixation point. From the left eye view in Figure 14, it can be seen that the circular outlines are quite distinct and cause artifacts in the view which can be uncomfortable for the user in its current form.

3.1.3 Artifact Removal

From Figures 12 and 14, it can be observed that some artifacts exist in the resulting images where there is an abrupt change in the blur σ parameter. In order to eliminate/minimize them, a technique proposed by Perry and Geisler

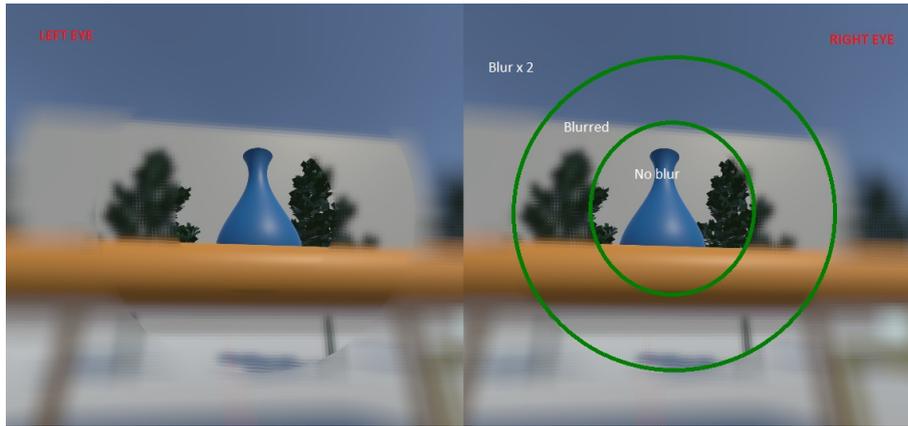


Figure 14: Stereoscopic view of the multi-region foveation output. The central region has no blur applied while the other two regions (highlighted in green for sake of visualization only) have different blurs applied to them.

[96] for blending multi-resolution images using the transfer function of the resolution map is used. Their approach is adapted to the VR system on the transitional regions, i.e., regions with abrupt σ variations. Instead of the transfer function, radial distances between the transitional regions from the fixation point are used in our technique.

A transitional region R_t is introduced and the surrounding regions are defined as either inner R_i or outer R_o based on the location with respect to the fixation point. Likewise, their corresponding radii to the fixation point are defined as r_k with $k = 1, 2, 3$ and $r_k < r_{k-1}$. The blending function $B(x, y)$ is computed by:

$$B_k(x, y) = \begin{cases} 0 & d(x, y) \leq r_k \\ \frac{d(x, y) - r_k}{r_{k-1} - r_k} & r_k < d(x, y) < r_{k-1} \\ 1 & d(x, y) \geq r_{k-1} \end{cases} \quad (5)$$

where $d(x, y)$ is the distance between the rendered pixel coordinates and the pixel coordinates of the point of fixation.

The output of Equation 5 approaches 1.0 as the considered pixel nears the outer region and approaches 0.0 as the pixel nears the inner region. Using the blending function, the output of the smoothing filter is determined by:

$$O(x, y) = B_k(x, y)I_k(x, y) + (1 - B_k(x, y))I_{k-1}(x, y) \quad (6)$$

where $I_k(x, y)$ and $I_{k-1}(x, y)$ are the outputs from the smoothing filters from k th and $(k - 1)$ th regions. This makes sure that a percentage from each blur level is taken based on the location of the pixel in the transitional region to determine the final output resulting in an artifact free scene.

To merge the output of the DoF blur and foveation, pixel-wise σ is computed for both. However, only the smaller σ is used for the smoothing filter for each pixel. Figure 15 shows an example output of the foveated DoF effect. The transitions between high acuity and blurred regions are smoother and the central 20° of eccentricity is free of artifacts.



Figure 15: Example of an output from the foveated depth-of-field blur effect.

3.2 Cybersickness Study

The developed foveated DoF technique was developed with the primary aim of reducing the onset of cybersickness in VR systems. In order to evaluate the effectiveness of the effects, a user study on cybersickness was conducted.

3.2.1 *Experimental Setup*

The developed system was implemented using Unity¹ operating on an Intel Core i7-9700K processor equipped with a NVIDIA GeForce GTX 1080 graphics card. An HTC Vive Pro Eye² device that has an integrated Tobii³ eye tracking system was used for interacting with the user. The HMD has a resolution of 1440×1600 pixels per eye and a 110° FoV. The eye tracking system has an accuracy of 0.5° – 1.1° and a binocular gaze data frequency of 120Hz. The Scosche Rhythm⁴ armband monitor was used to measure the user's heart rate.

A VR roller-coaster environment was designed to induce cybersickness. In order to have control over the experimental conditions, the roller-coaster was custom built in Unity. This allowed us to control and manipulate the experimental parameters, such as velocities, acceleration, and duration of the experiment. The track consists of seesaw and spiral motions placed at different points (see Figure 16). Figure 17 shows the cart velocity and acceleration components over a roller-coaster cycle. Various objects and buildings were closely placed around the roller-coaster tracks to have a cluttered environment. The cluttered environment ensures that the user's focus point changes rapidly and the effect of the foveated DoF blur is more prominent. Figure 18 shows the custom VR environment created for the experiment.

¹ <https://unity3d.com/get-unity/download/archive>

² <https://www.vive.com/us/product/vive-pro-eye/overview/>

³ <https://vr.tobii.com/integrations/htc-vive-pro-eye/>

⁴ <https://www.scosche.com/rhythm-plus-heart-rate-monitor-armband>

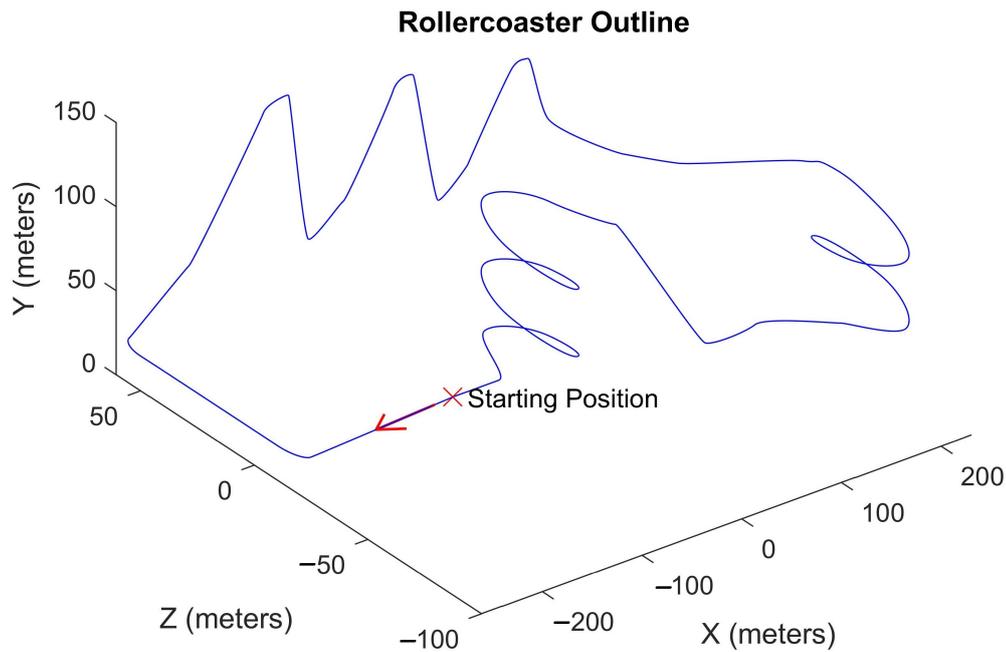


Figure 16: Roller-coaster track outline. The arrow indicates the direction of motion. The coordinate system follows the convention used in Unity, i.e., X: right direction; Y: up direction; Z: forward direction.

3.2.2 Procedure

Data was collected from 18 users (9 males and 9 females) aged from 18 to 46 years (mean 29.3 ± 7.6) who were recruited from students of the University of Genoa. The participants were volunteers and received no reward. All users had normal to corrected-to-normal acuity and normal stereo vision. All users except four were novice VR users.

We considered three experimental conditions: one with our Foveated Depth-of-field Blur (FD) enabled, and one with the Unity's Post-processing Stack Blur (GC)⁵ enabled, and one with No Blur (NB) present. The full fidelity NB condition acts as the control group. The Unity blur GC condition only implements the DoF effect using a seven-pass shader. It also uses the Bokeh effect to introduce spatial blur in the peripheral regions. The size of the Bokeh filter in the Unity blur condition and our foveated DoF condition were kept the same to ensure comparability. The Unity blur does not explicitly support eye-tracking or VR devices so a custom interface was developed to integrate the eye tracking module with the Unity blur effect to provide gaze-contingency.

All users underwent these three conditions in random order, i.e., 1/3rd of the users performed the FD session first, 1/3rd of the users performed the GC session first, and 1/3rd of the users performed the NB session first. This was to ensure that no bias was present in the experiment. Each session only had one condition active. A significant amount of time was provided between each

⁵ <https://docs.unity3d.com/Packages/com.unity.postprocessing@3.1/manual/Depth-of-Field.html>

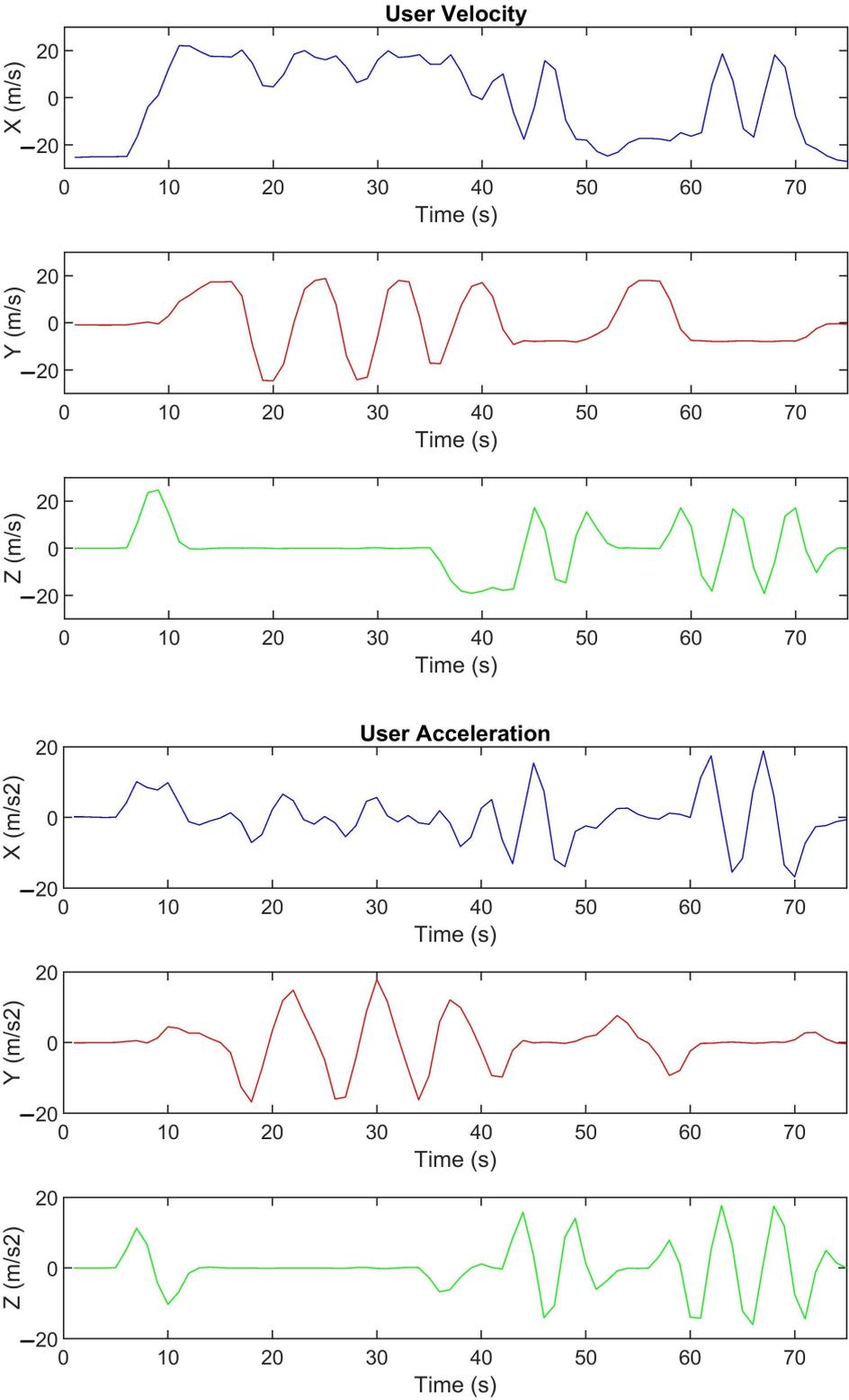


Figure 17: Instantaneous user velocity and acceleration components during each roller-coaster cycle. The coordinate system follows the convention used in Unity, i.e., X: right direction; Y: up direction; Z: forward direction. Seesaw motion: 8–32 s; spiral motion: 36–44 s and 48–64 s.

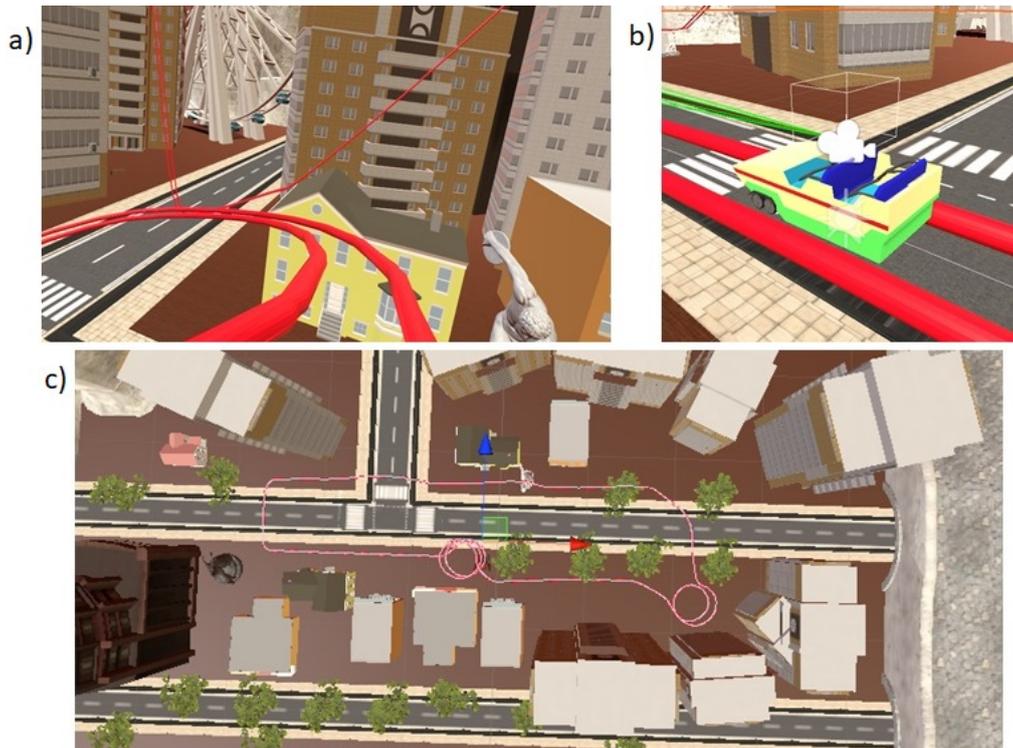


Figure 18: Roller-coaster virtual environment. a) User-view. b) Roller-coaster cart with VR camera attached. c) Top-view of the cluttered environment.

session to all users to recover from the after-effects of the previous condition. Participants were provided with a minimum of a 90-min break between the sessions. Most users opted to undergo the sessions on successive days. Before each session, all participants underwent an eye calibration process.

Each user session lasted for 5 min. This length of the experimental session was determined based on pre-testing trials which suggested that this time-frame was sufficient to induce SS based on the roller-coaster design. For quantitative evaluation, the user's positional data, gaze data, and heart rate were recorded. Heart rate data were recorded at 1 Hz frequency while all other data were recorded at approximately 50 Hz frequency.

To measure SS, users had to fill the Simulator Sickness Questionnaire (SSQ) [54]. The SSQ consists of 16 questions, to be answered on a 4-point Likert scale. The SSQ scores reflect the level of nausea, oculomotor disturbance, disorientation, and overall severeness of induced sickness. The questionnaire was filled by each user immediately before (Pre) and after (Post) each session. To measure user experience between each type of session, the Igroup Presence Questionnaire (IPQ) [98] was used. The IPQ consists of 14 questions, to be answered on a 7-point Likert scale. The IPQ scores reflect the level of spatial presence, involvement, experienced realism, and the general sense of being in the virtual world. Each user filled the IPQ after each session.

3.2.3 Data Analysis and Results

Data gathered from the experimental sessions was analyzed in order to have a better understanding of performance of the developed technique. Figures 19, 20 and 21 show the results of the SSQ questionnaire. It can be observed that our foveated DoF blur has a better performance over the no blur setup. A Wilcoxon rank sum test was performed to compare results of the different conditions. The cross-validation among the Pre states of the users who used different blurred systems showed no significant difference between them. The cross-validation between the Pre and Post states of users during each type of system shows a significant difference, i.e., the experimental environment caused a significant increase in the SSQ scores (see Table 4).

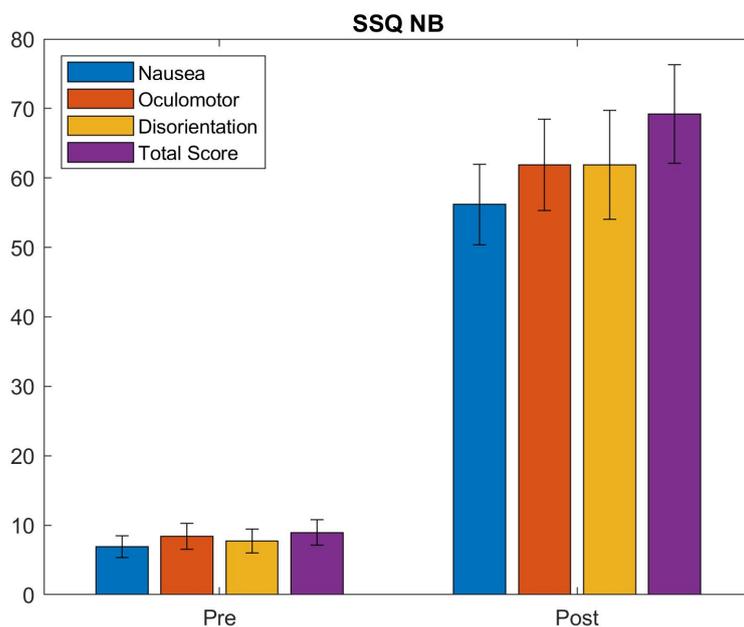


Figure 19: SSQ score for the no blur (NB) sessions. The questionnaire was filled before (Pre) and after (Post) each session. Each plot shows the group mean values and the standard deviations for the three sub-scales and the overall score.

Table 4: The Wilcoxon rank sum test confidence scores between Pre and Post states for the different subcategories of the SSQ test (N–Nausea; O–Oculomotor; D–Disorientation; TS–Total Score).

| | N | O | D | TS |
|----|-------------|-------------|-------------|-------------|
| NB | $p = 0.001$ | $p = 0.002$ | $p = 0.002$ | $p = 0.001$ |
| GC | $p = 0.001$ | $p = 0.003$ | $p = 0.004$ | $p = 0.001$ |
| FD | $p = 0.005$ | $p = 0.004$ | $p = 0.004$ | $p = 0.003$ |

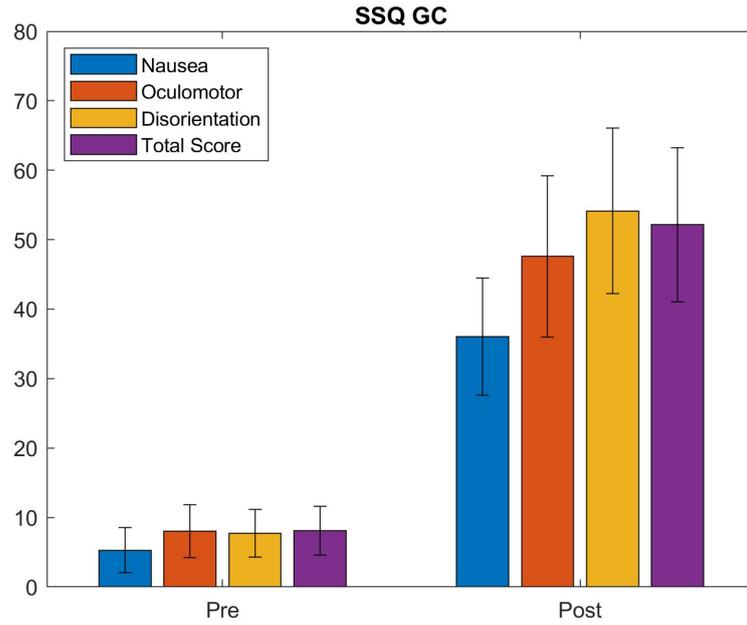


Figure 20: SSQ scores for the unity blur (GC) session. The questionnaire was filled before (Pre) and after (Post) each session. Each plot shows the group mean values and the standard deviations for the three sub-scales and the overall score.

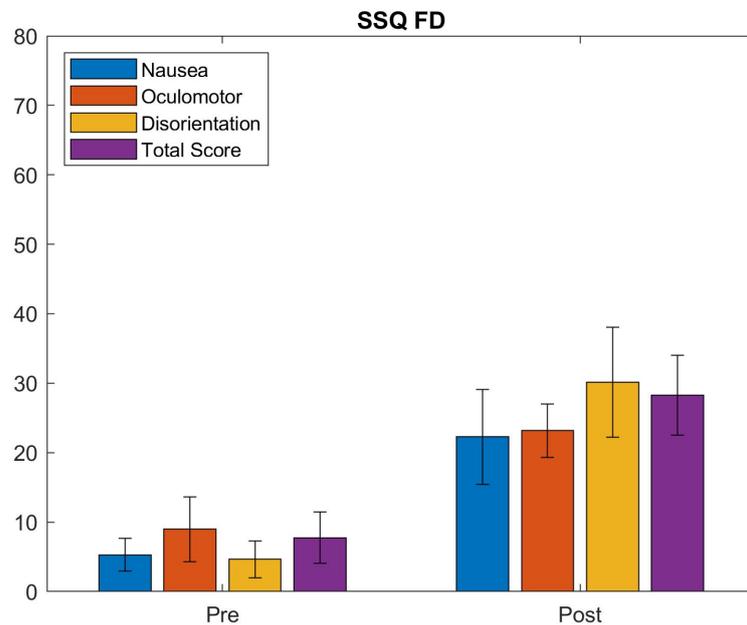


Figure 21: SSQ scores for the foveated depth-of-field effect (FD) session. The questionnaire was filled before (Pre) and after (Post) each session. Each plot shows the group mean values and the standard deviations for the three sub-scales and the overall score.

The differences between the Pre and Post scores (see Figure 22 and Table 5) show that the amount of increase in individual sub-scales is highest in NB sessions ranging between 49–54. The conditions with spatial blur incorporated (GC and FD) show the highest change in disorientation scores which is related to the vestibular disturbances. The amount of induced disorientation is similar in the NB and GC conditions. Although the range of individual sub-scores is different, the results demonstrate that the three conditions produce slightly different patterns of symptomatology, i.e., NB: $D \approx O \approx N$; GC: $D > O > N$; FD: $D > O \approx N$.

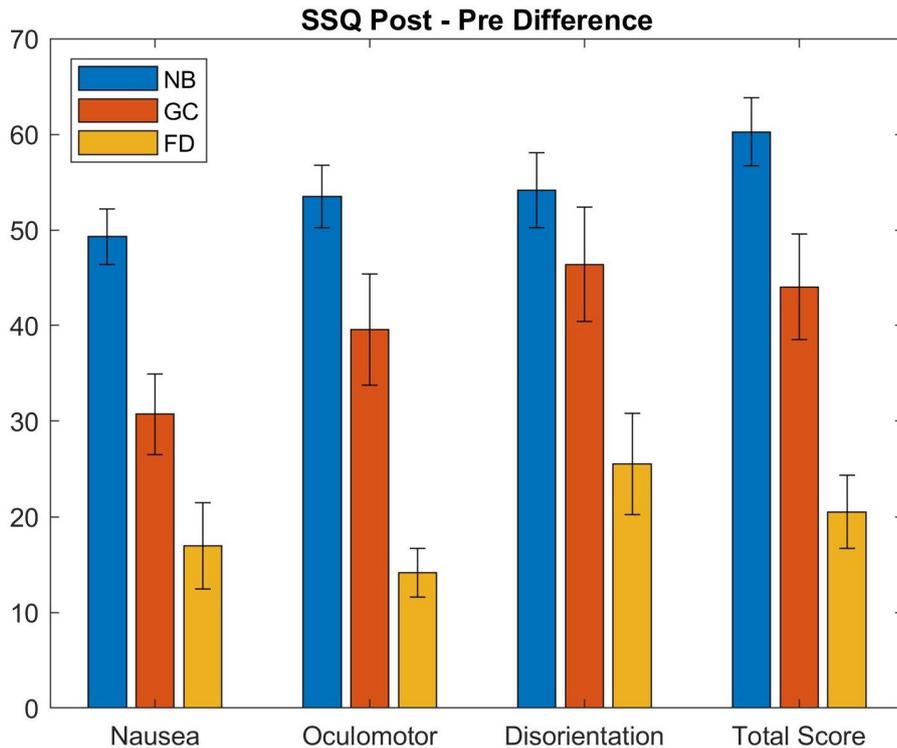


Figure 22: Comparison of the Post-Pre difference of the SSQ scores for each condition. The plot shows the changes in individual SSQ scores between the pre and post experiment conditions.

Table 6 shows a comparison between different techniques discussed earlier in Section 2.3.3 with our foveated DoF effects. We use the difference in the sickness scores between the no effect or full fidelity condition and the best performing parameters for each respective technique. The reported mean SSQ total scores were used where available. One of the user studies did not use the SSQ for the sickness evaluation. The study on peripheral visual effects [13] used a custom questionnaire instead. It can be observed that our foveated DoF blur approach outperforms the other methods.

Figure 23 shows the results of the IPQ questionnaire. A Wilcoxon rank sum test between the samples for Unity blur and our foveated DOF against the ones from the no blur sessions displayed no significant differences in the perceived sense of presence between the users of each type of session.

Table 5: The mean, standard deviation, and 95% confidence intervals of the Post-Pre difference of the SSQ scores for each condition (N–Nausea; O–Oculomotor; D–Disorientation; TS–Total Score).

| | MEAN (STD) | 95 % CI |
|-------|---------------|----------------|
| NB–N | 49.29 (5.81) | [43.14, 55.44] |
| NB–O | 53.48 (6.56) | [46.27, 60.69] |
| NB–D | 54.13 (7.83) | [46.08, 62.19] |
| NB–TS | 60.26 (7.16) | [52.65, 67.85] |
| GC–N | 30.74 (8.44) | [26.91, 34.57] |
| GC–O | 39.58 (11.61) | [33.65, 45.52] |
| GC–D | 46.40 (11.88) | [40.86, 51.94] |
| GC–TS | 44.05 (11.14) | [38.92, 49.17] |
| FD–N | 16.96 (9.07) | [12.97, 20.95] |
| FD–O | 14.18 (5.09) | [10.56, 17.79] |
| FD–D | 25.52 (10.56) | [21.05, 29.99] |
| FD–TS | 20.51 (7.63) | [16.57, 24.42] |

Table 6: Comparison among different techniques for reducing cybersickness. Δs is the reduction in the mean sickness scores between the no effect condition and the best performing condition/parameters.

| TECHNIQUE | HMD | VE / TASK | Δs |
|-------------------------------------|---------------------|---------------------|--------------|
| Dynamic FOV [30] | Oculus Rift DK2 | Reach waypoints | 5.6% |
| Rotation blurring [12] | Oculus Rift DK2 | FPS game | 17.9% |
| FOV reduction (vignetting) [87] | HTC Vive | Follow butterfly | 30.1% |
| Dynamic blurring (saliency) [86] | HTC Vive | Race track | 35.2% |
| Peripheral visual effects [13] | HTC Vive | Find objects | 49.1% |
| Static peripheral blur [71] | HTC Vive Pro | Maze | 54.8% |
| Unity depth blur | HTC Vive Pro Eye | Rollercoaster | 26.9% |
| Foveated DoF (ours) | HTC Vive Pro Eye | Rollercoaster | 66.0% |

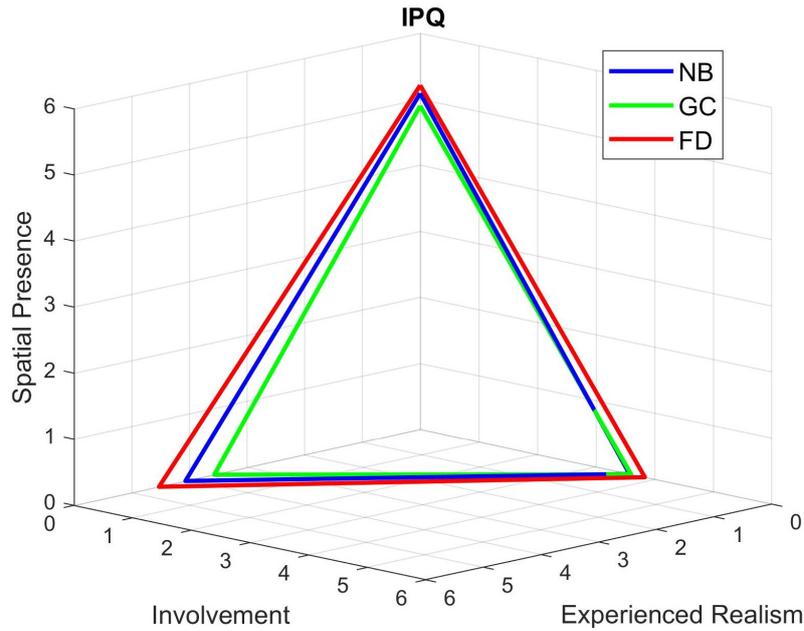


Figure 23: IPQ scores for the cybersickness experiment. The questionnaire was filled after each session. NB: Involvement 3.57, Experienced Realism 4.07, Spatial Presence 5.09; GC: Involvement 3.60, Experienced Realism 3.57, Spatial Presence 4.90; FD: Involvement 3.83, Experienced Realism 4.53, Spatial Presence 5.21.

Another parameter to observe discomfort in users is the heart rate fluctuations. However, at the moment, there is no psycho-physiological parameter that can satisfactorily measure and predict sickness [85, 112], measurements like the finger temperature, reaction time, and heart rate were correlated with cybersickness by Nalivaiko et al. [83]. Figure 24 shows the mean heart rate fluctuations, averaged over all the users, and the standard deviation during a roller-coaster cycle. It can be observed that our foveated DoF blur results in a stable heart rate and only a small increase from the resting heart rate. On the contrary, the heart rate fluctuation in the no blur sessions is more abrupt. The Unity blur sessions have a median performance.

Spatio-temporal data of the user's movement (see Figure 17) suggest that the spiral/torsional motion has a more adverse effect as compared to seesaw motion (up and down movements). We computed the Pearson's correlation coefficients between the heart rate fluctuations and the velocity and acceleration data. The results indicate a strong correlation between each other (r-value: NB = 0.87; GC = 0.81; FD = 0.75). It should be noted that the plots in Figure 24 do not begin from the origin because, in each session, there are four roller-coaster cycles, and the plot shows the mean heart rate of the participant, i.e., only in the first cycle, the participants have the resting heart rate while, in the subsequent cycles, there is an after-effect from the previous cycles.

In order to better understand how a user behaves/interacts with a VR device, the gaze data collected from the experimental sessions was also analyzed. Approximately 4% of the eye tracking data was discarded. This was due to the fact

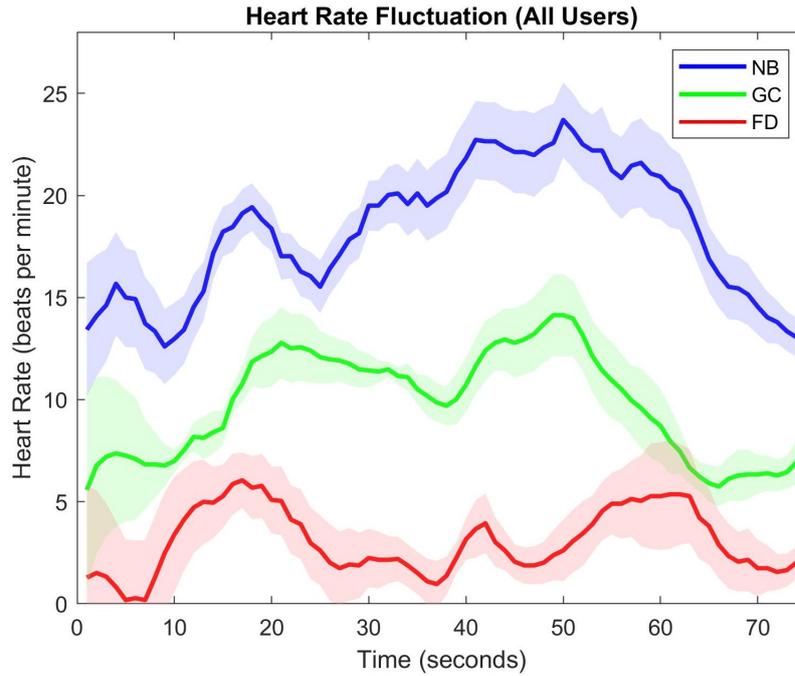


Figure 24: Average heart rate fluctuations from a resting heart rate during a roller-coaster cycle. Origin on the heart rate axis represents the resting heart rate.

that, during the experiment, for some frames, the users either blinked/closed their eyes or there was faulty sensor reading. Figure 25 shows the combined heatmap of all users. It can be observed that the users tend to fixate mostly on the center of the scene.

Positional and orientation data of the user revealed that, when they had to focus on an object further away from the center, they preferred to move their heads instead of just the gaze. This observation is in support of studies conducted by Kenny et al. [55] on First Person Shooter (FPS) games and by Sitzmann et al. [113] on saliency which highlighted that user gaze is mostly directed towards the center of the view (approximately 85% of the time). Consequently, it can be assumed that gaze related user behavior in VR is similar, verifying the assumptions taken in other user studies in the absence of eye tracking [16, 43].

Saccadic movements of the users' eyes were also analyzed. The angular speeds of the eye were computed from the eye tracking data. In humans, angular speed of the eye usually varies between $200^{\circ}/s$ to $500^{\circ}/s$, but can go up to $900^{\circ}/s$ [66]. Thus, for analysis, we considered mainly the saccades having relatively higher speed ranges to determine whether the motion of the eye has any influence on the induced level of cybersickness. Table 7 describes the peak angular speed measured for each user and how many times speeds of greater than $200^{\circ}/s$ was achieved. It can be noticed that, during our blur algorithm integrated sessions, saccades were shorter/slower as compared to the other sessions. A Kolmogorov-Smirnov test was performed on the angular speed data. The statistical analysis showed a significant difference in the distribution

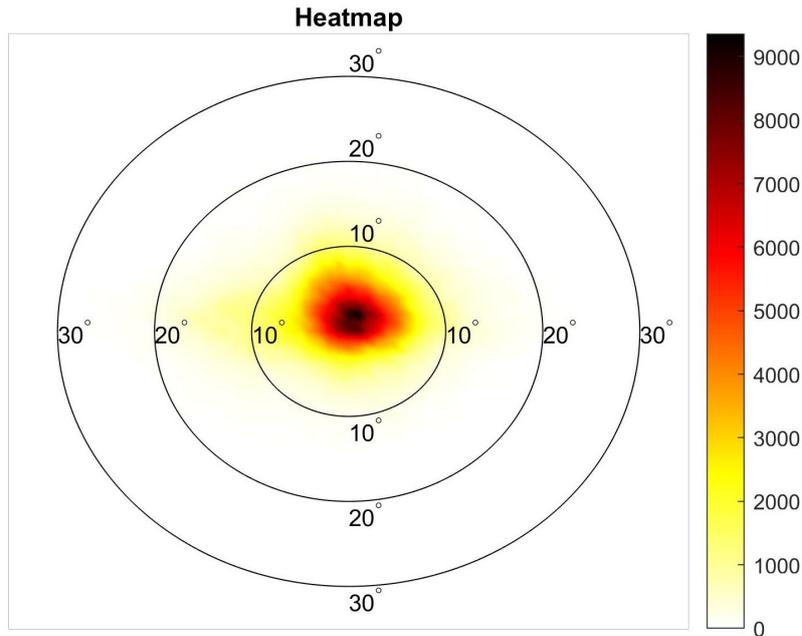


Figure 25: Heatmap of the visual field for user gaze combined for all sessions performed. The circles are centered at the center of the HMD screen and indicate the visual angle (e.g., the 10° circle represents the central 20° of visual eccentricity). The colors represent how frequent the user fixated at that particular location on the HMD screen with white representing 0 and black representing 9358.

of the angular speed data with a 95% Confidence Interval (CI) for the three conditions.

Figure 26 describes the number of occurrences for speeds higher than 350°/s. It should be noted that speeds lower than this value had a similar trend in all the three conditions, so they are not shown here. Previously, the SSQ revealed that the level of sickness in the NB sessions is higher than our blur system. Correspondingly, there may be a correlation between the occurrences of faster saccades with the level of induced sickness. The temporal analysis revealed that higher peaks were observed mostly during the seesaw motion.

A possible explanation for lower amplitudes in our system could be that the encompassed blur reduces the amount of detail in the periphery. Consequently, the saccades are shorter. This peripheral reduction mimics the popular approach of reducing the FoV to minimize cybersickness [30]. However, in our approach, the peripheral content is still visible, albeit at a lower acuity; thus, the level of presence is not compromised unlike the FoV reduction approach.

We also analyzed how age affects cybersickness. It is widely assumed that motion sickness is more prevalent in younger participants; however, past studies on cybersickness in VR have revealed contradicting conclusions. Studies by Arns et al. [4] and Hakkinen et al. [38] revealed that younger participants suffer less from SS, whereas a meta-analysis by Saredakis et al. [106] showed the opposite. We divided the participants into two groups, young and old. The

| USER | NB | | GC | | FD | |
|-------|---------|--------|---------|--------|---------|--------|
| | >200°/s | PEAK | >200°/s | PEAK | >200°/s | PEAK |
| AT | 106 | 810°/s | 89 | 502°/s | 59 | 354°/s |
| CT | 132 | 784°/s | 108 | 544°/s | 96 | 497°/s |
| EV | 88 | 859°/s | 99 | 743°/s | 74 | 556°/s |
| GB | 136 | 546°/s | 90 | 650°/s | 101 | 549°/s |
| HR | 115 | 773°/s | 125 | 663°/s | 97 | 568°/s |
| KK | 78 | 593°/s | 71 | 539°/s | 84 | 542°/s |
| LH | 132 | 731°/s | 93 | 707°/s | 103 | 581°/s |
| MB | 87 | 581°/s | 116 | 582°/s | 63 | 431°/s |
| MM | 112 | 703°/s | 95 | 697°/s | 88 | 553°/s |
| ND | 101 | 802°/s | 107 | 718°/s | 71 | 655°/s |
| NR | 86 | 824°/s | 119 | 702°/s | 105 | 603°/s |
| OQ | 88 | 595°/s | 92 | 629°/s | 95 | 612°/s |
| SA | 106 | 697°/s | 105 | 735°/s | 94 | 514°/s |
| SR | 97 | 710°/s | 82 | 657°/s | 68 | 570°/s |
| TB | 113 | 688°/s | 89 | 617°/s | 87 | 545°/s |
| UG | 115 | 591°/s | 84 | 623°/s | 89 | 511°/s |
| US | 92 | 597°/s | 111 | 502°/s | 89 | 533°/s |
| YK | 67 | 351°/s | 142 | 661°/s | 67 | 508°/s |
| Total | 1999 | 859°/s | 1923 | 743°/s | 1619 | 655°/s |

Table 7: Comparison of angular speed during saccadic motion for each user. Number of occurrences of speeds greater than 200°/s and the peak speed observed are shown.

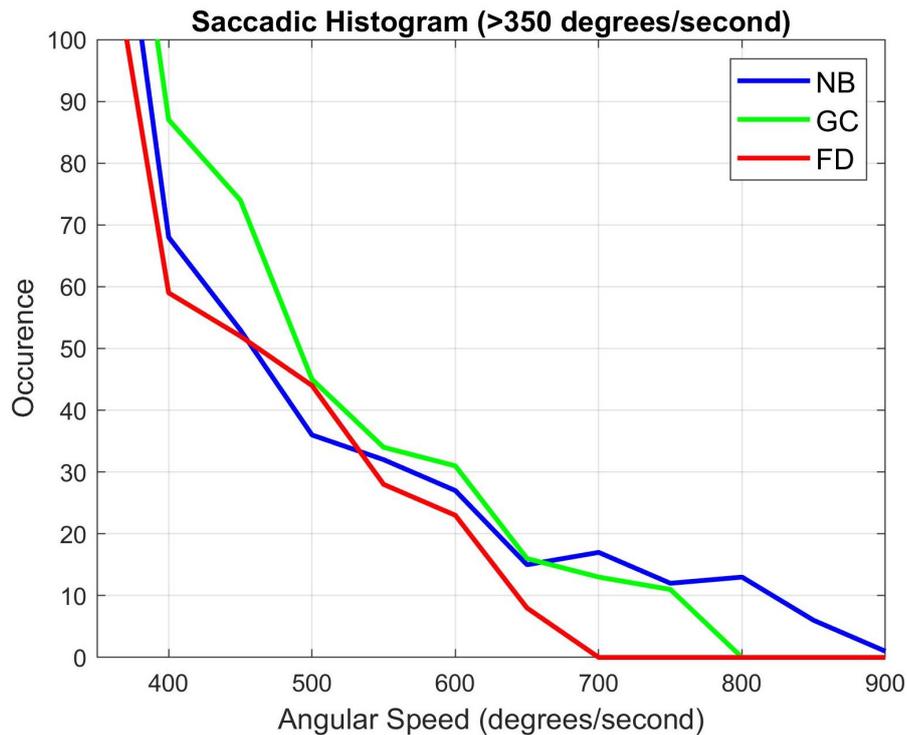


Figure 26: Histogram for angular speed greater than $350^{\circ}/s$ of the eye for all users during a saccade.

younger group is comprised of people aged between 18 and 26 years while the rest comprised the older group. There were 10 users in the younger group and eight users in the older group. Figure 27 shows the difference in the total score of SSQ for the two age groups. A Wilcoxon rank sum test was performed. In the FD condition, no statistical difference was found in the SSQ scores and heart rate distributions ($p > 0.45$). However, in the NB and GC conditions, the older participants suffered more from cybersickness ($p < 0.05$).

The participants were also sub-grouped with respect to gender. Figure 28 shows the difference in the total score of SSQ for the two gender groups. A Wilcoxon rank sum test was also performed; however, no statistically significant difference was found between the two groups ($p > 0.65$). It should be noted that age and gender do not exclusively influence sickness. Factors such as neuroticism, prior VR experience, etc. also simultaneously affect cybersickness. Wider studies on age and gender may be required to fully understand how these factors influence cybersickness as highlighted by Chang et al. [17].

Using the data recorded from the cybersickness user study, the frame processing times were also computed in order to have a better understanding of the computational overhead added by the blurring techniques. Data from the NB sessions acted as the reference for comparison. The average processing times and their equivalent frame rates are summarized in Table 8. There is no overlap between the processing time of the three conditions within a 95% CI. It can be observed that our system offers better computational performance than

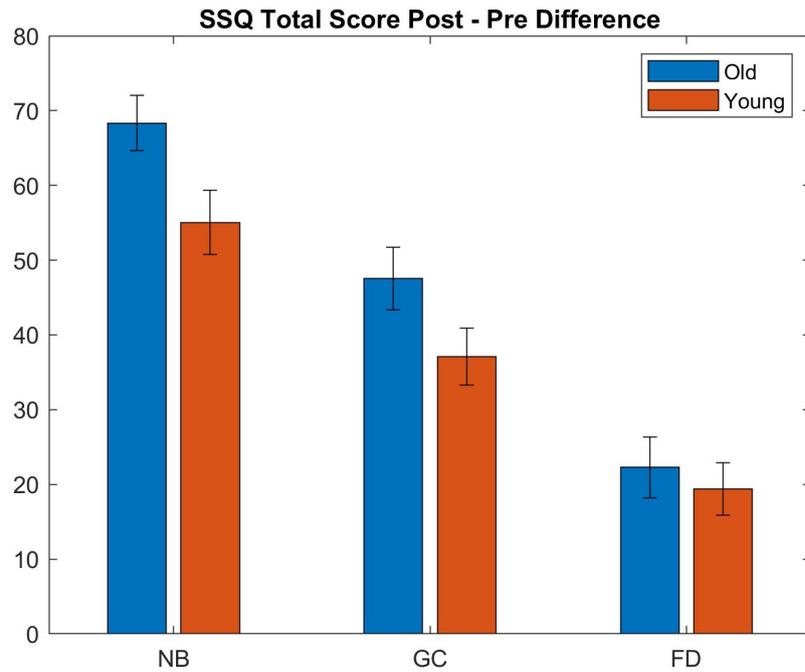


Figure 27: Comparison of the Post-Pre difference of the SSQ scores for each condition with respect to age groups. The plot shows the changes in individual SSQ total scores between the Pre and Post experiment conditions for the two age groups. Old: NB 68.34, GC 47.55, FD 22.26; Young: NB 55.03, GC 37.06, FD 19.38.

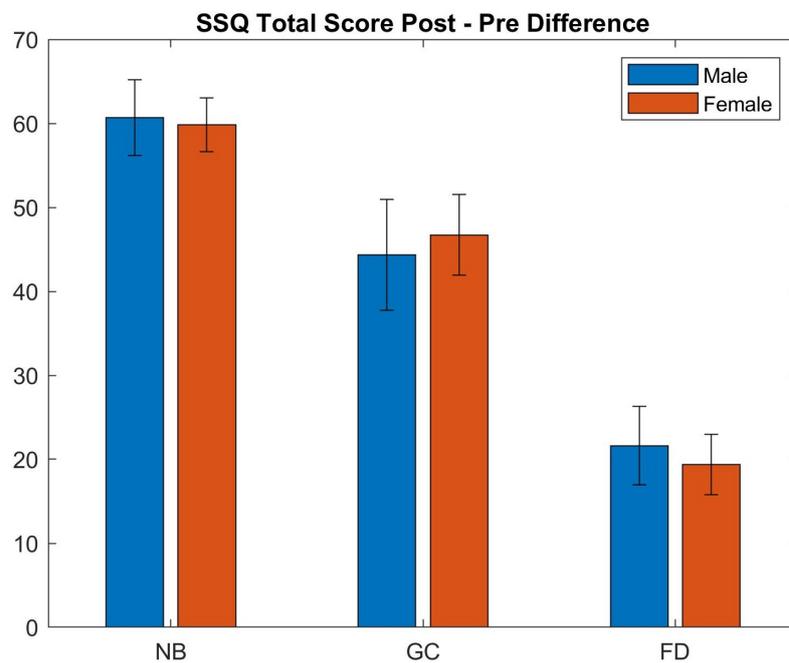


Figure 28: Comparison of the Post-Pre difference of the SSQ scores for each condition with respect to gender groups. The plot shows the changes in individual SSQ total scores between the Pre and Post experiment conditions for the two age groups. Male: NB 60.67, GC 44.37, FD 21.63; Female: NB 59.84, GC 46.72, FD 19.39.

Unity's blur even though the built-in blur in Unity only applies the DoF effect, whereas our system processes two different types of blur.

Table 8: Frame rate comparison.

| SYSTEM | PROCESSING TIME | 95 % CI | FRAME RATE |
|--------|-----------------|--------------------|------------|
| NB | 15.9 ms | [15.9 ms, 15.9 ms] | 63 Hz |
| GC | 17.2 ms | [17.1 ms, 17.3 ms] | 58 Hz |
| FD | 16.7 ms | [16.6 ms, 16.8 ms] | 60 Hz |

3.2.4 Outcome

The aim of this work was to develop a spatial blurring technique inspired by nature for VR applications. The spatial blurring system adopts a hybrid approach to incorporate foveation and DoF effects into the rendering pipeline. The developed technique implemented using a shader program can be used as a post-processing step to any VR application. It should be noted that for optimal performance, the VR headset used should be equipped with an eye tracking system. The developed algorithm provides sufficient frame rate to not affect the sense of presence.

The user study provides clues to how the onset of cybersickness can be mitigated using spatial blurring effects. Two types of blurring systems were used to evaluate the performance. The results show that spatial blurring can indeed reduce the amount of induced cybersickness as supported by both subjective and objective measures. Overall, there was a 27% reduction in sickness scores for the Unity DoF blur setup. Using our effects, the reduction was much higher as the sickness scores improved by 66%. These improvements are with respect to the normal viewing condition. The randomisation of the order of the sessions ensured that no bias was present in the experimental setup.

Although the use of SSQ for assessing cybersickness has existed for many years, using physiological measures have gained popularity only recently. Self-reported questionnaires only allow analysis of cybersickness before and after the experimental session. In some experimental procedures, it can be done at fixed time instances. However, the use of physiological signals allows continuous monitoring of how the user is feeling. Incorporating physiological signals such as heart rate allowed us to establish which part of the experimental session contributed more to the sickness. From example, as observed from Figure 24, circular or spiral motion contributed more adversely to the onset of cybersickness as compared to straight line motion.

4

Depth Perception in Virtual Environments

Depth perception is the visual ability to perceive the environment in 3D and the distance of objects. In humans, depth perception is performed by a variety of depth cues. These can be typically classified as monocular or binocular. Monocular cues are the ones which can be represented in only Two-Dimensional (2D) and requires only one eye. Motion parallax, occlusion and relative size are examples of such. On the other hand, binocular cues makes use of stereopsis. Convergence is an example. When humans use a XR device such as an HMD, they are often presented with conflicting cues. This diminishes their ability to judge object distances properly. The focus of this chapter is to understand how visual stimuli can be adjusted to allow users to better perceive depth in XR systems. For this, first a user study is presented on the spatial blurring technique developed earlier. Then, an inverse blurring system is described and validated through two user studies which can potentially reduce Vergence Accommodation Conflict (VAC), thus, allowing users to perceive object depth better.

4.1 Depth Perception Study

In literature, many researchers have proposed that blurring, especially when done using DoF effects, can potentially improve the mismatch present between accommodation and convergence [16, 72]. As a first step to understanding how visual stimuli can affect depth perception in VR devices, a user study was conducted using the spatial blurring effects developed previously.

4.1.1 *Experimental Setup*

The system was implemented using Unity¹ operating on an Intel Core i7-9700K processor equipped with a NVIDIA GeForce GTX 1080 graphics card. HTC

¹ <https://unity3d.com/get-unity/download/archive>

on a virtual keypad integrated into the scene using a laser pointer attached to the HTC Vive Pro controller (see Figure 30).

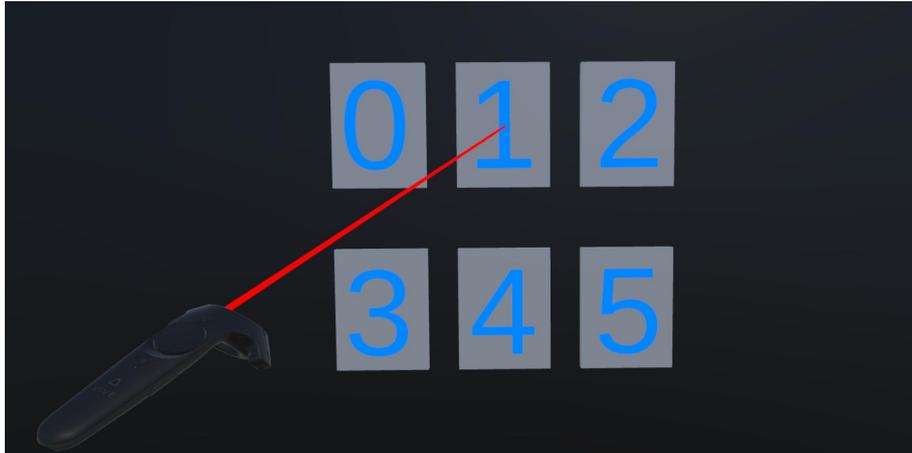


Figure 30: Keypad display used to obtain answers from the test subjects.

Figure 31 shows a sketch of a trial. The user is positioned approximately 2m from the table. Objects to be observed on the table are shown in colored circles and squares. The red square is the reference object while the green objects are the ones placed at the same scene depth as the reference object. The user has to look at the red square and perceive how many other objects are at the same distance from the user. In this scenario, the correct answer is 2.

Each user performed three sessions with each session having 30 trials. In each session, 15 trials were without the foveated DoF effect and 15 with the blur enabled. The no blur sessions acted as the control group. The sequence was random, as a result the two conditions are switched randomly during the experiment, not introducing any bias. Likewise, for each session, the order of reference objects was also randomly generated without repetition. User answers for depth perception were recorded for qualitative evaluation.

After completing the experiment, the subjects were asked to fill a subjective questionnaire in order to evaluate their experience with using the system. The open questionnaire was composed of the following questions:

- Q1) Do you feel any kind of dizziness after using the system?
- Q2) Did you notice any artifacts while changing the fixation point / were the transitions from blur to sharp and vice versa smooth?
- Q3) Which system was more realistic/immersive?
- Q4) Which system do you prefer for depth perception?

4.1.3 Data Analysis and Results

The main objective parameter to determine user performance is the accuracy of the perceived depth during each trial. We computed the true distances of each

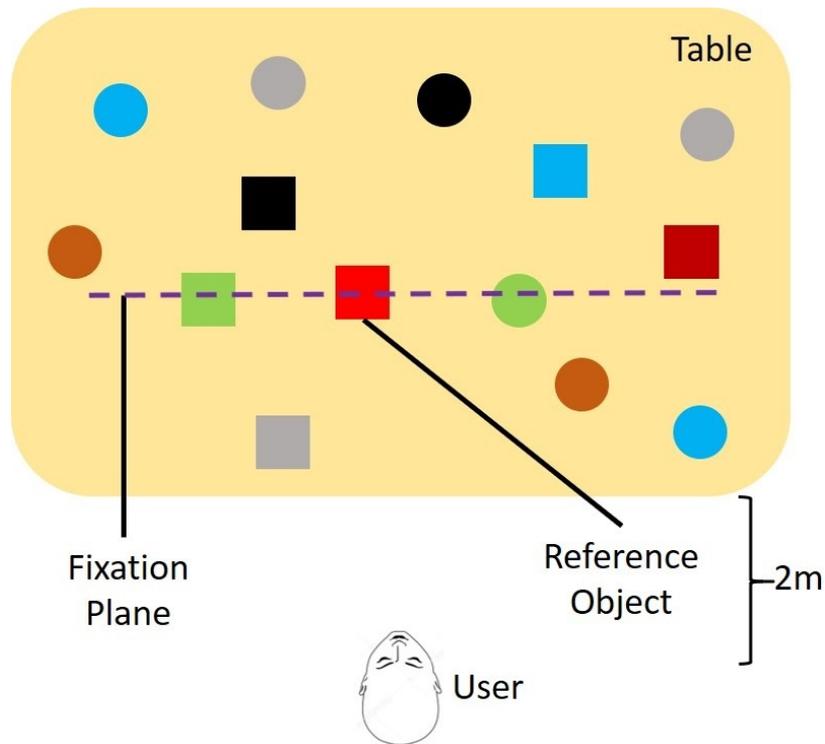


Figure 31: A sketch of the test scenario. Colored squares and circles are the objects to be observed (see Figures 29 and 30 for the actual virtual scene observed).

object from the user view point in each session and compared it with answers provided by the user. The number of trials where the user answered correctly is summarized in Table 9. User performance in the foveated DoF sessions was better. It can be noted that depth perception in virtual environments is not an easy task as compared to the real world.

We further investigated the user responses by calculating the error in their outputs. The error function used was the mean absolute average. Performance of each user can be seen in Figure 32. Comparing the errors between the trials with blur disabled and enabled, it can be observed that the performance either improved considerably or stayed the same. User performance did not deteriorate for any subject. Two of the users (GB and RK) had a high error reduction in their output.

In order to understand whether the user output was biased towards one side (giving a lower output than the true value or vice versa), the user performance was compared by computing the mean error (see Figure 33). It can be noticed that most of the users were overestimating the objects at the same scene depth, i.e., they gave a higher answer than the true value.

A comparison of the combined performance of the subjects is presented in Table 10. An overall error reduction of approximately 27% for depth perception was observed.

Generally, users found the transitions smooth and did not perceive any noticeable artifacts. Ten of the users preferred the system with blur enabled. One user (KM) indicated that sometimes the blur provided a distraction while an-

Table 9: Accuracy of each user.

| USER | BLUR DISABLED | BLUR ENABLED |
|---------|---------------|--------------|
| AT | 6.7% | 17.8% |
| CB | 17.8% | 20.0% |
| DG | 28.9% | 31.1% |
| GB | 4.4% | 17.8% |
| KK | 11.1% | 26.7% |
| KM | 20.0% | 24.4% |
| NF | 15.6% | 20.0% |
| NZ | 8.9% | 13.3% |
| RH | 20.0% | 31.1% |
| RK | 4.4% | 26.7% |
| TK | 22.2% | 33.3% |
| YK | 31.1% | 35.6% |
| Overall | 15.9% | 24.8% |

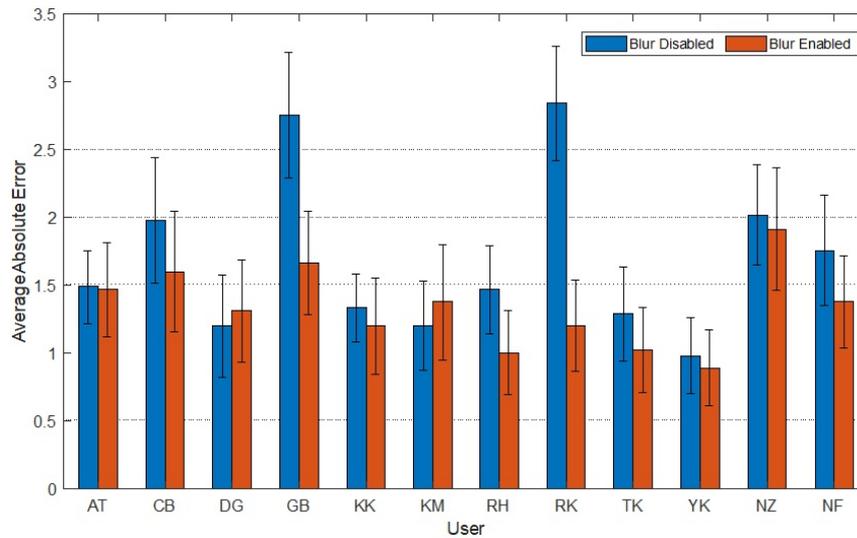


Figure 32: Mean absolute error with its standard deviation for each subject.

Table 10: Group mean error along with the standard deviation.

| ERROR | BLUR DISABLED | BLUR ENABLED |
|---------------|-----------------|-----------------|
| Mean Absolute | 1.69 ± 0.20 | 1.33 ± 0.09 |
| Average | 1.04 ± 0.33 | 0.54 ± 0.18 |

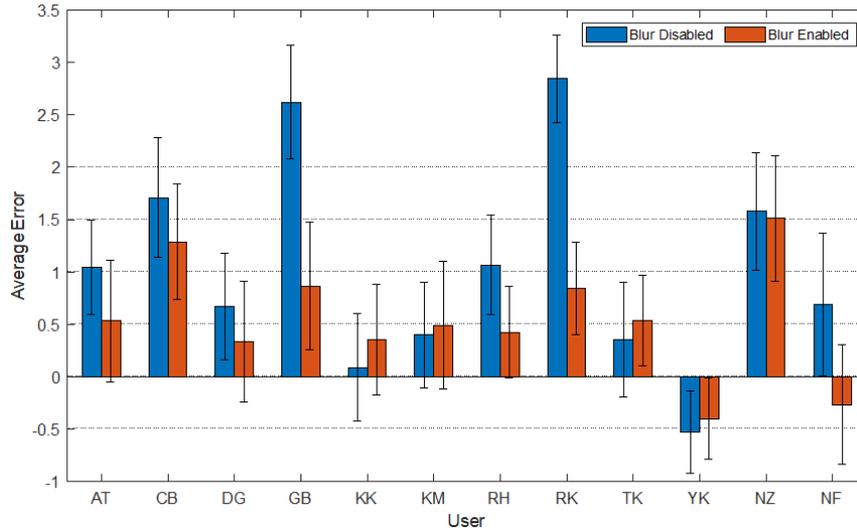


Figure 33: Mean error with its standard deviation for each user.

other (DG) indicated that the blur caused confusion in completing the tasks. It should be noted that none of the subjects knew what the purpose of the blur effect was or how it was calculated prior to completion of the experiment. Only 1 user (KM) indicated about feeling a minor headache after using the system though he was slightly nauseous before using the system as well. None of the other users felt any such symptoms. One user (GB) who was familiar with VR devices found the blur enabled system to be more realistic and provided with a better sense of immersion.

Due to the random arrangement of objects, in some trials a few smaller objects were occluded by bigger objects placed in the line of sight. The test subjects were unable to notice their presence. Similarly, sometimes the users had to move their head to bring the reference object into focus, as a result, some objects were out of the FoV. These factors accounted for some of the errors in the user performance.

4.1.4 Outcome

The aim of this user study was to understand how spatial blurring affects depth perception in VR systems. The task was to estimate object depths solely through visual stimuli. To this aim, users were shown a cluttered virtual environment and they had to estimate objects at a particular distance. All users underwent two experimental conditions, one with the foveated DoF effects developed in Section 3.1 and the other without any effects.

The data analysis showed a 27% reduction in error with the spatial blur enabled. The general trend showed that users were overestimating the number of objects at the reference depth plane. It was also observed that incorporating spatial blurring effects is not detrimental to user performance. Overall, user performance was not high, indicating that VR users have difficulty perceiving depth properly in VEs.

4.2 Inverse Blurring

When light rays enter the eyes through the cornea, they diffract to form a focused image on the retina. The diffraction pattern can be modeled as a Point Spread Function (PSF). If this PSF is known, it is possible to identify the optical requirement of corrective lenses that are necessary to adjust the light rays entering the eyes. In the image processing domain, this operation can be analogously expressed by a deconvolution operator [90]. Convolution is the technique popularly used to apply filters to images. Deconvolution is the inverse process of convolution. Primarily, it is a computationally intensive process that can be used to recover the blurring in an image [88]. This process can also be referred to as inverse blurring or deblurring. Common deconvolution algorithms include inverse filtering, Wiener filtering and iterative approaches such as Lucy-Richardson algorithm. With the recent boom in the deep learning domain, many new algorithms have been proposed [60, 63, 100, 107, 137, 140]. However, the iterative processes and the deep learning models are not viable solutions when it comes to AR/VR applications as fast processing is of utmost importance in order to update the scene in real-time and these methods have a very high processing and memory cost. For this reason, we developed our inverse blurring system based on the Wiener deconvolution.

4.2.1 Wiener Deconvolution

Generally, given an image i , the convolution operation with a blurring filter f can be defined as:

$$b = f * i + n \quad (7)$$

where $*$ is the convolution operator, n is the noise in the system and b is the resulting blurred image.

In the Fourier or frequency domain, Equation 7 can be written as:

$$B = FI + N \quad (8)$$

where B , F , I and N are the Fourier transforms of b , f , i and n respectively. Typically, this blurred image can be corrected through the inverse procedure:

$$I' = \frac{B - N}{F} \quad (9)$$

However, this approach is not optimal as it amplifies the noise in the system. Instead, the Wiener deconvolution is considered more optimal for this type of task as it is insensitive to small variations in the signal power spectrum. The Wiener filter assumes that the image is modeled as a random process whose 2^{nd} order statistics along with noise variances are known. The image restoration model can be written as:

$$I' = H_W B \quad (10)$$

where I' is the estimation of the original image and H_W is the Wiener filter.

Assuming the PSF is real and symmetric and the power spectrum of the original image and the noise is unknown, then the Wiener filter can be defined as:

$$H_W = \frac{H}{|H|^2 + \frac{1}{SNR}} \quad (11)$$

where SNR is the Signal-to-Noise Ratio (SNR) and H is the estimate of the PSF of the blur. For out-of-focus distortions such as those that are naturally present in human visual system, a circular PSF (see Figure 34) is considered a good approximation [36]. Such a PSF can be defined by only one parameter, R which is the radius of the circle. Therefore, only two parameters (R and SNR) need to be tuned to implement the Wiener deconvolution.

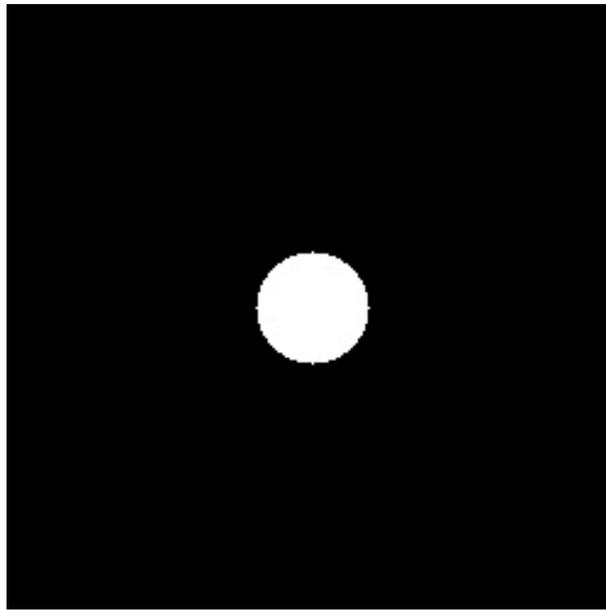


Figure 34: Example of a circular point spread function.

4.2.2 Implementation and Parameter Tuning

The Wiener filter was implemented using *OpenCV for Unity*³ package by Enox Software. Each virtual scene frame is processed as an independent image. The image needs to be transformed to the frequency domain using Discrete Fourier Transform (DFT). However, a limitation arises. Whilst most XR devices operate in rectangular images, DFT can be applied only to square images. Hence, a pre-processing step needs to be undertaken. There are two ways to transform a rectangular image into a square one:

- **Cropping:** The image is cropped by removing the unwanted regions in the periphery. A squared region centered at the center of the image is

³ <https://assetstore.unity.com/packages/tools/integration/opencv-for-unity-21088>

extracted to be used by the inverse blurring algorithm. This approach leads to loss of data. A work around to the lost data is to superimpose the restored image onto the original image. However, this leads to the peripheral regions being unprocessed resulting in pronounced artifacts.

- **Image Resizing:** The image is resized to change the aspect ratio. Typically most entertainment devices are designed for a 16:9 aspect ratio. The image is transformed to a square aspect ratio. This approach has the advantage that it can be incorporated into the down-sampling process that is already present in the system to lower computation costs. Also, less information is lost as the restored image can be resized back to the original aspect ratio during the up-scaling post-process.

The latter approach is adopted to obtain a down-sampled square image at half-resolution. The resulting image is transformed into the frequency domain. Then the Wiener filter is computed using the circular PSF in the frequency domain. The filter is applied to the image and transformed back into the spatial domain using Inverse Discrete Fourier Transform (IDFT). The overall process is illustrated in Figure 35. The process is performed individually on each Red-Green-Blue (RGB) channel. Although the implementation allows for each RGB channel to have a different set of parameters (R and SNR), for the purpose of this work, same values are used.

The parameter R is dependent on the amount of blur present in the image while SNR is a measure based on the noise present in the system. Since, R is more significant to the restoration/deblurring process, it is recommended to tune it first. R can be determined based on the distance between the user and the objects in view. Since the amount of blur present in the human visual system has already been discussed previously in Section 3.1, it is not discussed here.

To obtain the optimal values of SNR corresponding to each R value, a tuning process was carried out. A variety of virtual scenes containing virtual objects were created. They were blurred using the spatial blurring technique discussed previously in Section 3.1. The inverse blurring was applied to the resulting blurred images. This ensured that the original image and the values of parameter R are already known and the value of SNR corresponding to each value of R can be determined.

To achieve this, the blurred images with known R were deblurred with different values of SNR . Example outputs are illustrated in Figure 36. To assess the quality of deblurring, image quality was measured for each image using the metrics discussed in Section 2.4.3 such as PSNR and mean-SSIM. VIF was also considered but not utilized in the final tuning process as it is developed based on natural scenes statistics which do not always work well with artificial/virtual scenes.

The corresponding PSNR and mean-SSIM are shown in Figures 37 and 38 respectively. The plots show the best fit curves. A cut-off value of 25 was chosen for PSNR. Similarly, 0.8 was chosen as cut-off for mean-SSIM. Using these thresholds, the range of possible values for SNR corresponding to each R was obtained.

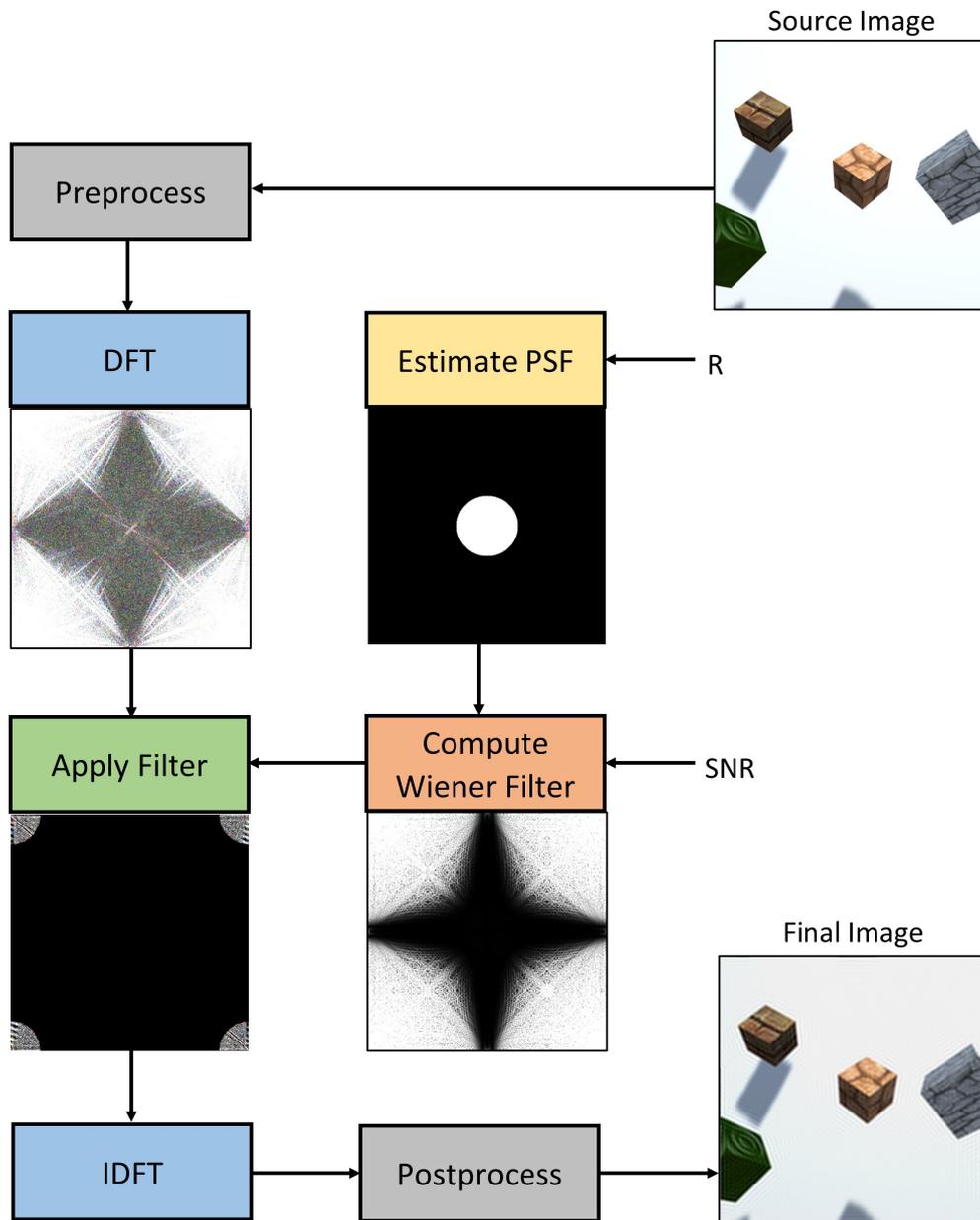


Figure 35: Process flow of the proposed inverse blurring method showing the intermediate outputs.

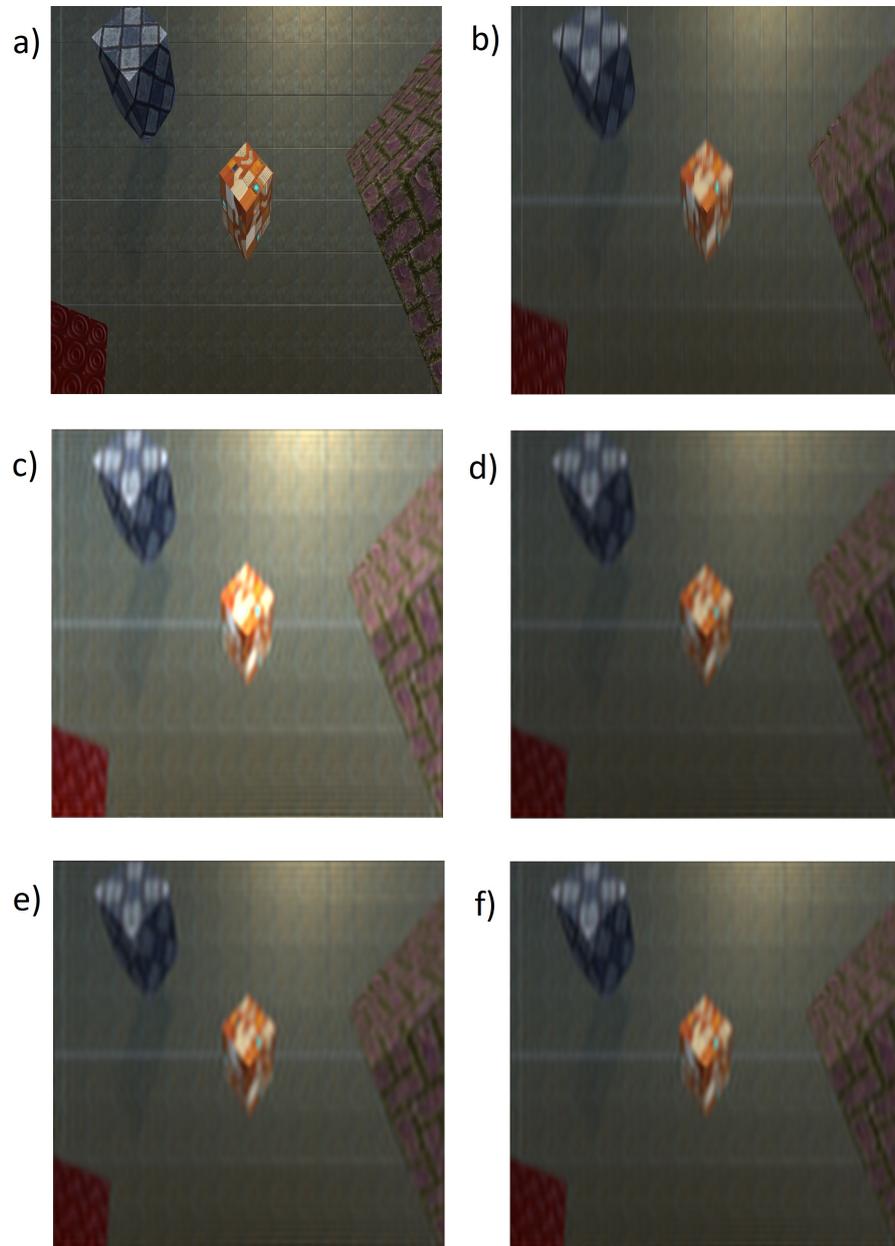


Figure 36: Illustration of the parameter tuning process. a) Source image. b) Blurred image. c) Output with parameters: $R=32$, $SNR=4500$. d) Output with parameters: $R=50$, $SNR=7500$. e) Output with parameters: $R=65$, $SNR=8200$. f) Output with parameters: $R=70$, $SNR=10000$.

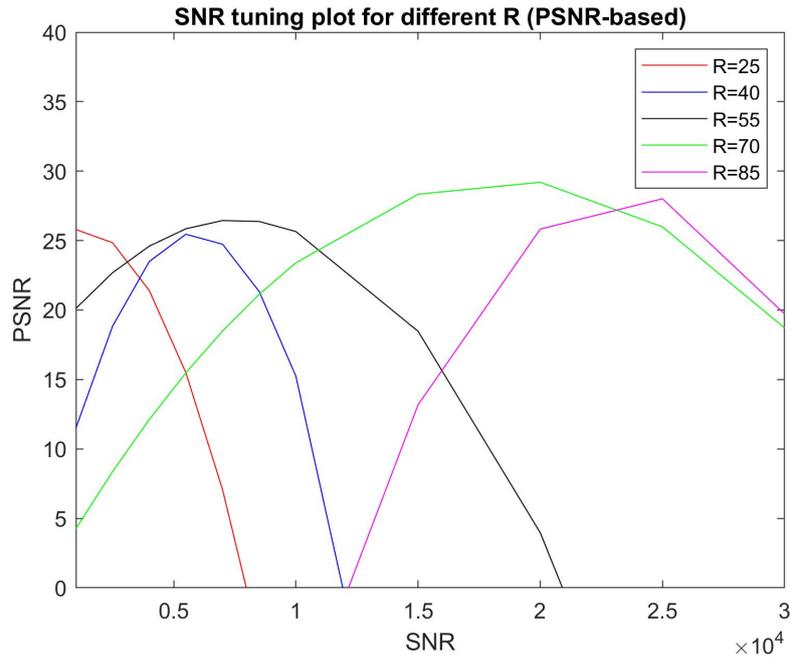


Figure 37: PSNR values versus SNR computed for different values of R. The curves represent the best fit line.

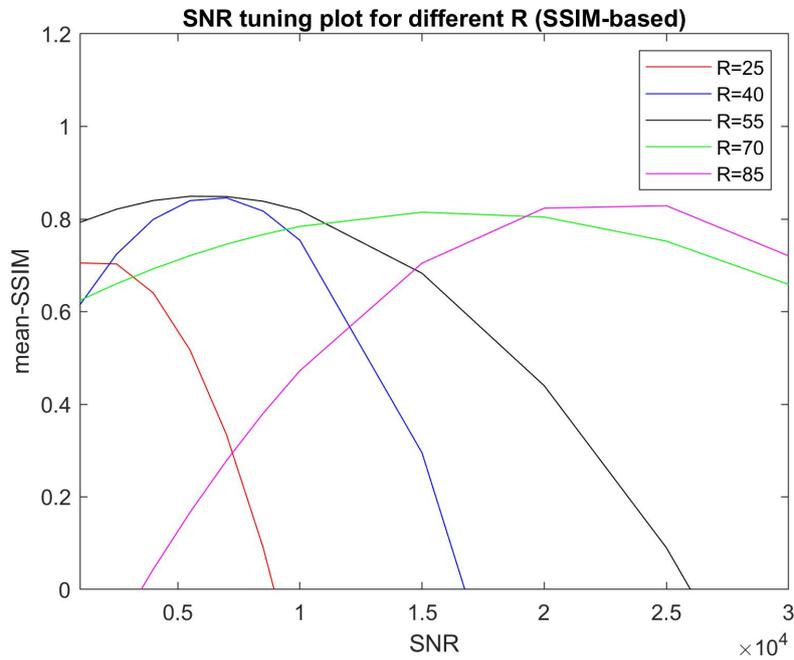


Figure 38: Mean-SSIM values versus SNR computed for different values of R. The curves represent the best fit line.

To further fine tune the parameters, the FovVideoVDP [77] metric was utilized which has been developed based on the foveation that occurs in the human visual system (see Section 2.4.3 for details on the metric). The threshold used for FovVideoVDP was 0.90. For the fine-tuning, the effect was applied to the original image and not the blurred image. Example outputs are illustrated in Figure 39. Using this, the values of SNR corresponding to different values of R are highlighted in Table 11. It should be noted that these values are subjective to the type of application and virtual scene and it is recommended to fine tune the values for each application using the process explained.

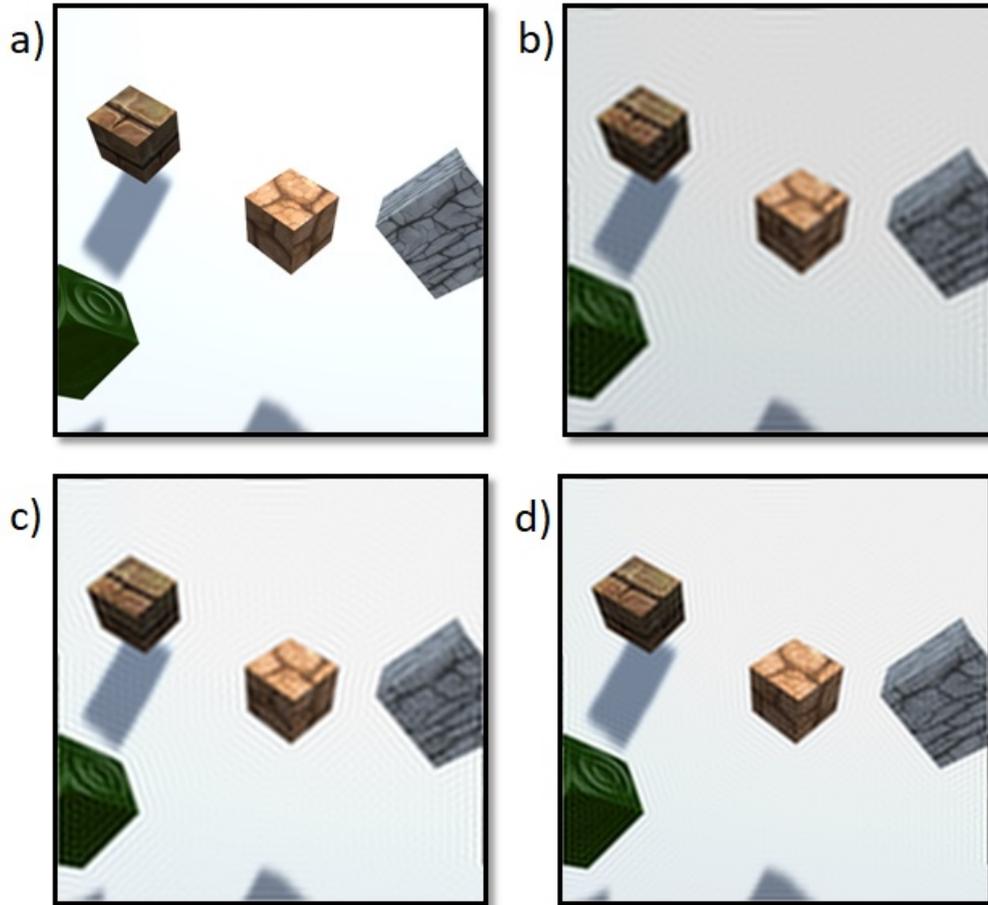


Figure 39: Fine tuning of the parameters. a) Source image. b) Output with parameters: $R=50$, $SNR=8000$. c) Output with parameters: $R=80$, $SNR=20000$. d) Output with parameters: $R=110$, $SNR=37000$.

| | | | | | | | | | |
|-------|------|------|-------|-------|-------|-------|-------|-------|-------|
| R | 40 | 50 | 60 | 70 | 80 | 90 | 100 | 110 | 120 |
| SNR | 5000 | 8000 | 11500 | 15500 | 20000 | 25500 | 31500 | 37000 | 45500 |

Table 11: Optimal values of SNR for different values of R .

4.3 Reaching Experiments

In order to understand whether inverse blurring effects can help reduce the effects of VAC, a study on depth perception was carried out. The task utilized was a reaching task in which the users were asked to reach a series of virtual positions using their finger.

4.3.1 Experimental Setup

The developed system was implemented using Unity⁴ operating on an Intel Core i7-9700K processor equipped with a NVIDIA GeForce 1080 graphics card. A 47-inch LG 3D screen⁵ which supports 1080p resolution at 60Hz frequency was used for interacting with the user. Users wore 3D polarized glasses to view the virtual objects in 3D. A Microsoft Kinect Sensor v2⁶ was used to measure the reaching positions.

The 3D screen was placed on an office desk and the user was seated 125cm away from the screen. The user viewpoint was set at the middle of the 3D screen. The Kinect was fixated below the screen (as shown in Figure 40).

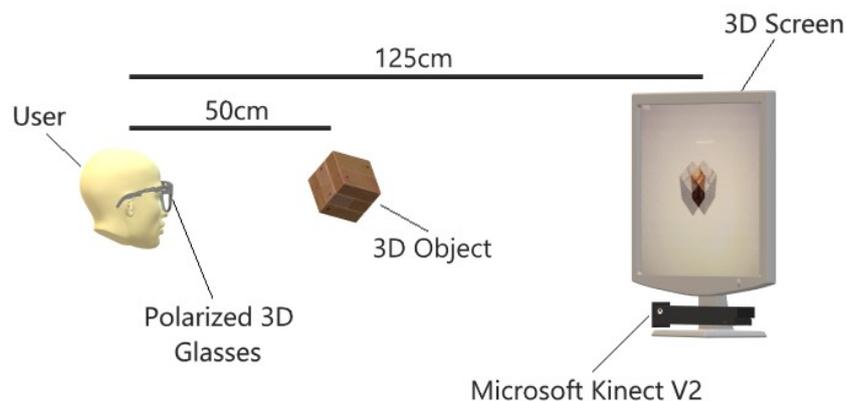


Figure 40: Reaching experiment overview. The user is seated at a distance of 125cm from the screen. The user's eye level is kept at the center of the screen. Virtual objects at different positions are visualized using polarized 3D glasses. Microsoft Kinect v2 tracks user positions.

A simplistic virtual 3D environment was created containing a spherical object of radius 1cm that spawned at different locations. The spherical object or ball acts as the target position that the user will have to try to reach. From literature, it is known that the average human adult's IPD is 63mm [31, 82, 127].

⁴ <https://unity3d.com/get-unity/download/archive>

⁵ <https://www.lg.com/it/supporto/prodotto/lg-47LM615S.API>

⁶ <https://developer.microsoft.com/en-us/windows/kinect/>

However, many studies have shown that the average IPD for females is lower than males and age also affects the IPD [74, 92]. For this purpose two settings for the IPD were chosen (63mm and 58mm).

For tracking the finger, the Kinect v2 Software Development Kit (SDK) for Unity⁷ was used. The Kinect SDK offers tracking of 25 body joints covering the whole body (see Figure 41). For the purpose of the experiment, only two joints were tracked, namely the Head and the HandTipRight joints. The Head joint acted as a reference to ensure the user stayed at the defined viewing position. The HandTipRight joint represented the position of the right hand index finger.

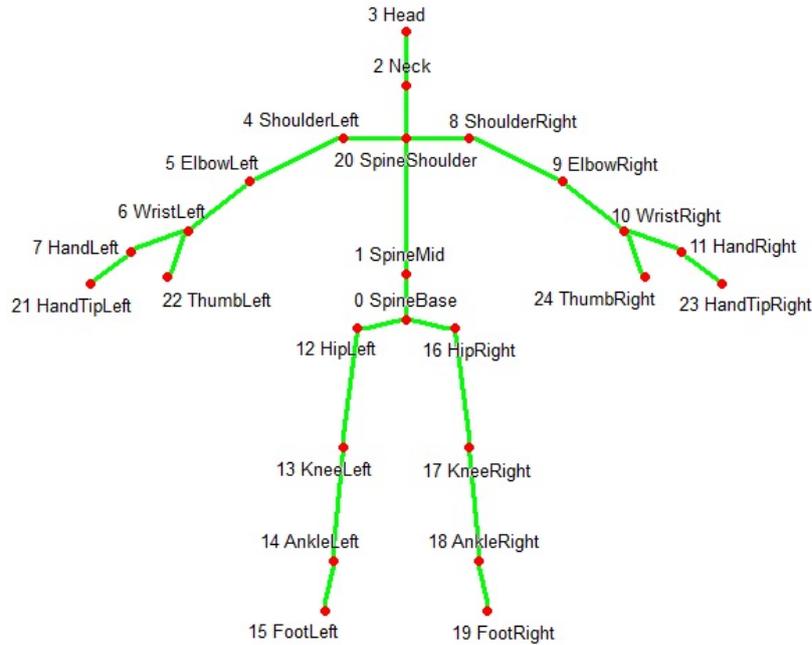


Figure 41: Kinect v2 joint hierarchy. A total of 25 body joints can be tracked.

From literature, the position of the Kinect v2 reference frame is known as it is at the center of the depth camera [138] (see Figure 42). Using a transformation, it is possible to convert object positions from Kinect reference frame to Unity reference frame and vice versa. The transformation was done using the following equations:

$${}^u_k T = \begin{bmatrix} R & | & t \end{bmatrix} \quad (12)$$

where k and u represent Kinect and Unity respectively. R and t are the rotation matrix and translation respectively as defined by:

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}, \quad t = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (13)$$

⁷ <https://go.microsoft.com/fwlink/p/?LinkId=513177>



Figure 42: Location of the Kinect v2 reference frame. Values are in mm.

To perform a linear transformation, a simple matrix multiplication operation can be performed to find the point ${}_k p$ expressed in the Kinect reference frame to the same point ${}_u p$ expressed in Unity reference frame using:

$${}_u p = {}_k^u T {}_k p \quad (14)$$

To ensure the perceived distances corresponded to actual distances, a pre-experiment session was conducted in which three people took part. A cube of size 5x5x5 cm was created in the virtual scene and a similar one was created in the real world (see Figure 43). The virtual cube was shown to the users at different depths and the users were asked to place the real cube where they see the virtual cube. Using a tape measure, the position of the real cube was measured and compared with the cube position in the virtual scene computed using Equation 14.

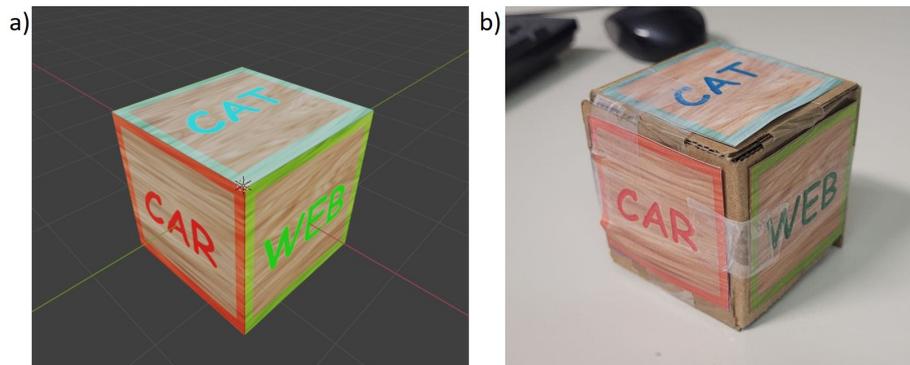


Figure 43: Distance verification. a) Virtual cube. b) Real cube.

4.3.2 Procedure

Data was collected from 23 users (13 males and 10 females) aged from 23 to 54 years (mean 30.65 ± 7.15) who were recruited from students and faculty members of the University of Genoa. All participants were volunteers and received no reward. All users had normal to corrected-to-normal acuity. Users who normally wore corrective glasses or lenses wore them underneath the polarized glasses.

The target positions were vertices of a 20x20x20 cm cube located 5cm apart. Therefore, the total possible positions were 125. In each session, there were 50

trials. The sequence of the target positions was randomly generated without repetition.

The user was asked to reach the position of the ball with their right hand index finger. Once, they felt that they have reached the target position, they were asked to hold steady their finger and press a button on the keyboard with their left hand to register the position.

The users were asked to undergo the sessions using the normal IPD setting. If the user reported issues fusing the stimuli, the session was stopped and the setting was changed to the lower IPD setting.

Two conditions were considered: normal viewing and inverse blurring viewing. In the normal viewing, the stimuli was presented in full fidelity. This session acted as the control group to have a reference performance. The stimuli in the inverse blurring session was presented with our effects enabled. The parameters used were based on the tuning process explained in Section 4.2.2 and optimal values found in Table 11. All users underwent the experimental conditions in random order, i.e., half performed the normal session first and half performed the inverse blurring session first. This was done to ensure no bias was present in the system.

For quantitative analysis, the finger positions were used. In order to also have a qualitative or subjective measure, a symptom questionnaire was used. The questionnaire used was developed by Hoffman et al. for their study on Vergence Accommodation Conflict (VAC) [44] and later adapted by Shibata et al. for assessing discomfort in stereo display applications [111]. The questionnaire asked the users to rate their symptoms on a 5-point Likert scale, where 1 indicated no symptom and 5 indicated severe symptom. The questions were:

- Q1) How tired are your eyes?
- Q2) How clear is your vision?
- Q3) How tired and sore are your neck and back?
- Q4) How do your eyes feel?
- Q5) How does your head feel?

The users filled the symptom questionnaire after each of the two sessions. When both sessions were completed, the user was asked to fill a session comparison questionnaire which was also adapted from the work of Shibata et al. [111]. In this questionnaire, the users were asked to rate their experience on a 5-point Likert scale. A lower rating meant that the users preferred the first session and higher rating indicated the preference towards the second session. The questions asked were:

- Q1) Which session was more fatiguing?
- Q2) Which session irritated your eyes the most?
- Q3) Which session gave you more headache?
- Q4) Which session did you prefer?

4.3.3 Data Analysis and Results

Data from 5 users was discarded. Three of these had a very high mean error (>25cm) and two had all their reaching positions in the same depth plane. This indicated that these users were not able to fuse the stimuli properly.

The error between the expected finger position and the perceived finger position was calculated. The mean errors along with their standard deviation are reported in Table 12. It can be seen that there is a small difference between the performance in the X (horizontal) and Y (vertical) planes. However, there is an improvement of around 2.23 cm in the Z (depth) plane. Error in the euclidean space was also calculated and a decrease in the error can be noticed.

| ERROR | NORMAL VIEWING | INVERSE BLURRING |
|--------------------|-----------------|------------------|
| X | 3.08 ± 1.44 | 3.05 ± 1.35 |
| Y | 3.78 ± 3.34 | 2.76 ± 2.20 |
| Z | 7.31 ± 3.43 | 5.08 ± 2.59 |
| Euclidean distance | 9.83 ± 4.21 | 7.50 ± 3.05 |

Table 12: Mean average absolute errors. All values in cm.

In order to understand, how the behaviour is in each of the 5 depth planes. The mean error for each of the planes was plotted along with the best fit line as shown in Figure 44. The distances are measured from the screen so the higher depth value indicates the ball is closer to the user. It can be seen that the errors increases when the object is in the near FoV.

Similar analysis was also done in the horizontal and vertical planes. The error plots are shown in Figures 45 and 46 respectively. In the horizontal plane, value 0 indicates the position at the center of the display which is also the center of the user-view. The error is higher when the target position was towards the left of the user. A potential reason for this could be that the user was asked to reach the locations with their right hand so the relative distance is higher. In the vertical plane, a similar trend can be seen. As the distance from the user increases, the error also increases.

A similar plot was also plotted for the euclidean distance (see Figure 47). The error grows as the distance increases, however, the rate of growth is more gradual in the case of the inverse blurring condition, as was the case in the depth plane.

In order to observe whether the error differences have a statistical significance, a T-test was performed. The p-values obtained are reported in Table 13. There is a significant difference in the depth plane and the euclidean space. However, no statistically significant difference was observed in the horizontal and vertical planes.

To understand whether the users were underestimating the positions or overestimating them, the position in the individual planes (XY, XZ and YZ) were observed. All the target positions were translated to a reference point

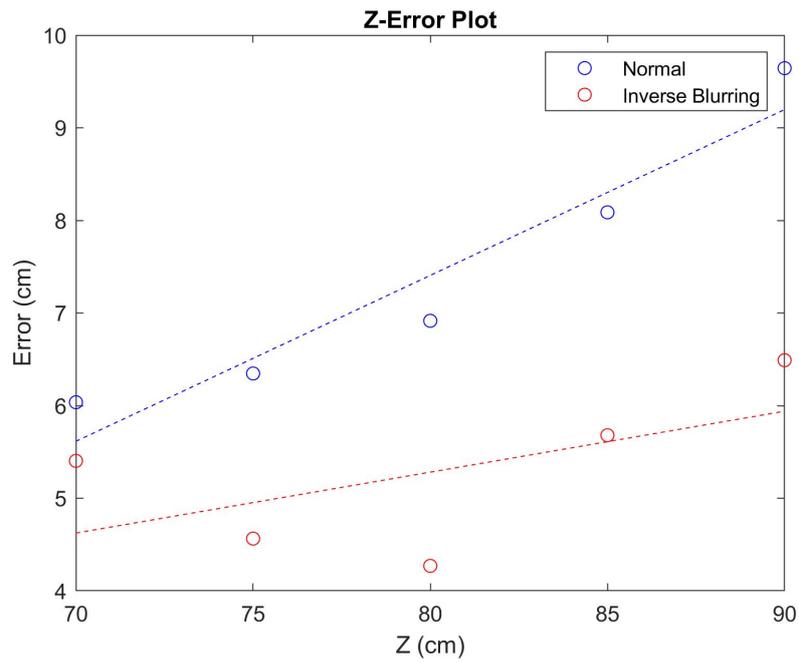


Figure 44: Z-error plot for the five depth levels. The dashed line represents the best fit line.

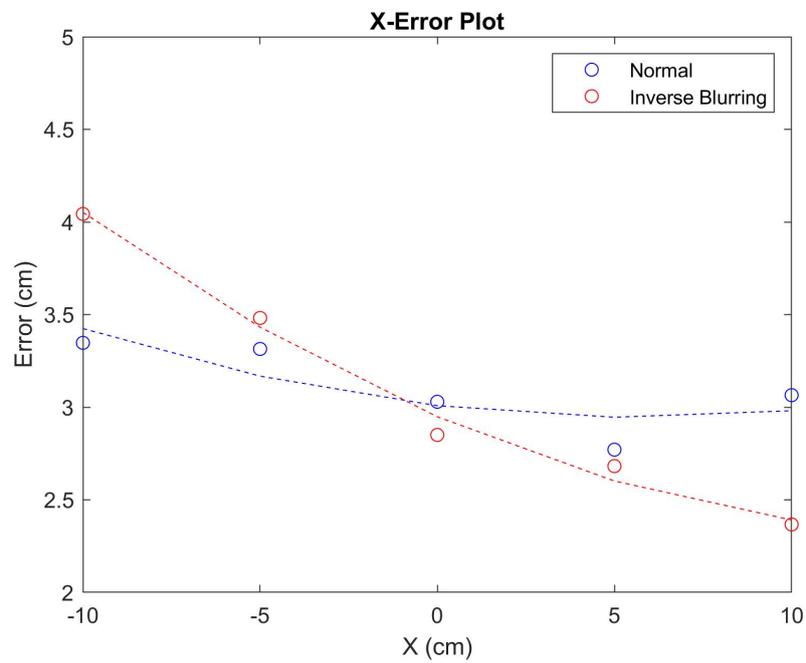


Figure 45: X-error plot for the five levels. The dashed line represents the best fit line.

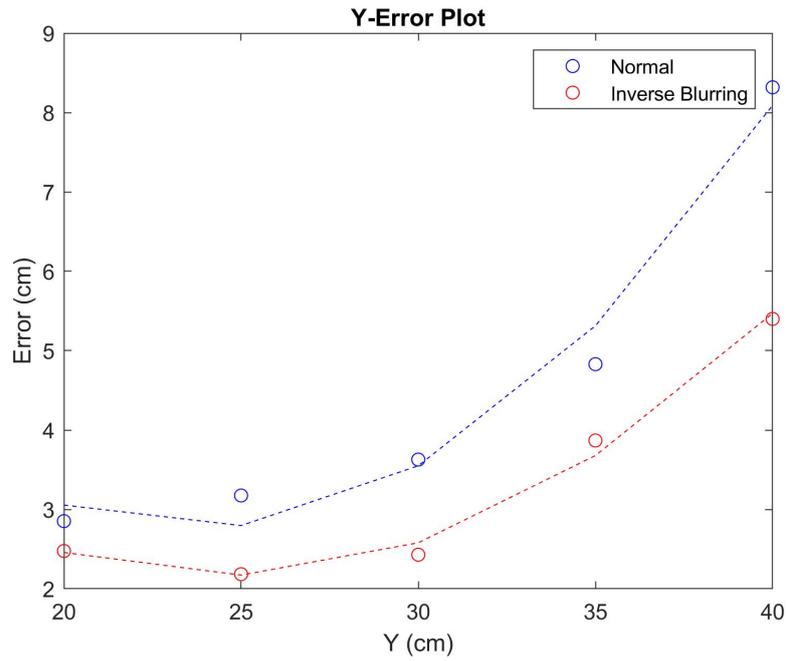


Figure 46: Y-error plot for the five levels. The dashed line represents the best fit line.

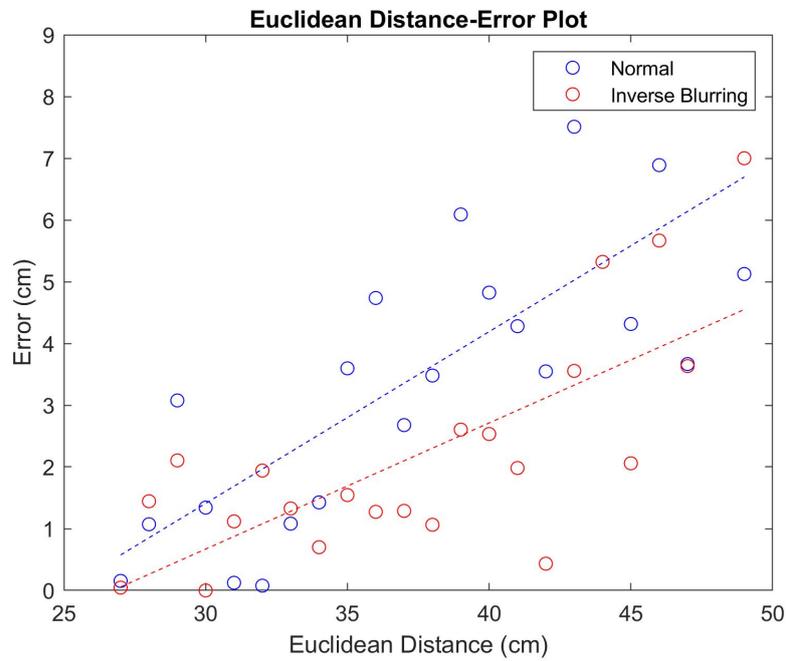


Figure 47: Euclidean distance-error plot. The dashed line represents the best fit line.

| T-TEST DATA | <i>p</i> -VALUE |
|--------------------|-----------------|
| X | 0.95 |
| Y | 0.30 |
| Z | 0.02 |
| Euclidean distance | 0.04 |

Table 13: T-test results.

(0,0,0). A heatmap was drawn of the three planes. Figures 48 and 49 show the heatmap for the two conditions in the XY plane. A similar pattern can be observed for the two conditions. The spread is symmetric, i.e., there is equal amount of overestimation and underestimation. There is no significant difference in the performance of the two experimental conditions.

Figures 50 and 51 show the heatmap for the two conditions in the XZ plane. It can be seen that there is slight underestimation in the normal condition and the points are more spread out. However, when observing the inverse blurring condition there is more symmetry to the points and the points are more closer to the reference position.

Figures 52 and 53 show the heatmap for the two conditions in the YZ plane. Again, a similar pattern is observed. There is an underestimation for the normal condition and the positions are more spread out. The error is reduced in the inverse blurring condition.

The time taken to perform the task was also computed (see Figure 14). The users took similar times for each condition indicating that there was no influence on how much time they spent to reach the position and how long it took them to fuse the stereoscopic stimuli.

| SESSION | TIME (s) |
|------------------|--------------|
| Normal | 1.77 ± 2.07 |
| Inverse Blurring | 1.68s ± 1.84 |

Table 14: Mean time taken to perform the reaching task.

Next, the subjective measure were analyzed. Figures 54 and 55 show the group mean values along with the standard deviations for the symptom questionnaire. The values are slightly lower for the inverse blurring condition, however, there is no significant difference between the two conditions due to the large variance. More pronounced symptoms can be observed for Q3. This is due to the physical load of the task as many users reported that their arms were tired after the sessions. The results for the session comparison questionnaire are shown in Figure 56. There is slight preference for the session with the inverse blurring effect.

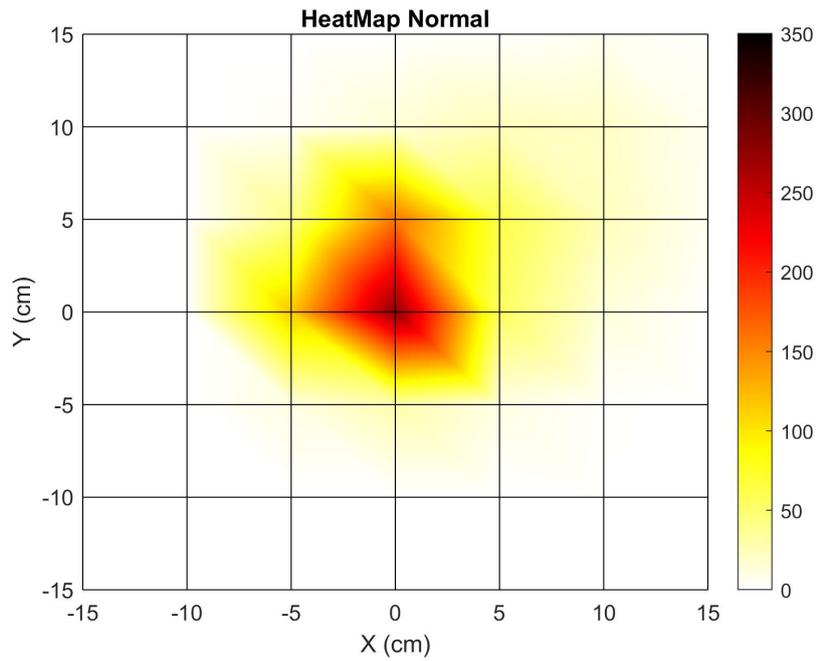


Figure 48: Heatmap for the finger positions with respect to a fixed position $(0,0,0)$ in the XY plane for the normal viewing condition.

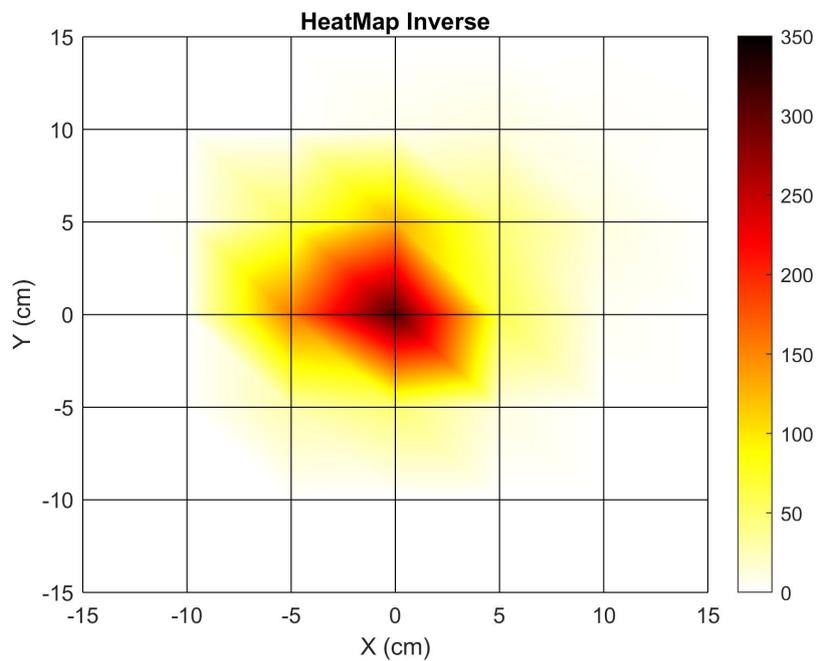


Figure 49: Heatmap for the finger positions with respect to a fixed position $(0,0,0)$ in the XY plane for the inverse blurring condition.

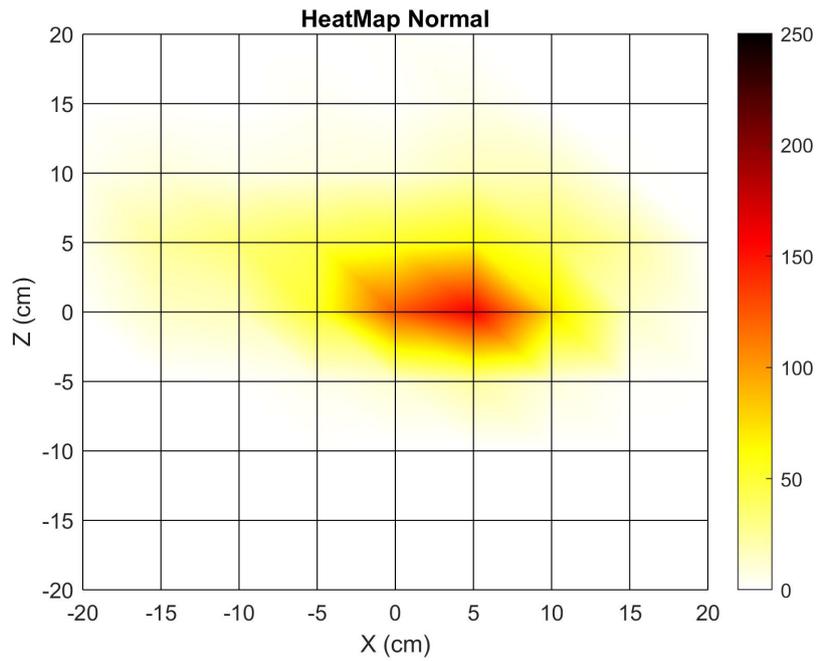


Figure 50: Heatmap for the finger positions with respect to a fixed position $(0,0,0)$ in the XZ plane for the normal viewing condition.

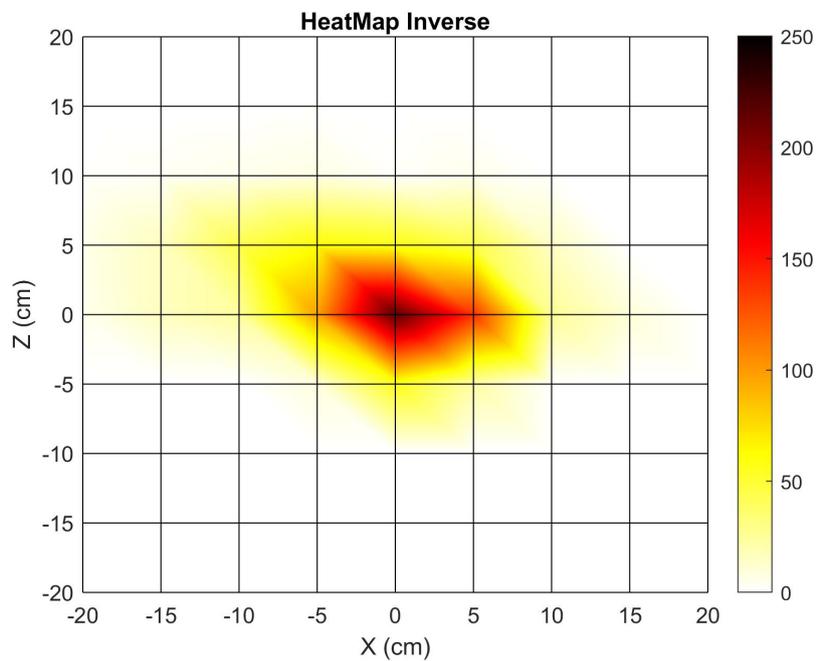


Figure 51: Heatmap for the finger positions with respect to a fixed position $(0,0,0)$ in the XZ plane for the inverse blurring condition.

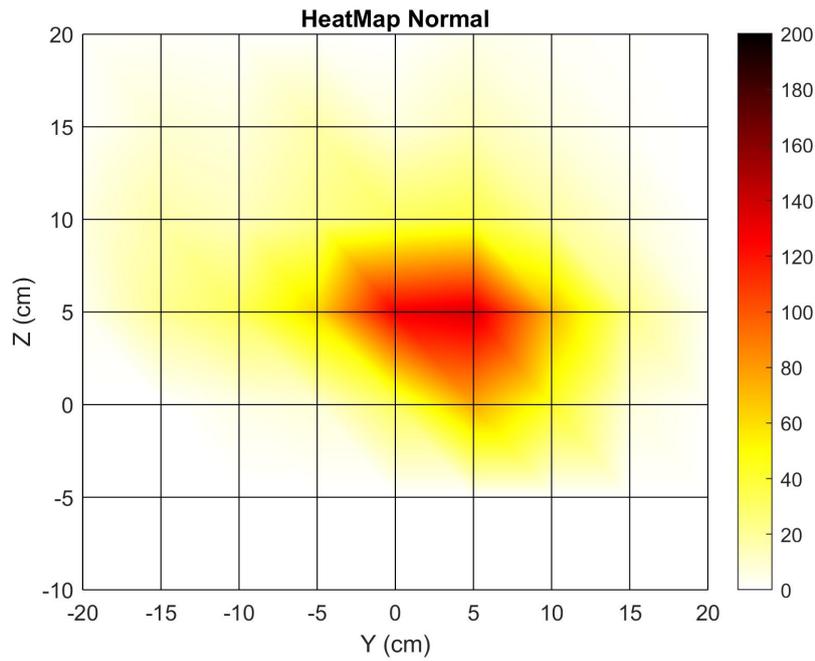


Figure 52: Heatmap for the finger positions with respect to a fixed position $(0,0,0)$ in the YZ plane for the normal viewing condition.

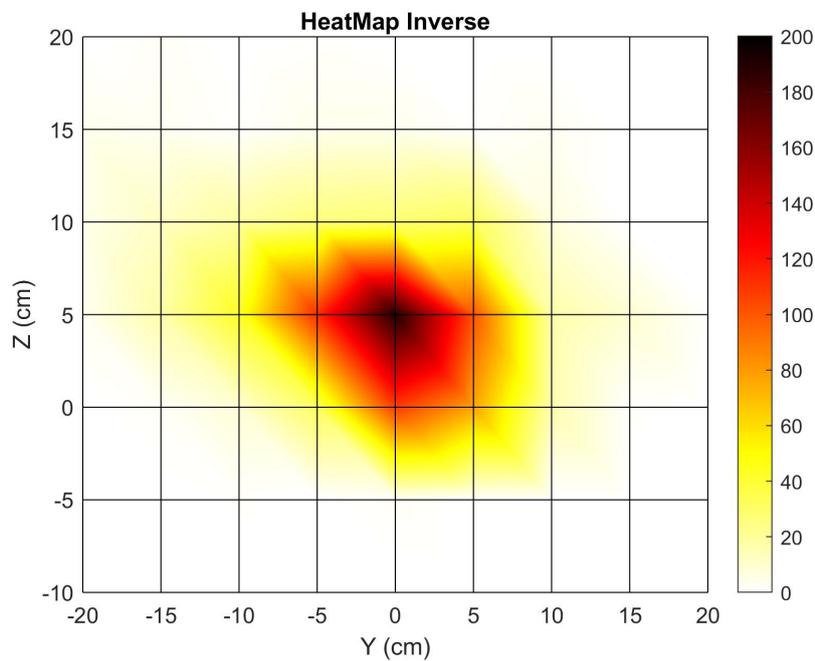


Figure 53: Heatmap for the finger positions with respect to a fixed position $(0,0,0)$ in the YZ plane for the inverse blurring condition.

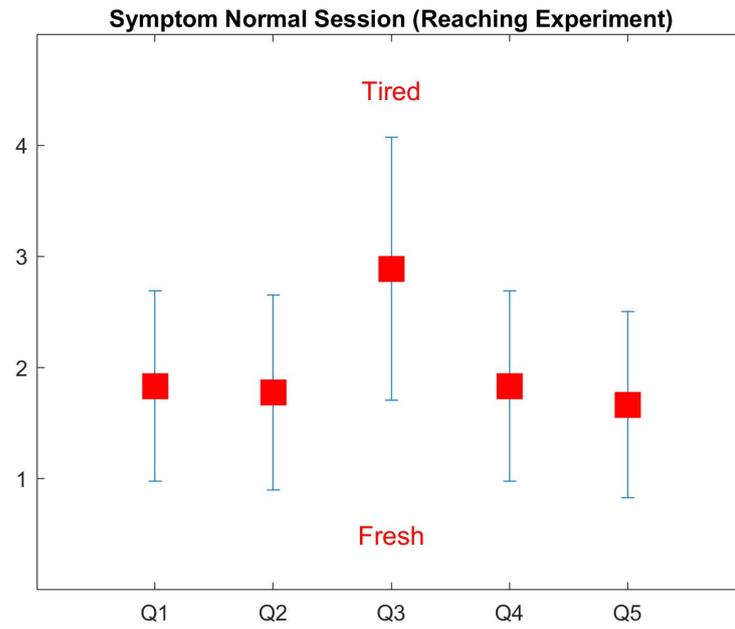


Figure 54: Symptom questionnaire scores for the normal viewing condition. Q1: 1.83; Q2: 1.78; Q3: 2.89; Q4: 1.83; Q5: 1.67.

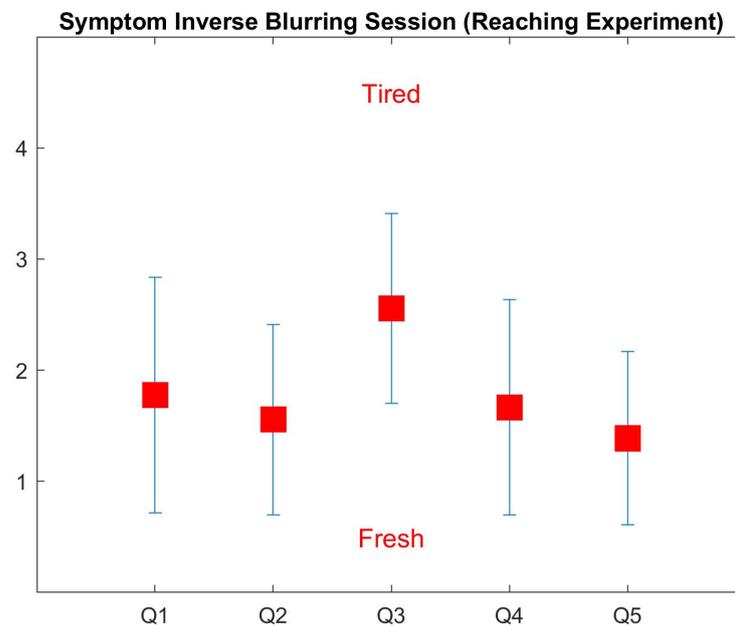


Figure 55: Symptom questionnaire scores for the inverse blurring condition. Q1: 1.78; Q2: 1.56; Q3: 2.56; Q4: 1.67; Q5: 1.39.

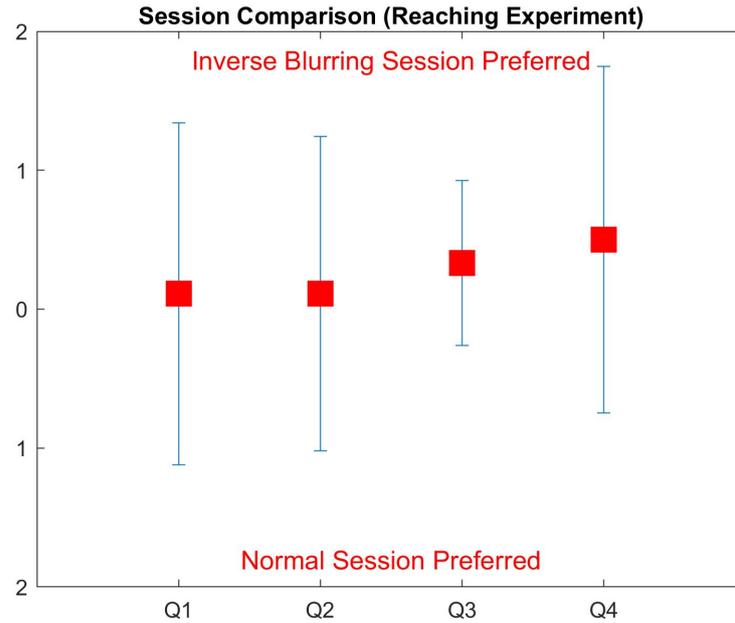


Figure 56: Session comparison questionnaire scores for the reaching task. Q1: 0.11; Q2: 0.11; Q3: 0.33; Q4: 0.50.

4.3.4 Outcome

The aim of this work was to develop a system for immersive media devices that incorporates inverse blurring distortions with the aim of mitigating VAC. For this purpose, the Wiener deconvolution was used as the deblurring technique. It was chosen because unlike other superior deblurring algorithms, the computational load is less and it is insensitive to small variations in the signal power spectrum. Other approaches either use an iterative procedure which requires a high processing time resulting in an undesired low frame rate or they are based on deep learning models which have high memory requirements.

The purpose of this user study was to understand how distortions to visual stimuli created by inverse blurring can affect depth perception. The study was based on a reaching task where users were asked to reach different positions in the personal space with their right hand index finger. Experimental analysis showed that an improvement of 36% was achieved in depth perception. An interesting observation was that users were underestimating the distances which supports various studies found in literature. However, the underestimation was lower with the inverse blurring sessions as compared to normal viewing.

Some users reported issues fusing the visual stimuli, so an additional user study was carried out. The new user study uses objects that are bigger and textured to help with stimuli fusion. The study was also aimed as support for the improvement in depth perception observed during the reaching experiment.

4.4 Spatial Awareness Experiments

Some of the people reported issues with fusing the stimuli in the reaching experiment. An alternate experiment was conducted to verify the improvement in the performance with the inverse blurring condition.

4.4.1 *Experimental Setup*

The developed system was implemented using Unity⁸ operating on an Intel Core i7-9700K processor equipped with a NVIDIA GeForce 1080 graphics card. A 47-inch LG 3D screen⁹ which supports 1080p resolution at 60Hz frequency was used for interacting with the user. The user wore 3D polarized glasses to view the virtual objects in 3D. An on-screen Eye Tribe¹⁰ tracker was used to track user's eye movements. The eye tracker has an accuracy of 0.5°–1° and an operating range of 45–75cm.

A similar VE to the reaching experiment was created. Two virtual textured cubes of size 10x10x10 cm were placed equally distant from the center of the screen (one towards the left and the other towards the right). The distance between the cube was 60cm in the horizontal plane and 0 cm in the vertical plane. Ten depth levels were created with 5 cm intervals. A plus sign was placed at the center of the screen.

4.4.2 *Procedure*

Data was collected from 24 users (14 males and 10 females) aged from 23 to 54 years (mean 30.65 ± 7.00) who were recruited from students and faculty members of the University of Genoa. All participants were volunteers and received no reward. All users had normal to corrected-to-normal acuity. Users who normally wore corrective glasses or lenses wore them underneath the polarized glasses.

The user was seated 80 cm from the 3D screen. The user was told to fixate on the cross. A stimuli containing the two cubes was shown (see Figure 57). The depth level of each cube was randomly selected. Each user session lasted for 50 trials. The stimuli was shown for 800 ms. This time was chosen based on studies found in literature which suggested that humans take around 500–800 ms to respond and fuse the stimuli depending on the distance [15, 42, 44, 51, 123]. When the stimuli disappeared, the users were asked to select which of the two cubes was closer to them. The users made the selection by pressing the arrow keys on a keyboard, i.e., left arrow key if they perceived the left cube as closer to them or the right arrow key if they perceived the right cube was closer. The choice was forced, i.e., even if they perceived the two cubes at the

8 <https://unity3d.com/get-unity/download/archive>

9 <https://www.lg.com/it/supporto/prodotto/lg-47LM615S.API>

10 <https://imotions.com/hardware/the-eye-tribe-tracker/>

same depth, they had to make a selection. This approach was based on the Two-Alternative Forced Choice (2AFC) paradigm.

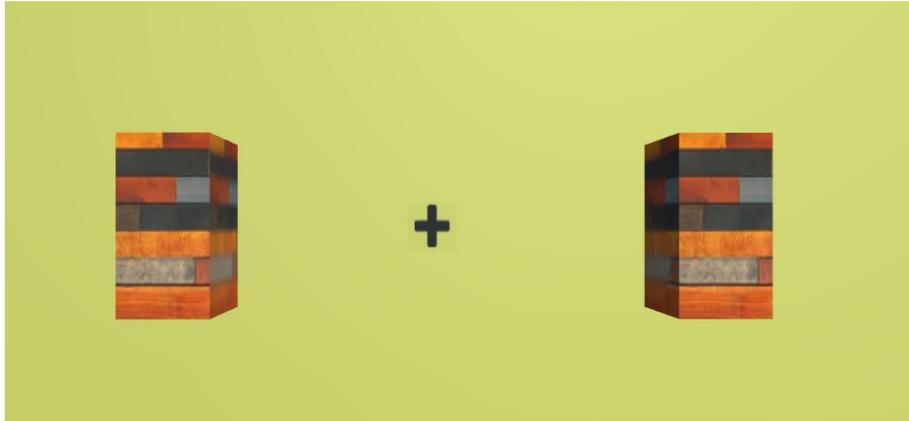


Figure 57: Spatial awareness experiment stimuli. The two cubes appear at different depth planes. The plus sign at the center of the screen is placed at the depth plane of the 3D screen.

Since the observed objects appear at different depth, users may use other depth cues such as relative size to estimate which object is closer. However, the purpose of the user study is to only study depth perception via accommodation and convergence. For this reason, the objects were scaled in such a way that they occupied the same number of pixels on the screen irrespective of their depth. This ensured that users only used accommodation and convergence to make their selection.

In each trial, before showing the stimuli, the users were asked to fixate on the plus sign. They were given 500ms to do this. This was done to ensure that the starting gaze condition is similar for all trials and also to give the users some time to focus back on the screen in case they looked on the keyboard to make the selection.

Two conditions were considered during the experiment: normal viewing and inverse blurring viewing. In the normal viewing, the stimuli was presented in full fidelity. This session acted as the control group to have a reference performance. The stimuli in the inverse blurring session was presented with our effects enabled. The parameters used were based on the tuning process explained in Section 4.2.2 and optimal values found in Table 11. All users underwent the experimental conditions in random order, i.e., half performed the normal session first and half performed the inverse blurring session first. This was done to ensure no bias was present in the system. Users had to fill a post session questionnaire similar to the one used in the reaching experiment.

4.4.3 *Data Analysis and Results*

The probability of getting the correct answer is 50% when choosing randomly. For this reason a threshold was set. Four users had more than 50% error rate

for both conditions indicating either they did not fully understand the task or were guessing randomly. So their data was discarded from the analysis.

The number of correct and incorrect answers for all users were computed. The individual performances are shown in Figures 58 and 59. The group means along with their standard deviations are summarized in Table 15. It can be observed that the error is much lower in the inverse blurring condition, indicating that the effect lowers the conflict caused by accommodation and convergence in stereo displays. It should be noted that in some trials, the two cubes appeared at the same depths. Those trials were considered neither as correct or incorrect.

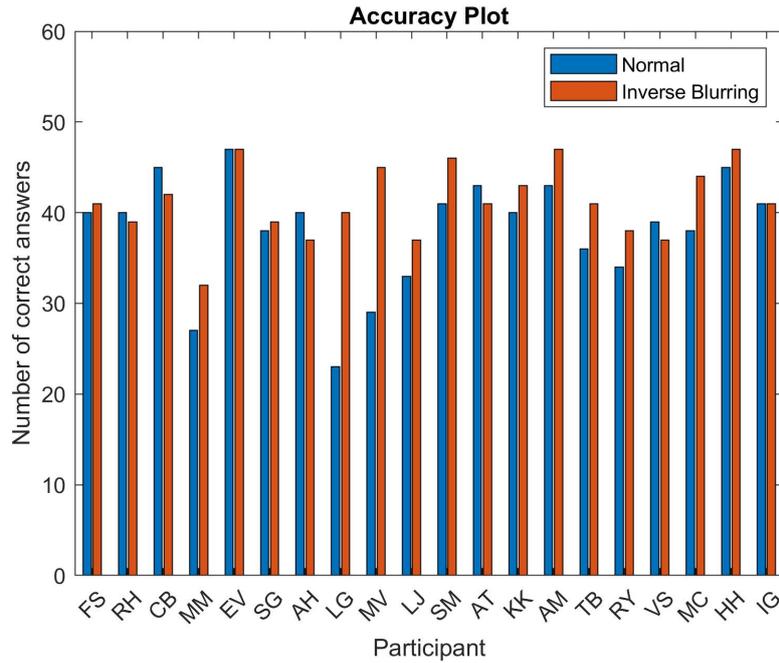


Figure 58: Individual user performance for the spatial awareness task. The bars indicate the number of correct answers for the two experimental conditions.

| | NORMAL | INVERSE BLURRING |
|-----------|----------------|------------------|
| Correct | 38.1 ± 6.2 | 41.2 ± 4.0 |
| Incorrect | 11.3 ± 5.7 | 6.9 ± 3.7 |

Table 15: Group mean performance.

To understand whether the results have statistical significance, the discrimination sensitivity can be computed for the 2AFC task as was done by Maiello et al. in their study [76]. The data for each user and condition was converted into discrimination d' [133]. A bootstrapping procedure was used to compute the group confidence levels on d' measurements [28]. Mean d' were computed for each user and condition from the original data re-sampled with replacement 5000 times. These bootstrapped distributions were then collapsed across

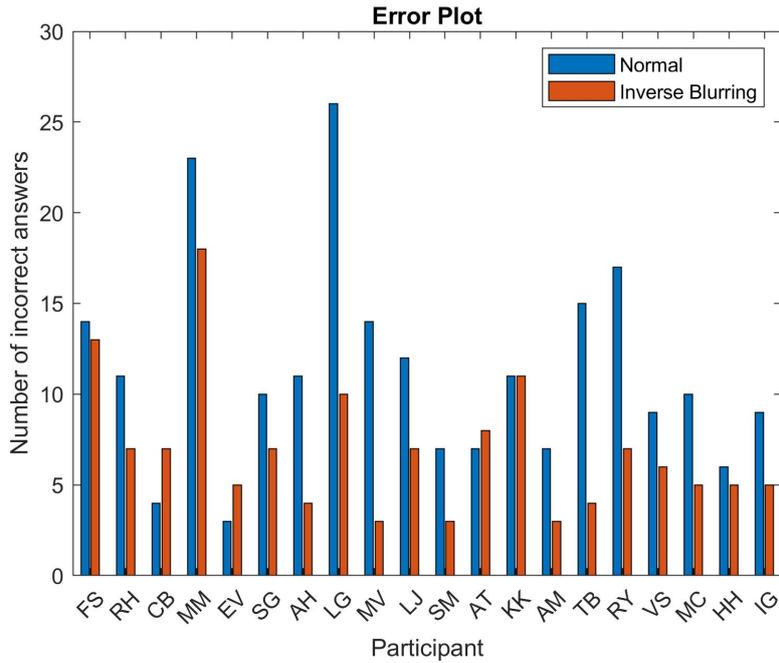


Figure 59: Individual errors in user performance for the spatial awareness task. The bars indicate the number of incorrect answers for the two experimental conditions.

observers to obtain group distributions for each condition. The group distributions were fitted over a Gaussian distribution from which the 2.5th and 97.5th quantiles were taken as the 95% CI.

Figure 60 shows the discrimination for the two experimental conditions. A mean discrimination of 1.46 was observed for the normal viewing session whereas the discrimination increased to 2.02 when the inverse blurring condition was presented. The increase is statistically significant. The results are summarized in Table 16.

| | NORMAL | INVERSE BLURRING |
|--------|--------------|------------------|
| Mean | 1.46 | 2.02 |
| 95% CI | [1.20, 1.93] | [1.85, 2.40] |

Table 16: Discrimination sensitivity for the two experimental conditions.

The time it took the users to make the selection was also computed. This time is computed after the 1.3s it took to display the stimuli had passed. The mean time taken along with the standard deviation are reported in Table 17. In both conditions, user took similar times to make their selections.

Next, the subjective measures were analyzed. Figures 61 and 62 show the mean values along with the standard deviations for the symptom questionnaire. The values are slightly lower for the inverse blurring condition, however, there is no significant difference between the two conditions. In this experi-

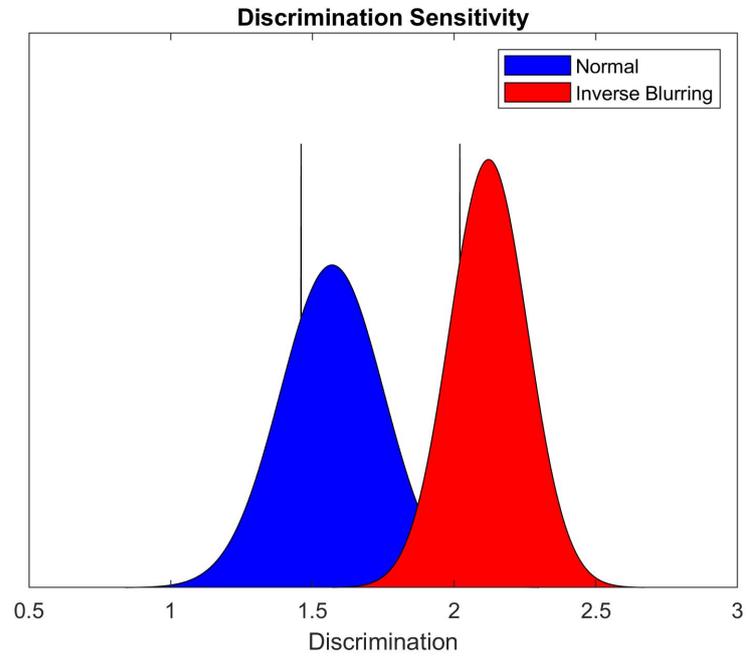


Figure 60: Discrimination sensitivity plot. The vertical bars are the true group means.

| SESSION | TIME (s) |
|------------------|------------------|
| Normal | 0.65 ± 1.18 |
| Inverse Blurring | $0.43s \pm 0.58$ |

Table 17: Mean time taken to perform the spatial awareness task.

ment, there was no longer more pronounced symptoms for Q3 as the physical load of the task was less. The results for the session comparison questionnaire are shown in Figure 63. There is slight preference for the session with the inverse blurring effect.

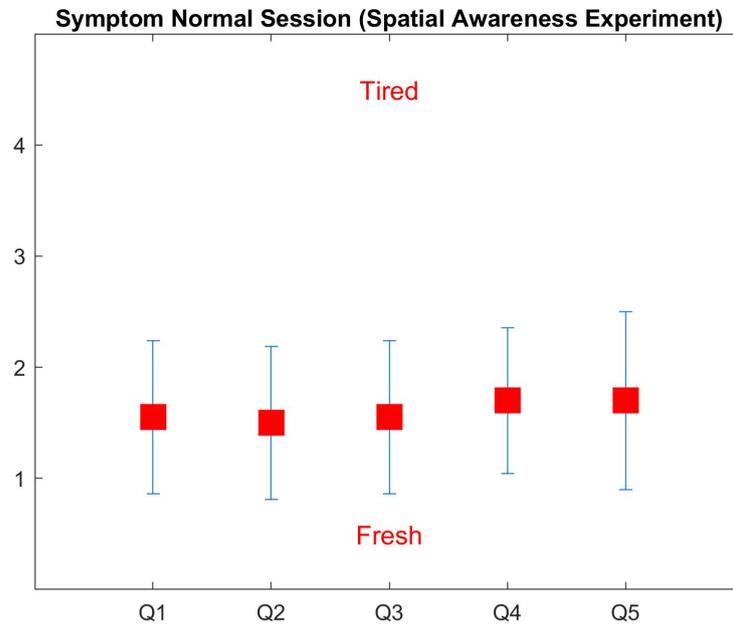


Figure 61: Symptom questionnaire scores for the normal viewing condition. Q1: 1.55; Q2: 1.50; Q3: 1.55; Q4: 1.70; Q5: 1.70.

During the experimental sessions, the frame processing times were also computed in order to have an insight about the processing cost required by the inverse blurring technique. The algorithm took on average 16ms to process each frame, resulting in a 62Hz frame rate. This shows that the processing cost of the technique is not high and it does not affect the standard frame rate of XR devices. However, it should be noted that the frame processing time is dependent on the screen size and the processing power available.

4.4.4 Outcome

This user study aimed at further supporting the observations of the reaching experiment. The task utilized in this study was a spatial awareness one, where the users were asked to indicate which of the two objects appeared closer to them. The sizes of the objects were adjusted based on their depths to make sure that they occupied the same amount of pixels on the display. This ensured that the user only exploited convergence and accommodation to estimate the object distances. Other depth cues such as difference in object sizes were made redundant in the experimental setup.

The data analysis supported the conclusions of the reaching experiment. An improvement of 48% in depth perception was observed. After undergoing the

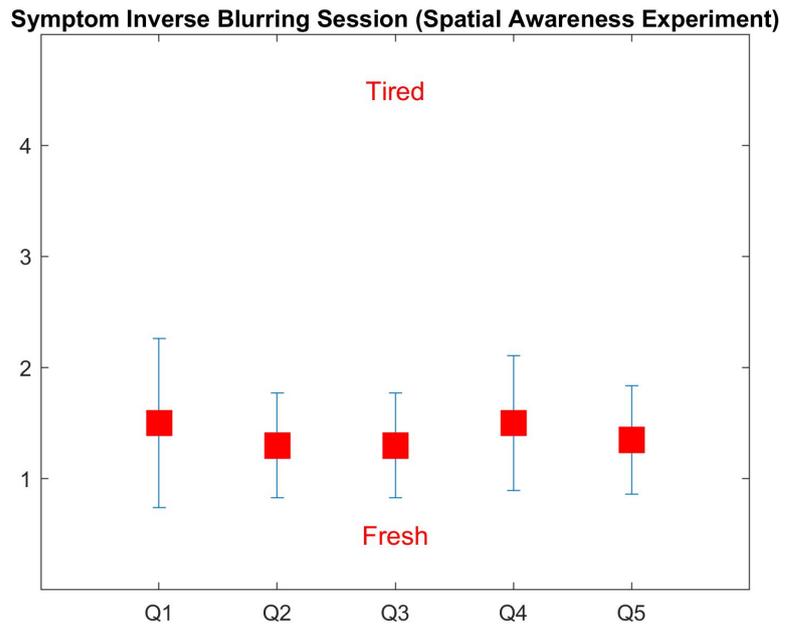


Figure 62: Symptom questionnaire scores for the inverse blurring condition. Q1: 1.50; Q2: 1.30; Q3: 1.30; Q4: 1.50; Q5: 1.35.

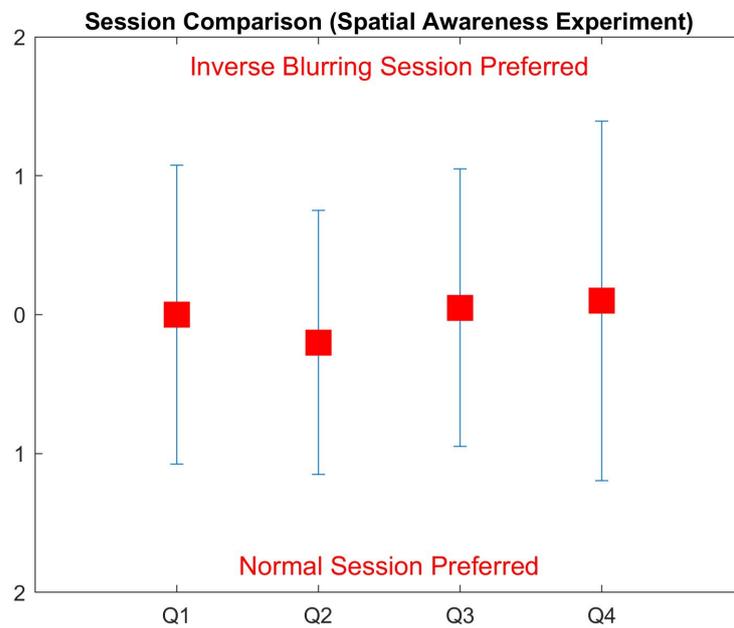


Figure 63: Session comparison questionnaire scores for the spatial awareness task. Q1: 0.00; Q2: -0.20; Q3: 0.05; Q4: 0.10.

two experimental sessions, many users highlighted that they found discriminating the object depths better with the inverse blurring system. Although the majority of the users performed the task better with the inverse blurring system, the qualitative measures produced a mixed conclusion. From a subjective viewpoint, many users preferred the normal viewing condition. A potential explanation for this could be that humans are used to the aesthetics of normal viewing and any distortions in the scene such as those introduced by the deblurring system are often considered as artifacts.

In this experimental setup, no user reported difficulties in fusing the stereoscopic stimuli. This indicates that the spatial limitations of the hardware need to be considered while designing the application. Also, textured objects allow users to fuse the stimuli better so their use in stereoscopic displays is highly encouraged.

PART IV

Discussion

The following part provides a summary of the research work done and a perspective on some open questions and challenges for future developments to conclude the thesis.

5

Conclusions

The field of Virtual Reality (VR) has had a significant boost in the recent decade. In the past, due to the bulky nature of VR devices and their high costs, its applications were more or less confined to military training and flight simulators. However, with the advent of compact commercial HMDs, their application domain has expanded. These devices are lightweight and are much more affordable. These days, the applications of VR technology can be found in gaming, medical training and minimally invasive surgeries, data visualization, remote/virtual meetings, edutainment, physical and cognitive rehabilitation, product design, search and rescue mission training and infotainment. However, the current technology, although quite advanced, is not perfect in bringing the human experience of the real world into the virtual one. A lot of factors contribute to this such as sensory conflict, restricted Field-of-View (FoV), lower resolution, lack of proper haptic feedback, embodiment, etc. The focus of this thesis was mainly on addressing the sensory conflict and the subsequent effects caused by it.

The contribution of this thesis is in two domains. Firstly, a framework was developed to incorporate spatial blurring into VR applications with the aim of mitigating the onset of cybersickness. Second is the development of a system, addressing the Vergence Accommodation Conflict (VAC) in immersive media with the aim of improving depth perception.

Although, the literature is quite vast for the first contribution, the current work focuses on a hybrid approach. Broadly speaking, in the literature, two main approaches are highlighted. They either suggest the use of Depth-of-Field (DoF) effects or the use of foveated rendering. DoF is based on the properties of lens and occurs commonly in cameras. Depending on the distance between objects in the environment, a certain amount of blur is present in the peripheral regions. On the other hand, foveated rendering techniques are designed to limit the spatial resolution of the peripheral regions in order to cater for the growing processing and memory demand of the XR industry. The literature deals with the two approaches separately. Foveated rendering is not studied as an approach to help cybersickness whereas DoF is proposed as a measure to remove the visual artifacts introduced by screen effects. Very few attempts have been made to combine these effects, however, the corresponding

literature makes use of advanced technologies such as ray tracing which are not readily available in most current XR systems. On the other hand, the developed technique uses a hybrid approach with the two approaches embedded into the screen effect. The technique is aimed at reducing the onset of cybersickness and can be applied as a post-processing effect to any XR application to achieve a more natural virtual environment.

The developed technique uses the Bokeh filter as the smoothing filter. The algorithm is implemented in four steps using a four-pass shader. The first pass uses the gaze location to compute the radial distances of each pixel from the focus point and also computes the Circle of Confusion (CoC) values. The second and third shader passes apply the DoF and foveation effects respectively to the source image. The last shader pass combines the two effects in a congruent way to produce an artifact-free scene. The effects were computed at half-resolution and later up-sampled back to the original resolution. The system offers sufficient real-time capabilities that are able to meet the frame rate requirements necessary to cope with fast eye movements. Since the system requires user gaze as input, it is essential to have an eye-tracking capable VR device.

A experimental study was carried out in order to understand whether such effects can actually reduce cybersickness. A custom VR roller-coaster environment was created. This was done so that we have control over the experimental conditions such as velocity, acceleration, time, etc. Three experimental conditions were evaluated. The first one is the full fidelity condition where no effects was applied to the VR scene. This condition was included to act as the control group to have a reference. The second experimental condition was the VR environment with the developed spatial blur enabled. The third was the blur technique present in the Unity post-processing stack. It should be noted that this technique only implements the DoF effect. 18 participants took part in this user study. The HMD used was the HTC Vive Pro Eye which has an integrated Tobii eye-tracking system. For qualitative analysis, Simulator Sickness Questionnaire (SSQ) and Igroup Presence Questionnaire (IPQ) were used. User heart rate and gaze measurements were used for quantitative analysis.

The analysis showed that there was a statistically significant reduction in the onset of cybersickness by incorporating the spatial blurring effects into the VR environment. More specifically, there was a 27% and 66% reduction in the sickness scores for the Unity blur and foveated DoF effects respectively when compared to the normal condition. This observation was also supported by the heart rate and gaze analysis. Temporal data suggested that circular or spiral motion tended to affect negatively to cybersickness when compared to linear motion. The study group was also divided based on gender and age. The analysis showed that older users were slightly more susceptible to cybersickness. However, no statistically significant difference was observed for the age groups when undergoing session with the foveated DoF effects enable. The gender groups produced no difference between the experimental conditions.

There are obvious differences between the scenes presented in the three conditions which may help understand why there is lower sickness induced in the

systems with spatial blur incorporated. The no blur condition presented the entire VR scene in high focus which contradicts natural viewing. The Unity blur condition mimics how lenses work while our technique considers depth-of-field and foveation effects together as in natural vision leading to a more realistic scene. Another possible explanation to why a reduced sickness is observed is optic flow. Motion in the periphery can cause sickness. Motion is detected by the visual system and hence the motion is seen, but no motion or little motion is sensed by the vestibular system. By reducing the amount of information in the peripheral region, the users are less susceptible to this sensory conflict.

The second contribution of the thesis deals with understanding depth perception in a Virtual Environment (VE). Although there is abundant literature for depth perception, it mainly focuses on personal space and relies on interaction modalities. On the other hand, in the current work, the emphasis is put on understanding how the scene can be visually altered to improve the depth perception. To this end, firstly, a user study was carried out in which the previously developed foveated DoF effects were used.

Users were asked to identify how many objects in a cluttered environment were at the same distance to a reference object. Two conditions were considered. The first condition was the VR environment with the foveated DoF effects enabled. The second was the normal viewing condition which acted as the control group. 12 participants took part in the study. User errors were used to quantify the user performance. The analysis showed a 27% reduction in error with the spatial blur enabled. User performance either improved or stayed the same, indicating that such effects do not have a detrimental effect on user performance. Most of the users were overestimating the number of objects at the same scene depth, i.e., a higher answer was given than the correct one. Generally, in both user studies, the users found the transition in the spatial blurring effects to be smooth and did not perceive any noticeable artifacts.

One of the most predominant reasons for poor depth perception in VR environments is the Vergence Accommodation Conflict (VAC). It causes the user to experience mismatching cues to how far the object actually is. To address this mismatch, the current work developed a system to introduce inverse blurring into VE. The purpose of the system is to emulate how visual perception works in humans. Humans when viewing objects in the real world, tend to have a natural blur integrated into their vision which is dictated by anatomical setup of the eye. Light rays from objects hitting the fovea in the retina form a sharp image while the object reduces in clarity as the light rays hit the retina elsewhere. The diffraction pattern of these light rays can be modelled as a PSF. If the PSF is known, the light rays can be adjusted to form a more natural viewing experience.

A Wiener deconvolution based system was developed. The system was implemented using OpenCV for Unity package. Firstly, a pre-processing step transforms the image into the frequency domain. The filter was computed based on two parameters, SNR and R which are the signal-to-noise ratio and radius of the circular PSF approximation. The filter is applied to the image

and the resulting image is transformed back into the spatial domain. During the pre-processing process, the image is squared and down-sampled to half-resolution. Once the effect has been applied, the image is up-scaled to its original resolution. This is done to boost processing times and lower memory requirements. The developed system can be incorporated into any immersive media device.

In order to verify the usefulness of the system, two user studies were conducted on depth perception. The first one involved a reaching task and was done in the personal space. While the second one comprised of a spatial awareness task based in the personal and action space. For both task, an LG 3D screen was used to interact with the user while the user wore 3D polarized glasses. Two experimental conditions were considered. The first was the one with the inverse blurring enabled while the second one was the normal viewing condition which acted as the control group.

For the reaching task, the users were shown a series of virtual target locations using a small spherical object. The user had to reach the target position with their right hand index finger. A Kinect v2 was used to measure the finger locations. The finger positions represented the perceived position. The distance between the actual position and the perceived position was considered as the error in the system and used for qualitative analysis. A symptom questionnaire was used to assess any symptoms that arose by using the system while a session comparison questionnaire was used for subjective measures. 23 users participated in the user study.

The analysis showed a statistically significant reduction in error with the inverse blurring condition. A 36% reduction in error was observed for the Z-axis (depth). No significant difference was found in the errors in the X and Y-axis (horizontal and vertical). In general, the users were underestimating the distances with the error growing as the target position was brought closer to the user. There was no significant difference between the subjective measure as both systems were equally preferred by the participants.

For the spatial awareness experiments, the participants were shown two virtual cubes, one on each side of the user. The cubes were strategically placed to appear at different depths. The users had to identify which cube was closer to them. The cubes were scaled in such a manner that they occupied the same number of pixels irrespective of which depth they were actually at. This was done to negate other depth cues such as size differences and occlusion and to force the users to only use convergence and accommodation to perceive the depth.

The analysis showed that there was a 48% reduction in error by incorporating inverse blurring into the system. After undergoing the experimental sessions, many users stated that it was easier to identify the closer object during the inverse blurring condition which is supported by the quantitative analysis. Similar to the reaching experiment, the subjective measures did not produce any significant preference between the two conditions. A potential reason could be that users are more used to the aesthetics of the normal viewing condition, prompting some users to subconsciously side with it.

6

Perspective

Virtual Reality (VR) is currently an evolving technology with new headsets having many new features showing up in the market each year. The prime focus has been to increase the visual quality of the device such as larger FoV and better resolution along with making the interaction as natural as possible by integrating new tools such as haptic gloves. With each new advance in the field, entrepreneurs and researchers are constantly being drawn to the technology. The introduction of eye tracking technology in VR devices has opened up a new era of how user behaviour can be interpreted. Eye tracking gives an insight into where the user's attention is at each moment of the virtual experience and what visual elements trigger each response and behaviour. Eye tracking can contribute to a more immersive user experience by enabling more natural interactions through gaze. With the passage of time, it is fair to say that the VR technology will be able to perfectly replicate or even potentially replace real-world experiences. To this aim, it is imperative that the newer technology is based on human physiology and psychology, so that appropriate models and paradigms can be created that form the basis of VR devices.

A key challenge while rendering to any HMD is maintaining low latency which is important for achieving a higher sense of immersion and reducing visual fatigue. With considerable increase in pixel densities found in latest HMDs, the computational load is ever-increasing. High quality rendering at low latencies is beyond the current capabilities of VR hardware and software solutions. Similarly, discrepancies exist between how humans perceive the virtual world and how they perceive the real world causing significant fatigue. In VR devices, the VE is rendered in pin-sharp focus, allowing them to extract information from all regions of the projected scene. This is contrary to how humans perceive in the real world, where humans focus on a certain object by altering their accommodation and convergence while the surrounding objects appear blurred. The work presented on spatial blurring and cybersickness offers a new paradigm on how visual fatigue can be reduced in modern consumer headsets. By altering the visual stimuli based on techniques inspired by the human physiological system, we can bridge the gap between the real world visual experience and its virtual counterpart.

The experiments conducted on cybersickness, offered an insight on how to design VR applications. For example, it was found that circular motion affects more adversely to cybersickness as compared to linear motion. Such information can be crucial to how VR applications require users to behave in the virtual world.

The work done on depth perception can help understand how humans use different cues to perceive virtual objects. An important feature of XR is to project images on to real world objects and also the ability to bring real world objects inside the virtual world. When users are navigating the VE, they are also physically moving in the real world. Recently, researches have focused on mapping the objects in the real world and placing objects of similar size and shape into the virtual world in order to avoid collisions. For this to work perfectly, users need to perceive the depths as true as possible. Therefore, it is essential to eliminate any conflicting depth cues such as VAC for the XR technologies.

To conclude, the work done in this thesis offers a peek into how the VR technology can be adapted to mimic how human visually perceive their surroundings in the real world. Once the discrepancies in the virtual and real world experiences have been eliminated, it will open the VR field to a wider range of audience and applications. Currently, the technology is only used for short sessions. However, as the technology improves, humans may well be able to fully immerse themselves in virtual worlds for weeks, potentially even their entire lives.

Bibliography

- [1] Rachel Albert, Anjul Patney, David Luebke and Joochwan Kim. 'Latency Requirements for Foveated Rendering in Virtual Reality'. In: *ACM Trans. Appl. Percept.* 14.4 (Sept. 2017). ISSN: 1544-3558. DOI: 10.1145/3127589. URL: <https://doi.org/10.1145/3127589>.
- [2] Thomas Alsop. *Augmented reality (AR) and virtual reality (VR) market size worldwide from 2016 to 2024*. Statista, 2021.
- [3] Samuel Ang and John Quarles. 'GingerVR: An Open Source Repository of Cybersickness Reduction Techniques for Unity'. In: *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 2020, pp. 460–463.
- [4] Laura L. Arns and Melinda M. Cerney. 'The relationship between age and incidence of cybersickness among immersive environment users'. In: *IEEE Proceedings. VR 2005. Virtual Reality, 2005*. 2005, pp. 267–268. DOI: 10.1109/VR.2005.1492788.
- [5] Giorgio Ballestin, Manuela Chessa and Fabio Solari. 'A Registration Framework for the Comparison of Video and Optical See Through Devices in Interactive Augmented Reality'. In: *IEEE Access* 9 (2021), pp. 64828–64843. DOI: 10.1109/ACCESS.2021.3075780.
- [6] Brian A. Barsky and Todd J. Kosloff. 'Algorithms for Rendering Depth of Field Effects in Computer Graphics'. In: *Proceedings of the 12th WSEAS International Conference on Computers*. ICCOMP'08. Heraklion, Greece: World Scientific, Engineering Academy and Society (WSEAS), 2008, pp. 999–1010. ISBN: 978-960-6766-85-5. URL: <http://dl.acm.org/citation.cfm?id=1513605.1513775>.
- [7] Behnam Bastani, Eric Turner, Carlin Vieri, Haomiao Jiang, Brian Funt and Nikhil Balram. 'Foveated Pipeline for AR/VR Head-Mounted Displays'. In: *Information Display* 33.6 (2017), pp. 14–35. DOI: 10.1002/j.2637-496X.2017.tb01040.x. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/j.2637-496X.2017.tb01040.x>.
- [8] Leif P. Berg and Judy M. Vance. 'Industry Use of Virtual Reality in Product Design and Manufacturing: A Survey'. In: *Virtual Real.* 21.1 (Mar. 2017), 1–17. ISSN: 1359-4338. DOI: 10.1007/s10055-016-0293-9. URL: <https://doi.org/10.1007/s10055-016-0293-9>.
- [9] Philip J. Bos, Liwei Li, Douglas Bryant, Afsoon Jamali and Achintya K. Bhowmik. '28-2: Invited Paper: Simple Method to Reduce Accommodation Fatigue in Virtual Reality and Augmented Reality Displays'. In: *SID Symposium Digest of Technical Papers* 47.1 (2016), pp. 354–357. DOI:

- 10.1002/sdtp.10678. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/sdtp.10678>.
- [10] Susan Bruck and Paul A. Watters. 'The factor structure of cybersickness'. In: *Displays* 32.4 (2011). Visual Image Safety, pp. 153–158. ISSN: 0141-9382. DOI: <https://doi.org/10.1016/j.displa.2011.07.002>. URL: <https://www.sciencedirect.com/science/article/pii/S014193821100059X>.
- [11] Gerd Bruder, Andreas Pusch and Frank Steinicke. 'Analyzing Effects of Geometric Rendering Parameters on Size and Distance Estimation in On-Axis Stereographics'. In: *Proceedings of the ACM Symposium on Applied Perception. SAP '12*. Los Angeles, California: Association for Computing Machinery, 2012, 111–118. ISBN: 9781450314312. DOI: 10.1145/2338676.2338699. URL: <https://doi.org/10.1145/2338676.2338699>.
- [12] Pulkit Budhiraja, Mark Miller, Abhishek Modi and David Forsyth. 'Rotation Blurring: Use of Artificial Blurring to Reduce Cybersickness in Virtual Reality First Person Shooters'. In: *ArXiv abs/1710.02599* (Oct. 2017). preprint.
- [13] Helmut Buhler, Sebastian Misztal and Jonas Schild. 'Reducing VR Sickness Through Peripheral Visual Effects'. In: *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. Reutlingen, Germany, 2018, pp. 517–9.
- [14] Norman Burningham, Zygmunt Pizlo and Jan P. Allebach. 'Image Quality Metrics'. In: *Encyclopedia of Imaging Science and Technology*. John Wiley & Sons, Ltd, 2002. ISBN: 9780471443391. DOI: <https://doi.org/10.1002/0471443395.img038>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/0471443395.img038>.
- [15] F. W. Campbell and G. Westheimer. 'Dynamics of accommodation responses of the human eye'. In: *The Journal of Physiology* 151.2 (1960), pp. 285–295. DOI: <https://doi.org/10.1113/jphysiol.1960.sp006438>. URL: <https://physoc.onlinelibrary.wiley.com/doi/abs/10.1113/jphysiol.1960.sp006438>.
- [16] Kieran Carnegie and Taehyun Rhee. 'Reducing Visual Discomfort with HMDs Using Dynamic Depth of Field'. In: *IEEE Computer Graphics and Applications* 35.5 (Sept. 2015), pp. 34–41. ISSN: 0272-1716. DOI: 10.1109/MCG.2015.98.
- [17] Eunhee Chang, Hyun Taek Kim and Byounghyun Yoo. 'Virtual Reality Sickness: A Review of Causes and Measurements'. In: *International Journal of Human-Computer Interaction* 36.17 (2020), pp. 1658–1682. DOI: 10.1080/10447318.2020.1778351. URL: <https://doi.org/10.1080/10447318.2020.1778351>.
- [18] David Checa and Andrés Bustillo. 'A Review of Immersive Virtual Reality Serious Games to enhance Learning and Training'. In: *Multimedia Tools and Applications* 79 (Mar. 2020). DOI: 10.1007/s11042-019-08348-9.

- [19] Manuela Chessa, Guido Maiello, Peter J. Bex and Fabio Solari. 'A space-variant model for motion interpretation across the visual field'. In: *Journal of Vision* 16.2 (Aug. 2016), pp. 12–12. ISSN: 1534-7362. DOI: 10.1167/16.2.12. URL: <https://doi.org/10.1167/16.2.12>.
- [20] Viviane Clay, Peter König and Sabine U. König. 'Eye tracking in virtual reality'. In: *Journal of Eye Movement Research* 12.1 (Apr. 2019). DOI: 10.16910/jemr.12.1.3. URL: <https://bop.unibe.ch/JEMR/article/view/4332-Clay-final-sub>.
- [21] Sue V. G. Cobb, Sarah Nichols, Amanda Ramsey and John R. Wilson. 'Virtual Reality-Induced Symptoms and Effects (VRISE)'. In: *Presence: Teleoperators and Virtual Environments* 8.2 (Apr. 1999), pp. 169–186. DOI: 10.1162/105474699566152.
- [22] James E. Cutting and Peter M. Vishton. 'Chapter 3 - Perceiving Layout and Knowing Distances: The Integration, Relative Potency, and Contextual Use of Different Information about Depth*'. In: *Perception of Space and Motion*. Ed. by William Epstein and Sheena Rogers. Handbook of Perception and Cognition. San Diego: Academic Press, 1995, pp. 69–117. ISBN: 978-0-12-240530-3. DOI: <https://doi.org/10.1016/B978-012240530-3/50005-5>. URL: <https://www.sciencedirect.com/science/article/pii/B9780122405303500055>.
- [23] S. Oliver Daum and Heiko Hecht. 'Distance estimation in vista space'. In: *Attention, Perception, & Psychophysics* 71 (2009), pp. 1127–1137.
- [24] Simon Davis, Keith Nesbitt and Eugene Nalivaiko. 'A Systematic Review of Cybersickness'. In: *Proceedings of the 2014 Conference on Interactive Entertainment*. IE2014. Newcastle, NSW, Australia: Association for Computing Machinery, 2014, 1–9. ISBN: 9781450327909. DOI: 10.1145/2677758.2677780. URL: <https://doi.org/10.1145/2677758.2677780>.
- [25] Ciro Donalek et al. 'Immersive and collaborative data visualization using virtual reality platforms'. In: *2014 IEEE International Conference on Big Data (Big Data)*. 2014, pp. 609–614. DOI: 10.1109/BigData.2014.7004282.
- [26] John V. Draper, David B. Kaber and John M. Usher. 'Telepresence'. In: *Human Factors* 40.3 (1998). PMID: 9849099, pp. 354–375. DOI: 10.1518/001872098779591386.
- [27] Natalia Dużmańska, Paweł Strojny and Agnieszka Strojny. 'Can Simulator Sickness Be Avoided? A Review on Temporal Aspects of Simulator Sickness'. In: *Frontiers in Psychology* 9 (2018), p. 2132. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2018.02132. URL: <https://www.frontiersin.org/article/10.3389/fpsyg.2018.02132>.
- [28] Bradley Efron and Robert J. Tibshirani. *An Introduction to the Bootstrap*. 1st ed. Chapman and Hall/CRC, 1994. ISBN: 9780429246593.

- [29] Zhenan Feng, Vicente A. González, Robert Amor, Ruggiero Lovreglio and Guillermo Cabrera-Guerrero. 'Immersive virtual reality serious games for evacuation training and research: A systematic literature review'. In: *Computers & Education* 127 (2018), pp. 252–266. ISSN: 0360-1315. DOI: <https://doi.org/10.1016/j.compedu.2018.09.002>. URL: <https://www.sciencedirect.com/science/article/pii/S0360131518302380>.
- [30] Ajoy S. Fernandes and Steven K. Feiner. 'Combating VR sickness through subtle dynamic field-of-view modification'. In: *2016 IEEE Symposium on 3D User Interfaces (3DUI)*. Greenville, South Carolina, USA, 2016, pp. 201–210.
- [31] Hamid Fesharaki, Leila Rezaei, Fereidoun Farrahi, Taghi Banihashem and Ahmad Jahanbkhshi. 'Normal Interpupillary Distance Values in an Iranian Population'. In: *Journal of ophthalmic & vision research* 7 (July 2012), pp. 231–234.
- [32] Linus Franke, Laura Fink, Jana Martschinke, Kai Selgrad and Marc Stamminger. 'Time-Warped Foveated Rendering for Virtual Reality Headsets'. In: *Computer Graphics Forum* 40.1 (2021), pp. 110–123. DOI: <https://doi.org/10.1111/cgf.14176>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/cgf.14176>.
- [33] Alireza M. Gavgani, Keith V. Nesbitt, Karen L. Blackmore and Eugene Nalivaiko. 'Profiling subjective symptoms and autonomic changes associated with cybersickness'. In: *Autonomic Neuroscience* 203 (2017), pp. 41–50. ISSN: 1566-0702. DOI: <https://doi.org/10.1016/j.autneu.2016.12.004>. URL: <https://www.sciencedirect.com/science/article/pii/S1566070216301096>.
- [34] Benjamin Gerschütz, Marius Fechter, Benjamin Schleich and Sandro Wartzack. 'A Review of Requirements and Approaches for Realistic Visual Perception in Virtual Reality'. In: *Proceedings of the Design Society: International Conference on Engineering Design* 1.1 (2019), 1893–1902. DOI: [10.1017/dsi.2019.195](https://doi.org/10.1017/dsi.2019.195).
- [35] John F Golding. 'Motion sickness susceptibility questionnaire revised and its relationship to other forms of sickness'. In: *Brain Research Bulletin* 47.5 (1998), pp. 507–516. ISSN: 0361-9230. DOI: [https://doi.org/10.1016/S0361-9230\(98\)00091-4](https://doi.org/10.1016/S0361-9230(98)00091-4). URL: <https://www.sciencedirect.com/science/article/pii/S0361923098000914>.
- [36] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. 4th ed. Pearson, 2016. ISBN: 9780134734804.
- [37] Brian Guenter, Mark Finch, Steven Drucker, Desney Tan and John Snyder. 'Foveated 3D Graphics'. In: *ACM Trans. Graph.* 31.6 (Nov. 2012). ISSN: 0730-0301. DOI: [10.1145/2366145.2366183](https://doi.org/10.1145/2366145.2366183). URL: <https://doi.org/10.1145/2366145.2366183>.

- [38] Jukka Hakkinen, Tero Vuori and Monika Puhakka. 'Postural stability and sickness symptoms after HMD use'. In: *IEEE International Conference on Systems, Man and Cybernetics*. Vol. 1. 2002, pp. 147–152. DOI: 10.1109/ICSMC.2002.1167964.
- [39] Morton L Heilig. *Sensorama simulator*. US Patent 3,050,870. Aug. 1962.
- [40] Robert T. Held, Emily A. Cooper and Martin S. Banks. 'Blur and Disparity Are Complementary Cues to Depth'. In: *Current Biology* 22.5 (2012), pp. 426–431. ISSN: 0960-9822. DOI: <https://doi.org/10.1016/j.cub.2012.01.033>. URL: <http://www.sciencedirect.com/science/article/pii/S0960982212000632>.
- [41] Robert T. Held, Emily A. Cooper, James F. O'Brien and Martin S. Banks. 'Using Blur to Affect Perceived Distance and Size'. In: *ACM Trans. Graph.* 29.2 (Apr. 2010), 19:1–19:16. ISSN: 0730-0301. DOI: 10.1145/1731047.1731057. URL: <http://doi.acm.org/10.1145/1731047.1731057>.
- [42] Gordon Heron, William N. Charman and Clifton Schor. 'Dynamics of the accommodation response to abrupt changes in target vergence as a function of age'. In: *Vision Research* 41.4 (2001), pp. 507–519. ISSN: 0042-6989. DOI: [https://doi.org/10.1016/S0042-6989\(00\)00282-0](https://doi.org/10.1016/S0042-6989(00)00282-0). URL: <https://www.sciencedirect.com/science/article/pii/S0042698900002820>.
- [43] Sébastien Hillaire, Anatole Lécuyer, Rémi Cozot and Géry Casiez. 'Depth-of-Field Blur Effects for First-Person Navigation in Virtual Environments'. In: *IEEE Computer Graphics and Applications* 28.6 (Nov. 2008), pp. 47–55. ISSN: 0272-1716. DOI: 10.1109/MCG.2008.113.
- [44] David M. Hoffman, Ahna R. Girshick, Kurt Akeley and Martin S. Banks. 'Vergence–accommodation conflicts hinder visual performance and cause visual fatigue'. In: *Journal of Vision* 8.3 (Mar. 2008), pp. 33–33. ISSN: 1534-7362. DOI: 10.1167/8.3.33. URL: <https://doi.org/10.1167/8.3.33>.
- [45] Matt C. Howard. 'A meta-analysis and systematic literature review of virtual reality rehabilitation programs'. In: *Computers in Human Behavior* 70 (2017), pp. 317–327. ISSN: 0747-5632. DOI: <https://doi.org/10.1016/j.chb.2017.01.013>. URL: <https://www.sciencedirect.com/science/article/pii/S0747563217300134>.
- [46] Raabid Hussain, Alain Lalande, Caroline Guigou and Alexis Bozorg-Grayeli. 'Contribution of Augmented Reality to Minimally Invasive Computer Assisted Cranial Base Surgery'. In: *IEEE Journal of Biomedical and Health Informatics* 24.7 (2020), pp. 2093–2106. DOI: 10.1109/JBHI.2019.2954003.
- [47] Razeen Hussain, Manuela Chessa and Fabio Solari. 'Modelling Foveated Depth-of-field Blur for Improving Depth Perception in Virtual Reality'. In: *2020 IEEE 4th International Conference on Image Processing, Applications and Systems (IPAS)*. 2020, pp. 71–76. DOI: 10.1109/IPAS50080.2020.9334947.

- [48] Razeen Hussain, Manuela Chessa and Fabio Solari. 'Mitigating Cyber-sickness in Virtual Reality Systems through Foveated Depth-of-Field Blur'. In: *Sensors* 21.12 (2021). ISSN: 1424-8220. DOI: 10.3390/s21124006. URL: <https://www.mdpi.com/1424-8220/21/12/4006>.
- [49] Razeen Hussain, Fabio Solari and Manuela Chessa. 'Simulated foveated depth-of-field blur for virtual reality systems'. In: *16th ACM SIGGRAPH European Conference on Visual Media Production*. London, UK, 2019. DOI: 10.13140/RG.2.2.35437.31208. URL: <https://www.cvmpp-conference.org/files/2019/short/24.pdf>.
- [50] Razeen Hussain, Fabio Solari and Manuela Chessa. *A Method based on Inverse Blurring to Mitigate Vergence Accommodation Conflict in Immersive Media Devices*. Patent [Under Review]. 2022.
- [51] Kenji Ibi. 'Characteristics of dynamic accommodation responses: comparison between the dominant and non-dominant eyes'. In: *Ophthalmic and Physiological Optics* 17.1 (1997), pp. 44–54. DOI: <https://doi.org/10.1046/j.1475-1313.1997.96000324.x>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1475-1313.1997.96000324.x>.
- [52] Lina Jansen, Selim Onat and Peter König. 'Influence of disparity on fixation and saccades in free viewing of natural scenes'. In: *Journal of Vision* 9.1 (Jan. 2009), pp. 29–29. ISSN: 1534-7362. DOI: 10.1167/9.1.29. URL: <https://doi.org/10.1167/9.1.29>.
- [53] David Johnson. *Introduction to and Review of Simulator Sickness Research*. Research Report. United States Army Research Institute for the Behavioral and Social Sciences, Apr. 2005.
- [54] Robert S. Kennedy, Norman E. Lane, Kevin S. Berbaum and Michael G. Lilienthal. 'Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness'. In: *The International Journal of Aviation Psychology* 3.3 (1993), pp. 203–220. DOI: 10.1207/s15327108ijap0303_3.
- [55] Alan Kenny, Hendrik Koesling and Declan Delaney. 'A preliminary investigation into eye gaze data in a first person shooter game'. In: *Proceedings of European Conference on Modelling and Simulation*. Riga, Latvia: Addison-Wesley Professional, 2005, pp. 733–738. ISBN: 978-1-63266-913-1.
- [56] Robert V. Kenyon, Moses Phenany, Dan Sandin and Thomas A. DeFanti. 'Accommodation and Size-Constancy of Virtual Objects'. In: *Annals of Biomedical Engineering* 36 (2007), pp. 342–348.
- [57] Behrang Keshavarz, Heiko Hecht and Ben D. Lawson. 'Visually Induced Motion Sickness: Causes, Characteristics, and Countermeasures'. In: *Handbook of Virtual Environments, 2nd ed.* 2014.
- [58] Mingyu Kim, Changyu Jeon and Jinmo Kim. 'A Study on Immersion and Presence of a Portable Hand Haptic System for Immersive Virtual Reality'. In: *Sensors* 17.5 (2017). ISSN: 1424-8220. DOI: 10.3390/s17051141. URL: <https://www.mdpi.com/1424-8220/17/5/1141>.

- [59] Eric Klein, J. Edward Swan, Gregory S. Schmidt, Mark A. Livingston and Oliver G. Staadt. 'Measurement Protocols for Medium-Field Distance Perception in Large-Screen Immersive Displays'. In: *2009 IEEE Virtual Reality Conference*. 2009, pp. 107–113. DOI: 10.1109/VR.2009.4811007.
- [60] Jaihyun Koh, Jangho Lee and Sungroh Yoon. 'Single-image deblurring with neural networks: A comparative survey'. In: *Computer Vision and Image Understanding* 203 (2021), p. 103134. ISSN: 1077-3142. DOI: <https://doi.org/10.1016/j.cviu.2020.103134>. URL: <https://www.sciencedirect.com/science/article/pii/S1077314220301533>.
- [61] Eugenia M. Kolasinski. *Simulator sickness in virtual environments*. Research Report. United States Army Research Institute for the Behavioral and Social Sciences, May 1995.
- [62] Gregory Kramida. 'Resolving the Vergence-Accommodation Conflict in Head-Mounted Displays'. In: *IEEE Transactions on Visualization and Computer Graphics* 22.7 (2016), pp. 1912–1931. DOI: 10.1109/TVCG.2015.2473855.
- [63] Orest Kupyn, Volodymyr Budzan, Mykola Mykhailych, Dmytro Mishkin and Jiří Matas. 'DeblurGAN: Blind Motion Deblurring Using Conditional Adversarial Networks'. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2018.
- [64] Joseph J. LaViola. 'A Discussion of Cybersickness in Virtual Environments'. In: *SIGCHI Bull.* 32.1 (Jan. 2000), 47–56. ISSN: 0736-6906. DOI: 10.1145/333329.333344. URL: <https://doi.org/10.1145/333329.333344>.
- [65] KE Laver, B Lange, S George, JE Deutsch, G Saposnik and M Crotty. 'Virtual reality for stroke rehabilitation'. In: *Cochrane Database of Systematic Reviews* 11 (2017). ISSN: 1465-1858. DOI: 10.1002/14651858.CD008349.pub4. URL: <https://doi.org/10.1002/14651858.CD008349.pub4>.
- [66] R. John Leigh and David S. Zee. *The Neurology of Eye Movements*. Oxford, UK: Oxford University Press, June 2015. ISBN: 9780199969289. DOI: 10.1093/med/9780199969289.001.0001.
- [67] Ajey Lele. 'Virtual reality and its military utility'. In: *Journal of Ambient Intelligence and Humanized Computing* 4 (Feb. 2011). DOI: 10.1007/s12652-011-0052-4.
- [68] Lan Li, Fei Yu, Dongquan Shi, Jianping Shi, Zongjun Tian, Jiquan Yang, Xingsong Wang and Phoenix Zhang. 'Application of virtual reality technology in clinical medicine'. In: *American Journal of Translational Research* 9 (Sept. 2017), pp. 3867–3880.

- [69] Yang Li, Jin Huang, Feng Tian, Hong-An Wang and Guo-Zhong Dai. 'Gesture interaction in virtual reality'. In: *Virtual Reality & Intelligent Hardware* 1.1 (2019), pp. 84–112. ISSN: 2096-5796. DOI: <https://doi.org/10.3724/SP.J.2096-5796.2018.0006>. URL: <https://www.sciencedirect.com/science/article/pii/S2096579619300075>.
- [70] Qiufeng Lin, Xianshi Xie, Aysu Erdemir, Gayathri Narasimham, Timothy P. McNamara, John Rieser and Bobby Bodenheimer. 'Egocentric Distance Perception in Real and HMD-Based Virtual Environments: The Effect of Limited Scanning Method'. In: *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization*. APGV '11. Toulouse, France: Association for Computing Machinery, 2011, 75–82. ISBN: 9781450308892. DOI: 10.1145/2077451.2077465. URL: <https://doi.org/10.1145/2077451.2077465>.
- [71] Yun-Xuan Lin, Rohith Venkatakrishnan, Roshan Venkatakrishnan, Elham Ebrahimi, Wen-Chieh Lin and Sabarish V. Babu. 'How the Presence and Size of Static Peripheral Blur Affects Cybersickness in Virtual Reality'. In: *ACM Trans. Appl. Percept.* 17.4 (Nov. 2020). ISSN: 1544-3558. DOI: 10.1145/3419984. URL: <https://doi.org/10.1145/3419984>.
- [72] Jingyu Liu, Claire Mantel and Søren Forchhammer. 'Perception-Driven Hybrid Foveated Depth of Field Rendering for Head-Mounted Displays'. In: *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2021, pp. 1–10. DOI: 10.1109/ISMAR52148.2021.00014.
- [73] Dillon J Lohr, Samantha Aziz and Oleg Komogortsev. 'Eye Movement Biometrics Using a New Dataset Collected in Virtual Reality'. In: *ACM Symposium on Eye Tracking Research and Applications*. ETRA '20 Adjunct. Stuttgart, Germany: Association for Computing Machinery, 2020. ISBN: 9781450371353. DOI: 10.1145/3379157.3391420. URL: <https://doi.org/10.1145/3379157.3391420>.
- [74] Colleen MacLachlan and Howard Howland. 'Normal values and standard deviations for pupil diameter and interpupillary distance in subjects aged 1 month to 19 years'. In: *Ophthalmic & physiological optics : the journal of the British College of Ophthalmic Opticians (Optometrists)* 22 (May 2002), pp. 175–82. DOI: 10.1046/j.1475-1313.2002.00023.x.
- [75] Guido Maiello, Manuela Chessa, Peter J. Bex and Fabio Solari. 'Near-optimal combination of disparity across a log-polar scaled visual field'. In: *PLOS Computational Biology* 16.4 (Apr. 2020), pp. 1–28. DOI: 10.1371/journal.pcbi.1007699. URL: <https://doi.org/10.1371/journal.pcbi.1007699>.
- [76] Guido Maiello, Manuela Chessa, Fabio Solari and Peter J. Bex. 'The (In)Effectiveness of Simulated Blur for Depth Perception in Naturalistic Images'. In: *PLOS ONE* 10.10 (Oct. 2015), pp. 1–15. DOI: 10.1371/journal.pone.0140230. URL: <https://doi.org/10.1371/journal.pone.0140230>.

- [77] Rafał K. Mantiuk, Gyorgy Denes, Alexandre Chapiro, Anton Kaplanyan, Gizem Rufo, Romain Bachy, Trisha Lian and Anjul Patney. 'Fov-VideoVDP: A Visible Difference Predictor for Wide Field-of-View Video'. In: *ACM Trans. Graph.* 40.4 (July 2021). ISSN: 0730-0301. DOI: 10.1145/3450626.3459831. URL: <https://doi.org/10.1145/3450626.3459831>.
- [78] R.G. Menendez and J.E. Bernard. 'Flight simulation in synthetic environments'. In: *19th DASC. 19th Digital Avionics Systems Conference. Proceedings (Cat. No.00CH37126)*. Vol. 1. 2000, 2A5/1-2A5/6 vol.1. DOI: 10.1109/DASC.2000.886902.
- [79] Xiaoxu Meng, Ruofei Du, Matthias Zwicker and Amitabh Varshney. 'Kernel Foveated Rendering'. In: *Proc. ACM Comput. Graph. Interact. Tech.* 1.1 (July 2018), 5:1-5:20. ISSN: 2577-6193. DOI: 10.1145/3203199. URL: <http://doi.acm.org/10.1145/3203199>.
- [80] Harold M. Merklinger. 'A technical view of bokeh'. In: *Photo Techniques* 18 (1997), pp. 1-5.
- [81] Patrick Millais, Simon L. Jones and Ryan Kelly. 'Exploring Data in Virtual Reality: Comparisons with 2D Data Visualizations'. In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems. CHI EA '18*. Montreal QC, Canada: Association for Computing Machinery, 2018, 1-6. ISBN: 9781450356213. DOI: 10.1145/3170427.3188537. URL: <https://doi.org/10.1145/3170427.3188537>.
- [82] Walter K. Murphy and Daniel M. Laskin. 'Intercanthal and interpupillary distance in the black population'. In: *Oral Surgery, Oral Medicine, Oral Pathology* 69.6 (1990), pp. 676-680. ISSN: 0030-4220. DOI: [https://doi.org/10.1016/0030-4220\(90\)90346-T](https://doi.org/10.1016/0030-4220(90)90346-T). URL: <https://www.sciencedirect.com/science/article/pii/003042209090346T>.
- [83] Eugene Nalivaiko, Simon L. Davis, Karen L. Blackmore, Andrew Vakulin and Keith V. Nesbitt. 'Cybersickness provoked by head-mounted display affects cutaneous vascular tone, heart rate and reaction time'. In: *Physiology & Behavior* 151 (2015), pp. 583-590. ISSN: 0031-9384. DOI: <https://doi.org/10.1016/j.physbeh.2015.08.043>. URL: <https://www.sciencedirect.com/science/article/pii/S0031938415300974>.
- [84] Phillip E. Napieralski, Bliss M. Altenhoff, Jeffrey W. Bertrand, Lindsay O. Long, Sabarish V. Babu, Christopher C. Pagano, Justin Kern and Timothy A. Davis. 'Near-Field Distance Perception in Real and Virtual Environments Using Both Verbal and Action Responses'. In: *ACM Trans. Appl. Percept.* 8.3 (Aug. 2011). ISSN: 1544-3558. DOI: 10.1145/2010325.2010328. URL: <https://doi.org/10.1145/2010325.2010328>.
- [85] Keith Nesbitt, Simon Davis, Karen Blackmore and Eugene Nalivaiko. 'Correlating reaction time and nausea measures with traditional measures of cybersickness'. In: *Displays* 48 (2017), pp. 1-8. ISSN: 0141-9382. DOI: <https://doi.org/10.1016/j.displa.2017.01.002>. URL: <https://www.sciencedirect.com/science/article/pii/S0141938216301032>.

- [86] Guang-Yu Nie, Henry Been-Lirn Duh, Yue Liu and Yongtian Wang. 'Analysis on Mitigation of Visually Induced Motion Sickness by Applying Dynamical Blurring on a User's Retina'. In: *IEEE Transactions on Visualization and Computer Graphics* 26.8 (Aug. 2020), 2535–2545. ISSN: 1077-2626. DOI: 10.1109/TVCG.2019.2893668. URL: <https://doi.org/10.1109/TVCG.2019.2893668>.
- [87] Nahal Norouzi, Gerd Bruder and Greg Welch. 'Assessing Vignetting as a Means to Reduce VR Sickness during Amplified Head Rotations'. In: *Proceedings of the 15th ACM Symposium on Applied Perception*. SAP '18. Vancouver, British Columbia, Canada: Association for Computing Machinery, 2018. ISBN: 9781450358941. DOI: 10.1145/3225153.3225162. URL: <https://doi.org/10.1145/3225153.3225162>.
- [88] Tom O'Haver. *A Pragmatic Introduction to Signal Processing*. May 2020. ISBN: 9798611266687.
- [89] Ekaterina Olshannikova, Aleksandr Ometov, Yevgeni Koucheryavy and Thomas Olsson. 'Visualizing Big Data with augmented and virtual reality: challenges and research agenda'. In: *Journal of Big Data* 2 (2015), pp. 1–27.
- [90] Kohei Oshima, Kenneth R Moser, Damien Constantine Rompapas, J. Edward Swan, Sei Ikeda, Goshiro Yamamoto, Takafumi Taketomi, Christian Sandor and Hirokazu Kato. 'SharpView: Improved clarity of defocused content on optical see-through head-mounted displays'. In: *2016 IEEE Symposium on 3D User Interfaces (3DUI)*. 2016, pp. 173–181. DOI: 10.1109/3DUI.2016.7460049.
- [91] Thomas Oskam, Alexander Hornung, Huw Bowles, Kenny Mitchell and Markus Gross. 'OSCAM - Optimized Stereoscopic Camera Control for Interactive 3D'. In: *ACM Trans. Graph.* 30.6 (Dec. 2011), 1–8. ISSN: 0730-0301. DOI: 10.1145/2070781.2024223. URL: <https://doi.org/10.1145/2070781.2024223>.
- [92] Ebi Peter Osuobeni and Khaled Al-Musa. 'Gender Differences in Interpupillary Distance among Arabs'. In: *Optometry and Vision Science* 70 (1993), pp. 1027–1030.
- [93] Nitish Padmanaban, Robert Konrad, Tal Stramer, Emily A. Cooper and Gordon Wetzstein. 'Optimizing virtual reality for all users through gaze-contingent and adaptive focus displays'. In: *Proceedings of the National Academy of Sciences* 114.9 (2017), pp. 2183–2188. ISSN: 0027-8424. DOI: 10.1073/pnas.1617251114. URL: <https://www.pnas.org/content/114/9/2183>.
- [94] Anjul Patney, Marco Salvi, JooHwan Kim, Anton Kaplanyan, Chris Wyman, Nir Benty, David Luebke and Aaron Lefohn. 'Towards Foveated Rendering for Gaze-tracked Virtual Reality'. In: *ACM Trans. Graph.* 35.6 (Nov. 2016), 179:1–179:12. ISSN: 0730-0301. DOI: 10.1145/2980179.2980246. URL: <http://doi.acm.org/10.1145/2980179.2980246>.

- [95] Etienne Peillard, Yuta Itoh, Guillaume Moreau, Jean-Marie Normand, Anatole Lécuyer and Ferran Argelaguet. ‘Can Retinal Projection Displays Improve Spatial Perception in Augmented Reality?’ In: *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2020, pp. 80–89. DOI: 10.1109/ISMAR50242.2020.00028.
- [96] Jeffrey S. Perry and Wilson S. Geisler. ‘Gaze-contingent real-time simulation of arbitrary visual fields’. In: *Human Vision and Electronic Imaging*. Vol. 4662. 2002. DOI: 10.1117/12.469554. URL: <https://doi.org/10.1117/12.469554>.
- [97] Lisa Rebenitsch and Charles Owen. ‘Review on Cybersickness in Applications and Visual Displays’. In: 20.2 (2016). ISSN: 1359-4338. DOI: 10.1007/s10055-016-0285-9. URL: <https://doi.org/10.1007/s10055-016-0285-9>.
- [98] Holger Regenbrecht and Thomas Schubert. ‘Real and Illusory Interactions Enhance Presence in Virtual Environments’. In: *Presence* 11 (Aug. 2002), pp. 425–434. DOI: 10.1162/105474602760204318.
- [99] Stephan Reichelt, Ralf Häussler, Gerald Fütterer and Norbert Leister. ‘Depth cues in human visual perception and their realization in 3D displays’. In: *Three-Dimensional Imaging, Visualization, and Display 2010 and Display Technologies and Applications for Defense, Security, and Avionics IV*. Ed. by Bahram Javidi, Jung-Young Son, John Tudor Thomas and Daniel D. Desjardins. Vol. 7690. International Society for Optics and Photonics. SPIE, 2010, pp. 92–103. DOI: 10.1117/12.850094. URL: <https://doi.org/10.1117/12.850094>.
- [100] Dongwei Ren, Kai Zhang, Qilong Wang, Qinghua Hu and Wangmeng Zuo. ‘Neural Blind Deconvolution Using Deep Priors’. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2020.
- [101] Rebekka S. Renner, Erik Steindecker, Mathias Müller, Boris M. Velichkovsky, Ralph Stelzer, Sebastian Pannasch and Jens R. Helmert. ‘The Influence of the Stereo Base on Blind and Sighted Reaches in a Virtual Environment’. In: *ACM Trans. Appl. Percept.* 12.2 (Mar. 2015). ISSN: 1544-3558. DOI: 10.1145/2724716. URL: <https://doi.org/10.1145/2724716>.
- [102] Rebekka S. Renner, Boris M. Velichkovsky and Jens R. Helmert. ‘The Perception of Egocentric Distances in Virtual Environments - A Review’. In: *ACM Comput. Surv.* 46.2 (Dec. 2013). ISSN: 0360-0300. DOI: 10.1145/2543581.2543590. URL: <https://doi.org/10.1145/2543581.2543590>.
- [103] Gary E. Riccio and Thomas A. Stoffregen. ‘An ecological Theory of Motion Sickness and Postural Instability’. In: *Ecological Psychology* 3.3 (1991), pp. 195–240. DOI: 10.1207/s15326969eco0303_2. URL: https://doi.org/10.1207/s15326969eco0303_2.

- [104] Giuseppe Riva. 'Virtual reality: an experiential tool for clinical psychology'. In: *British Journal of Guidance & Counselling* 37.3 (2009), pp. 337–345. DOI: 10.1080/03069880902957056. URL: <https://doi.org/10.1080/03069880902957056>.
- [105] Klaus Rohrschneider. 'Determination of the Location of the Fovea on the Fundus'. In: *Investigative Ophthalmology & Visual Science* 45.9 (Sept. 2004), pp. 3257–3258. ISSN: 1552-5783. DOI: 10.1167/iovs.03-1157. URL: <https://doi.org/10.1167/iovs.03-1157>.
- [106] Dimitrios Saredakis, Ancret Szpak, Brandon Birckhead, Hannah A. D. Keage, Albert Rizzo and Tobias Loetscher. 'Factors Associated With Virtual Reality Sickness in Head-Mounted Displays: A Systematic Review and Meta-Analysis'. In: *Frontiers in Human Neuroscience* 14 (2020), p. 96. ISSN: 1662-5161. DOI: 10.3389/fnhum.2020.00096. URL: <https://www.frontiersin.org/article/10.3389/fnhum.2020.00096>.
- [107] Christian J. Schuler, Michael Hirsch, Stefan Harmeling and Bernhard Schölkopf. 'Learning to Deblur'. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38.7 (2016), pp. 1439–1451. DOI: 10.1109/TPAMI.2015.2481418.
- [108] Jean-Christophe Servotte, Manon Goosse, Suzanne Hetzell Campbell, Nadia Dardenne, Bruno Pilote, Ivan L. Simoneau, Michèle Guillaume, Isabelle Bragard and Alexandre Ghuysen. 'Virtual Reality Experience: Immersion, Sense of Presence, and Cybersickness'. In: *Clinical Simulation in Nursing* 38 (2020), pp. 35–43. ISSN: 1876-1399. DOI: <https://doi.org/10.1016/j.ecns.2019.09.006>. URL: <https://www.sciencedirect.com/science/article/pii/S1876139918303244>.
- [109] H.R. Sheikh and A.C. Bovik. 'Image information and visual quality'. In: *IEEE Transactions on Image Processing* 15.2 (2006), pp. 430–444. DOI: 10.1109/TIP.2005.859378.
- [110] Thomas B. Sheridan. 'Musings on Telepresence and Virtual Presence'. In: *Presence: Teleoperators and Virtual Environments* 1.1 (Feb. 1992), pp. 120–126. DOI: 10.1162/pres.1992.1.1.120. URL: <https://doi.org/10.1162/pres.1992.1.1.120>.
- [111] Takashi Shibata, Joohwan Kim, David M. Hoffman and Martin S. Banks. 'The zone of comfort: Predicting visual discomfort with stereo displays'. In: *Journal of Vision* 11.8 (July 2011), pp. 11–11. ISSN: 1534-7362. DOI: 10.1167/11.8.11. URL: <https://doi.org/10.1167/11.8.11>.
- [112] Avi Shupak and Carlos R. Gordon. 'Motion Sickness: Advances in Pathogenesis, Prediction, Prevention, and Treatment'. In: *Aviation, Space, and Environmental Medicine* 77.12 (2006), pp. 1213–1223. ISSN: 0095-6562. URL: <https://www.ingentaconnect.com/content/asma/asem/2006/00000077/00000012/art00001>.

- [113] Vincent Sitzmann, Ana Serrano, Amy Pavel, Maneesh Agrawala, Diego Gutierrez, Belen Masia and Gordon Wetzstein. 'Saliency in VR: How Do People Explore Virtual Environments?' In: *IEEE Transactions on Visualization and Computer Graphics* 24.4 (2018), pp. 1633–1642. DOI: 10.1109/TVCG.2018.2793599.
- [114] Mel Slater and Anthony Steed. 'A Virtual Presence Counter'. In: *Presence* 9.5 (2000), pp. 413–434. DOI: 10.1162/105474600566925.
- [115] Fabio Solari, Manuela Chessa and Silvio P. Sabatini. 'Design strategies for direct multi-scale and multi-orientation feature extraction in the log-polar domain'. In: *Pattern Recognition Letters* 33.1 (2012), pp. 41–51. ISSN: 0167-8655. DOI: <https://doi.org/10.1016/j.patrec.2011.09.021>. URL: <http://www.sciencedirect.com/science/article/pii/S0167865511002947>.
- [116] Hans Strasburger, Ingo Rentschler and Martin Jüttner. 'Peripheral vision and pattern recognition: A review'. In: *Journal of Vision* 11.5 (Dec. 2011), pp. 13–13. ISSN: 1534-7362. DOI: 10.1167/11.5.13. URL: <https://doi.org/10.1167/11.5.13>.
- [117] Ivan E. Sutherland. 'The Ultimate Display'. In: *Proceedings of the IFIP Congress*. 1965, pp. 506–508.
- [118] Heidi Sveistrup. 'Motor rehabilitation using virtual reality'. In: *Journal of neuroengineering and rehabilitation* 1 (Jan. 2005), p. 10. DOI: 10.1186/1743-0003-1-10.
- [119] Nicholas T. Swafford, José A. Iglesias-Guitian, Charalampos Koniaris, Bochang Moon, Darren Cosker and Kenny Mitchell. 'User, Metric, and Computational Evaluation of Foveated Rendering Methods'. In: *Proceedings of the ACM Symposium on Applied Perception*. SAP '16. Anaheim, California: ACM, 2016, pp. 7–14. ISBN: 978-1-4503-4383-1. DOI: 10.1145/2931002.2931011. URL: <http://doi.acm.org/10.1145/2931002.2931011>.
- [120] Vicente J. Traver and Alexandre Bernardino. 'A Review of Log-polar Imaging for Visual Perception in Robotics'. In: *Robot. Auton. Syst.* 58.4 (Apr. 2010), pp. 378–398. ISSN: 0921-8890. DOI: 10.1016/j.robot.2009.10.002. URL: <http://dx.doi.org/10.1016/j.robot.2009.10.002>.
- [121] Michel Treisman. 'Motion Sickness: An Evolutionary Hypothesis'. In: *Science* 197.4302 (1977), pp. 493–495. DOI: 10.1126/science.301659. URL: <https://www.science.org/doi/abs/10.1126/science.301659>.
- [122] Subarna Tripathi and Brian Guenter. 'A Statistical Approach to Continuous Self-Calibrating Eye Gaze Tracking for Head-Mounted Virtual Reality Systems'. In: *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*. 2017, pp. 862–870. DOI: 10.1109/WACV.2017.101.

- [123] Jill Tucker and William N. Charman. 'Reaction and response times for accommodation'. In: *American journal of optometry and physiological optics* 56.8 (Aug. 1979), 490—503. ISSN: 0093-7002. DOI: 10.1097/00006324-197908000-00003. URL: <https://doi.org/10.1097/00006324-197908000-00003>.
- [124] Eric Turner, Haomiao Jiang, Damien Saint-Macary and Behnam Bastani. 'Phase-Aligned Foveated Rendering for Virtual Reality Headsets'. In: *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. Los Alamitos, CA, USA: IEEE Computer Society, Mar. 2018, pp. 1–2. DOI: 10.1109/VR.2018.8446142. URL: <https://doi.ieeecomputersociety.org/10.1109/VR.2018.8446142>.
- [125] Okan Tarhan Tursun, Elena Arabadzhyska-Koleva, Marek Wernikowski, Radosław Mantiuk, Hans-Peter Seidel, Karol Myszkowski and Piotr Didyk. 'Luminance-Contrast-Aware Foveated Rendering'. In: *ACM Trans. Graph.* 38.4 (July 2019). ISSN: 0730-0301. DOI: 10.1145/3306346.3322985. URL: <https://doi.org/10.1145/3306346.3322985>.
- [126] Nicolas Vignais, Richard Kulpa, Sébastien Brault, Damien Presse and Benoit Bideau. 'Which technology to investigate visual perception in sport: Video vs. virtual reality'. In: *Human Movement Science* 39 (2015), pp. 12–26. ISSN: 0167-9457. DOI: <https://doi.org/10.1016/j.humov.2014.10.006>. URL: <https://www.sciencedirect.com/science/article/pii/S0167945714001833>.
- [127] F.G. Waack. *Stereo photography: an introduction to stereo photo technology and practical suggestions for stereo photography*. F. G. Waack, 1985. URL: <https://books.google.it/books?id=YnQvcgAACAAJ>.
- [128] Peng Wang, Peng Wu, Jun Wang, Hung-Lin Chi and Xiangyu Wang. 'A Critical Review of the Use of Virtual Reality in Construction Engineering Education and Training'. In: *International Journal of Environmental Research and Public Health* 15.6 (2018). ISSN: 1660-4601. DOI: 10.3390/ijerph15061204. URL: <https://www.mdpi.com/1660-4601/15/6/1204>.
- [129] Zhou Wang and A.C. Bovik. 'A universal image quality index'. In: *IEEE Signal Processing Letters* 9.3 (2002), pp. 81–84. DOI: 10.1109/97.995823.
- [130] Zhou Wang, A.C. Bovik, H.R. Sheikh and E.P. Simoncelli. 'Image quality assessment: from error visibility to structural similarity'. In: *IEEE Transactions on Image Processing* 13.4 (2004), pp. 600–612. DOI: 10.1109/TIP.2003.819861.
- [131] Martin Weier, Thorsten Roth, André Hinkenjann and Philipp Slusallek. 'Foveated Depth-of-Field Filtering in Head-Mounted Displays'. In: *ACM Trans. Appl. Percept.* 15.4 (Sept. 2018). ISSN: 1544-3558. DOI: 10.1145/3238301. URL: <https://doi.org/10.1145/3238301>.
- [132] Martin Weier, Thorsten Roth, Andre Hinkenjann and Philipp Slusallek. 'Foveated Depth-of-field Filtering in Head-mounted Displays'. In: *Proceedings of the 15th ACM Symposium on Applied Perception*. SAP '18. Van-

- couver, British Columbia, Canada: ACM, 2018, 18a:1–18a:1. ISBN: 978-1-4503-5894-1. DOI: 10.1145/3225153.3243894.
- [133] Thomas Wickens. ‘Elementary Signal Detection Theory’. In: *Elementary Signal Detection Theory* (Oct. 2001). DOI: 10.1093/acprof:oso/9780195092509.001.0001.
- [134] Christopher Wilson and Alessandro Soranzo. ‘The Use of Virtual Reality in Psychology: A Case Study in Visual Perception’. In: *Computational and mathematical methods in medicine* 2015 (Aug. 2015), p. 151702. DOI: 10.1155/2015/151702.
- [135] Bob G. Witmer and Jr. Wallace J. Sadowski. ‘Nonvisually Guided Locomotion to a Previously Viewed Target in Real and Virtual Environments’. In: *Human Factors* 40.3 (1998), pp. 478–488. DOI: 10.1518/001872098779591340.
- [136] Josef Wolfartsberger. ‘Analyzing the potential of Virtual Reality for engineering design review’. In: *Automation in Construction* 104 (2019), pp. 27–37. ISSN: 0926-5805. DOI: <https://doi.org/10.1016/j.autcon.2019.03.018>. URL: <https://www.sciencedirect.com/science/article/pii/S0926580518312093>.
- [137] Xiangyu Xu, Jinshan Pan, Yu-Jin Zhang and Ming-Hsuan Yang. ‘Motion Blur Kernel Estimation via Deep Learning’. In: *IEEE Transactions on Image Processing* 27.1 (2018), pp. 194–205. DOI: 10.1109/TIP.2017.2753658.
- [138] Lin Yang, Longyu Zhang, Haiwei Dong, Abdulhameed Alelaiwi and Abdulmotaleb El Saddik. ‘Evaluating and Improving the Depth Accuracy of Kinect for Windows v2’. In: *IEEE Sensors Journal* 15.8 (2015), pp. 4275–4285. DOI: 10.1109/JSEN.2015.2416651.
- [139] Eugenia Yiannakopoulou, Nikolaos Nikiteas, Despina Perrea and Christos Tsigris. ‘Virtual reality simulators and training in laparoscopic surgery’. In: *International Journal of Surgery* 13 (2015), pp. 60–64. ISSN: 1743-9191. DOI: <https://doi.org/10.1016/j.ijssu.2014.11.014>.
- [140] Kaihao Zhang, Wenhan Luo, Yiran Zhong, Lin Ma, Bjorn Stenger, Wei Liu and Hongdong Li. ‘Deblurring by Realistic Blurring’. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2020.
- [141] Aneeq Zia et al. ‘Surgical Visual Domain Adaptation: Results from the MICCAI 2020 SurgVisDom Challenge’. In: *ArXiv abs/2102.13644* (2021).