Models & Methods to Enhance the Navigational Experience in Extended Reality

Nilotpal Biswas

Models & Methods to Enhance the Navigational Experience in Extended Reality

Thesis submitted in partial fulfilment
of the requirements for the degree of

Doctor of Philosophy

in

Computer Science and Engineering

by

Nilotpal Biswas

Under the supervision of

Dr. Samit Bhattacharya

Department of Computer Science and Engineering
Indian Institute of Technology Guwahati

July 2024

This thesis is dedicated to my parents, my teachers, my friends and my family

- whose blessings, inspiration, unconditional love and support made my path of success.

Declaration

I hereby certify that

a. The work contained in this thesis is original and has been done by myself and the general supervision of my supervisor.

b. The work has not been submitted to any other Institute for any degree or diploma.

c. Whenever I have used materials (data, theoretical analysis, results) from other sources, I have given due credit by citing them in the text of the thesis and giving their details in the references. Elaborate sentences used verbatim from published work have been clearly identified and quoted.

d. No part of this thesis can be considered plagiarism to the best of my knowledge and understanding and take complete responsibility if any complaint arises.

Date : ____/_____/_____ **Nilotpal Biswas**

Place: Guwahati, India

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
गुवाहाटी - 781039
Indian Institute of Technology Guwahati
Guwahati - 781039

Dr. Samit Bhattacharya
Associate Professor

: +91 361 258 2362
:samit@iitg.ac.in

Department of Computer Science and Engineering

Thesis Certificate

This is to certify that the thesis entitled **"Models & Methods to Enhance the Navigational Experience in Extended Reality"** being submitted by **Nilotpal Biswas** to the Department of Computer Science and Engineering, Indian Institute of Technology Guwahati, is a record of bonafide research work carried out by him under my supervision and is worthy of consideration for the award of the degree of Doctor of Philosophy of the Institute.

To the best of my knowledge, no part of the work reported in this thesis has been presented for the award of any degree at any other institution.

Date : ____/_____/_____

Place: Guwahati, India

_____

Dr. Samit Bhattacharya
(Thesis Supervisor)

Acknowledgments

Firstly, I wish to extend my heartfelt gratitude to my supervisor, Dr. Samit Bhattacharya, for his invaluable guidance and advice throughout my tenure at the Indian Institute of Technology Guwahati. His mentorship not only taught me the essence of research but also encouraged me to cultivate a habit of critical thinking, discipline, and perseverance. Furthermore, I am indebted to him for the opportunity to serve as a teaching assistant in NPTEL and Coursera, which not only bolstered my expertise in the subject matters but also provided financial support after my institute assistantship concluded. I am equally grateful to the other esteemed members of my doctoral committee, Prof. P.K. Das, Prof. T. Venkatesh, and Dr. Keyur Sorathia, for their constructive feedback that significantly shaped the direction of my thesis. The privilege of collaborating with them over the past few years has indubitably transformed me in numerous positive aspects.

I express my gratitude to all the distinguished faculty members of the Department of Computer Science and Engineering and the Head of the Department for fostering a conducive learning environment through their classes and laboratories. My sincere thanks are also directed to the Indian Institute of Technology Guwahati for its exceptional campus infrastructure, which ensured a comfortable academic and living experience. I extend my gratitude to the Ministry of Education, Government of India, for the monthly assistantship provided. I am thankful to all the staff members of the CSE department for their comprehensive technical, logistic, and administrative support during my stay at the institute. Additionally, I appreciate the Department of Science and Technology (DST) for funding my registration at the international conferences ISMAR 2021 and ISMAR 2022.

I have been fortunate to have a circle of wonderful friends and colleagues: Dr. Alakesh, Anandi, Aradhana, Arnab, Dr. Debanjan, Hemraj, Hrishikesh, Manju, Meenaxi, Mrinal, Nayan, Nidhi, Prasurjya, Dr. Rajkumar, Saikat, Sameer, Sanjay, Dr. Shimmila, Dr. Sourav, Sruti, Sumita, Supriya, Suraj, Tanuj and Yashdeep, whose support and inspiration have been invaluable. A special thanks to Dr. Maithilee, who has been a constant source of support through the highs and lows of this PhD journey.

I am eternally thankful to my B.Tech supervisor, Dr. Ranjan Maity, for his guidance from my undergraduate days to the present. I am blessed to have had the companionship of my labmates in the UCCN lab — Dr. Subrata, Dr. Hema, Dr. Ujjwal, Swagat, Dr. Abdalganiy, Dr Moustafa, Shakeel, Dhruba, Arpit, Ambuj, Anuj, Anamitra, Vishal, Vishwas and Siddharth — with whom I have shared countless memorable moments, making the lab a second home to me. I am also grateful for the camaraderie with my junior colleagues, particularly Abir, Ajanta, Prachurya, Praveen, Sandeepan, Saurav, Shahil, Soumyadeep and Suvarthi.

Most importantly, I owe a profound debt of gratitude to my supportive and understanding parents, whose sacrifices have been pivotal in my growth and comfort. Their parenting has been a blessing to me. My heartfelt appreciation also goes to my school teachers, whose guidance and wisdom laid the foundational stones for my academic and personal development. Lastly, I wish to convey my love and gratitude to my broader family members and relatives, whose warmth and encouragement have enriched my life in various ways. I also want to acknowledge my cousins, who have always made me feel as though I am not without a sibling.

Abstract

This thesis delves into enhancing navigational experiences within Extended Reality (XR), which includes Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR). It tackles significant challenges in XR navigation, with a focus on travel and wayfinding. Through several novel contributions, it aims to enhance realism, reduce cybersickness (CS), and improve wayfinding, ultimately leading to improved navigation experiences.

The first major contribution is the development of a model predicting instantaneous walking speed for users in system-automated VR tours. Recognizing the gap in providing a realistic walking experience where user input is minimal, this model leverages user data to simulate natural walking speeds, enhancing the realism of system-automated virtual tours. A user study comparing our proposed model's effectiveness against traditional approach of showing constant speed confirmed its superiority in enhancing perceived realism, marking a contribution in system-automated VR walking experiences.

Addressing the pervasive issue of CS in VR, the thesis undertakes a comprehensive systematic review of CS, analyzing causes, measurement techniques, and mitigation strategies across 223 research articles. This exhaustive review culminates in a novel taxonomy for CS measurement and a holistic mitigation framework, providing a valuable resource for both researchers and practitioners. Furthermore, it identifies substantial research gaps, laying a solid foundation for future investigations in CS prevention and management.

Another notable contribution is the optimization of system-automated VR tour durations to minimize user discomfort and CS. By theoretically computing the optimal path and time for covering all locations within a virtual environment, this approach significantly reduces the physical and psychological strain on users, enhancing the overall VR tour experience, without compromising the realistic walking speed. This optimization strategy is particularly beneficial in applications ranging from tourism to education, where mundane travel like walking is crucial. We also discuss how we can use vibrotactile feedback to avoid CS, and establish a parameter called visual gain used in our proposed realistic walking speed prediction model.

Additionally, this thesis addresses the gap in recognizing users' affective states during user-controlled VR tours. Recognizing the affective state can personalize the VR tour experience, and it is noteworthy that the user's emotional state also influences the likelihood of experiencing CS. Utilizing the built-in sensors of consumer-grade HMDs, this thesis focuses on developing a model to predict the emotional state of a user engaged in controller-based VR locomotion.

The thesis also introduces an innovative Virtual Locomotion Technique (VLT) named

"BreathWalk," utilizing controlled breathing to navigate virtual environments. This method not only mitigates CS by leveraging the calming effects of diaphragmatic breathing but also offers a novel and immersive way to interact with virtual spaces. The comparative analysis with joystick based locomotion, highlights BreathWalk's effectiveness in reducing CS and increasing user preference for navigation, underscoring the potential of breath-based interaction in VR.

Furthermore, the thesis tackles the challenge of cluttered off-screen Point of Interest (POI) visualization in smartphone-based (or handheld) AR, especially in high-density settings. By distributing 3D arrows along the screen edges and introducing a clutter threshold model, this contribution significantly improves spatial awareness and reduces visual clutter. The proposed method outperforms existing techniques like 3DWedge+ and Halo3D in accuracy and efficiency, offering a clutter-free and intuitive way of wayfinding in a smartphone-based AR system with a large number of POIs.

Collectively, these contributions address the multifaceted challenges of navigation in XR, offering novel solutions to enhance realism, address CS, and improve navigational experience. The findings have broad implications for the development of XR applications, promising immersive, comfortable, and intuitive experiences for users across various domains.

Keywords: XR Navigation, Realistic walking experience, Virtual tour, Cybersickness, Cybersickness survey, Affective state recognition, Cybersickness mitigation, POI visualisation, Virtual Reality, Augmented Reality

Table of Contents

List of Figures

Page

List of Tables

Page

List of Algorithms

List of Acronyms

Acronym	Expansion
XR	Extended Reality
AR	Augmented Reality
VR	Virtual Reality
VLT	Virtual Locomotion Technique
POI	Point of Interest
CS	Cybersickness
FOV	Field of View
DOF	Depth of Field
HMD	Head Mounted Display
RNN	Recurrent Neural Network
HRV	Heart Rate Variability
EDA	Electrodermal Activity
EEG	Electroencephalograph
GIGO	Garbage In, Garbage Out
SSQ	Simulator Sickness Questionnaire
VRSQ	Virtual Reality Sickness Questionnaire
CSQ	Cybersickness in Virtual Reality Questionnaire
MISC	Misery Scale
FMS	Fast Motion Sickness Scale
MHQ	Motion History Questionnaire
MSSQ	Motion Sickness Susceptibility Questionnaire
VIMSQ	Visually Induced Motion Sickness Susceptibility Questionnair
MSAQ	Motion Sickness Assessment Questionnaire
GSR	Galvanic Skin Response

ML	Machine Learning
DL	Deep Learning
ERP	Event-Related Potential
ECG	Electrocardiograms
EGG	Electrogastrogram
EOG	Electrooculograph
PPG	Photoplethysmography
COP	Center of Pressure
EBM	Explainable Boosting Machine
KNN	K-Nearest Neighbors
LDA	Linear Discriminant Analysis
LSTM	Long Short-Term Memory
CNN	Convolutional Neural Network
SVM	Support Vector Machine
DESOM	Deep Embedded Self-Organizing Map
DMSRS	Dynamic Mono-Stereoscopic Rendering System
IPD	Interpupillary Distance
HMD	Head-Mounted Display
HMM	Hidden Markov Model
GVS	Galvanic Vestibular Stimulation
DB	Diaphragmatic Breathing
F+C	Focus+Context
O+D	Overview+Detail
SISO	Slow In Slow Out
CV	Cross-Validation
LOSOCV	Leave-One-Subject-Out Cross-Validation
SD	Standard Deviation
TSP	Travelling Salesman Problem
VRP	Vehicle Routing Problem
VRPP	Vehicle Routing Problem with Profit
TOP	Team Orientering Problem
SSQ	Simulator Sickness Questionnaire
CH	Cultural Heritage

SAM	Self-Assessment Manikin
LSTM	Long Term Short Memory
PH	Positive High
PL	Positive Low
NH	Negative High
NL	Negative Low
PQ	Presence Questionnaire

Chapter

1

Introduction

Extended Reality (XR) refers to a spectrum of technologies that merge the physical and virtual worlds, creating immersive and interactive experiences. This umbrella term encompasses Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR), each offering distinct levels of immersion and interaction. VR generates computer-generated environments that users can fully immerse themselves in. VR can be experienced in several ways, offering a wide range of applications across various industries. VR experiences are supported by a variety of platforms, from simple mobile VR setups to more advanced headsets and devices [1]. By simulating a completely virtual environment, VR facilitates an immersive experience that can be entirely different from the physical world. AR, on the other hand, overlays digital information onto the real world, enhancing the user's perception of their physical surroundings. AR is commonly experienced through smartphones or specialized glasses, enriching the user's environment with additional digital content, such as information pop-ups, directions, or virtual objects seamlessly integrated into the real world [2]. MR combines the real and virtual worlds, anchoring virtual objects to the real world and allowing interaction with them. It is a hybrid of VR and AR, offering the best of both technologies. MR headsets, like the HoloLens, combine the ability to see the real world along with superimposed holographic images, fostering innovative applications in various fields like gaming, education, healthcare, and industry [3]. While distinctions between AR and MR exist in their technical underpinnings, their user experience often blurs the lines. Both technologies superimpose digital layers onto the physical world, allowing users to manipulate

Figure 1.1: Reality Virtuality Continum

virtual objects within the real space. Whether navigating virtual creatures through your park or visualizing furniture placement in your home, the boundaries between AR and MR become less about the technical definition and more about the specific application and user perception. The progression of these technologies continues to redefine how we perceive and interact with our environment, offering vast potential for entertainment, education, tourism, and numerous other industries. Figure 1.1, taken from Kolivand et al. [4], illustrates the XR landscape and how VR/AR/MR is connected to both real and virtual environments.

.

In the realm of XR, interactions between the user and the virtual environment can be primarily categorized into three fundamental types: *Selection*, *Manipulation*, and *Navigation* [5]. Each of these interaction types plays a crucial role in how users engage with and experience the XR environment, offering unique functionalities and possibilities. *Selection* is the initial interaction type, where users choose or pick objects or elements within the virtual environment. This process is often facilitated through the use of handheld controllers, gestures, or even eye tracking, depending on the technology in use. The act of selection enables users to make choices, initiate actions, or interact with specific items in the XR space. This is a critical step in the interactive process, as it allows users to single out which objects or elements they wish to engage with from the multitude of possibilities presented within the virtual environment [6]. Following the act of selection is *Manipulation*, which refers to the ability of users to modify or change the virtual objects they have selected. This can include moving, rotating, resizing, or altering the properties of these objects. Manipulation allows users to engage deeply with digital elements, offering them the opportunity to manipulate these elements and potentially even create new content within the XR environment. This fosters a sense of agency and creativity, as users are not just passive observers but active

participants who can shape their virtual surroundings [7]. Lastly, *Navigation* encompasses how users move within the XR space. This interaction type is about the different methods and techniques that allow for the planning, exploration, and traversal of virtual environments. Navigation is fundamental to the XR experience, as it determines how users can move from one point to another, explore virtual spaces, and discover new aspects of the virtual world. It includes a variety of mechanisms, from teleportation and flying to walking in place, each designed to enhance the sense of immersion and realism within the XR environment [8].

1.1 Navigation in XR

Navigation in XR encompasses two essential components: *Travel* and *Wayfinding* [9].

1.1.1 Travel in XR

Travel represents the motor component, involving the physical movement or locomotion within the virtual environment. In the context of VR, travel encompasses different techniques facilitating movement within the virtual environment. These techniques, known as Virtual Locomotion Techniques (VLTs), encompass diverse means such as real walking [10], VR controllers [11], gestures [12], and specialized devices like omnidirectional treadmills [13]. VLTs enable users to traverse virtual spaces seamlessly, ensuring a heightened sense of immersion. However, in AR and MR, the concept of travel shifts from navigating virtual environments to interacting within the user's physical environment augmented by digital elements. In AR and MR, the user's physical movement in the real world constitutes their travel, while digital overlays or enhancements augment this real-world experience, influencing their exploration and interaction within the merged digital-physical environment [14, 15].

VR Tours

Travel in VR spans diverse fields like gaming, architecture, design, training, tourism, real estate, and education [16, 17, 18, 19]. In these applications, VR tours have emerged as a standout, gaining substantial prominence. VR tours redefine exploration by offering immersive and interactive experiences that enable individuals to delve into VR environments representing a wide array of destinations, properties, museums, and educational simulations

[17]. VR tours can be categorized into two primary modes: *User-Controlled* and *System-Automated* [20]. A user-controlled VR tour refers to a tour in which the user has control over their travel and navigation within the VR environment. This typically involves the user using VR controllers or other input modalities to move and explore the virtual space according to their preferences and actions. On the other hand, a system-automated VR tour involves the system or application controlling the user's travel within the VR environment. In this type of tour, the user's movement and navigation are guided or controlled by the system itself, often following a predetermined path or sequence. Each mode presents its own set of advantages and drawbacks, influencing the user experience significantly. A user-controlled VR tour can provide:

- Enhanced user agency: Users have complete control over navigation, offering a sense of freedom and personalization, which can have a significant positive effect on presence [21].

- Enhanced learning experience: Various locomotion methods, such as gaze/head-directed steering and walk-in-place, enable users to interact more intimately with the VR environment, thereby enhancing learning in the tour [22, 23].

- Engagement and immersion: Direct control fosters a deeper sense of immersion, enhancing the overall experience [24].

- Incomplete exploration: Users might miss important aspects or areas if they overlook or skip sections while navigating. This can result in an incomplete understanding or experience of the VR environment [25].

- Chances of getting lost or confused: Users might find it overwhelming or confusing to navigate through the virtual space, especially if it's complex or lacks clear directions.

On the other hand, using a System-Automated VR Tour can provide:

- Guided exploration: Users don't need to worry about navigation or decision-making. The system guides them through the tour, reducing the potential for confusion or disorientation, especially for those less familiar with VR environments [26].

- Focused presentation of information: The system can direct users' attention to specific elements, ensuring they do not miss crucial details. This controlled approach can be beneficial for delivering a focused narrative or educational content [27].

- Efficiency and convenience: Automated tours can be more time-efficient as users can passively experience the tour without the need to actively control their movement. This convenience can be particularly useful in scenarios where time is limited.

- Limited interactivity and personalization: Users have minimal to no control over their exploration. This lack of interactivity might lead to a less engaging experience for individuals who prefer more autonomy or desire personalized interactions [28].

- Lack of realism: Due to less physical involvement, the users find it less realistic than the user-controlled tours [29].

1.1.2 Wayfinding in XR

Wayfinding stands as the cognitive component of navigation in XR. It involves the mental processes, strategies, and tools used by users to orient themselves, understand their position, and find directions within the XR environment [5]. Wayfinding techniques can include maps, markers, waypoints, or visual indicators that aid users in navigating complex environments or reaching specific destinations [30]. Effective wayfinding is crucial for enhancing user orientation and ensuring a smooth and intuitive exploration experience within XR applications. Combining both travel and wayfinding elements creates a cohesive and user-friendly navigation system that contributes to a more immersive and enjoyable XR experience [31]. Among many techniques to enhance wayfinding experience, visualizing off-screen Points of Interest (POIs) is one of them. A POI is a location in a physical or digital environment that is of particular interest or significance to users. Off-screen POI visualization refers to the representation of POIs that are located outside the current field of view within an XR environment [32]. This technique is designed to provide users with directional and distance information about these off-screen POIs, enabling them to navigate and locate these points more effectively within the virtual or augmented environment. Enhancing wayfinding experiences through off-screen POI visualization involves improving the user's ability to locate and engage with points of interest that are not immediately visible on the

screen. This can be especially useful in scenarios where there are numerous POIs or when the user is navigating through dense environments where many POIs may be occluded or off-screen. The benefits of off-screen POI visualization in enhancing wayfinding experiences include:

- Improved Navigation: By providing directional and distance information about off-screen POIs, users can more effectively navigate and locate these points within the virtual or augmented environment, leading to a more efficient and seamless navigation experience [33].

- Spatial Awareness: Off-screen POI visualization contributes to enhancing users' spatial awareness within the virtual or augmented environment, allowing them to better understand the spatial relationships and locations of off-screen POIs in relation to their current position [34].

- Reduced Cognitive Load: The visualization of off-screen POIs can reduce the cognitive load on users by providing them with clear and intuitive cues for locating and interacting with these points of interest without requiring excessive mental effort [35].

- Enhanced User Experience: Knowing the positions of off-screen POIs can make virtual experiences more enjoyable and less frustrating by minimizing the chances of getting lost or missing critical elements of the environment [36].

- Support for Strategic Decision Making: Knowing the locations of off-screen POIs allows users to make strategic decisions based on a more comprehensive understanding of the environment [37].

1.2 Challenges and Motivations

This dissertation explores the challenges linked with navigation in XR. In the VR segment, the focus is on the aspect of travel, addressing the challenges encountered in both system-automated and user-controlled tours. In the AR domain, the emphasis is on the complexities of visualizing POI for wayfinding purposes. These challenges are each explored in the following.

1.2.1 Challenges in the VR travel systems

As discussed earlier, VR travel systems can be further sub-categorised as system-automated and user-controlled tours. In the following sections, we explore the challenges associated with each of these categories.

Challenges in the system-automated tours

While system-automated tours offer an effortless experience with reduced cognitive demands [29], allowing users to concentrate more on the experience itself, their lack of user involvement makes them less suitable for scenarios where a realistic walking experience is crucial. This issue has not been extensively explored in system-automated VR tours. Therefore, improving the walking experience in such tours is vital. A key factor influencing a user's experience in a VR environment is the travel speed. To create a realistic walking experience, it's essential to use the user's actual walking speed as the travel speed [38, 39]. However, determining the user's actual walking speed in an automated tour is challenging, as these systems typically do not allow for continuous user input, making it difficult to align the user's input with the walking speed. Previous research has employed a constant walking speed as an optic flow to increase realism [40]. Nevertheless, since walking speed varies during travel [41], we believe that displaying a visual speed that matches the user's instantaneous walking speed can enhance realism more effectively than showing a constant speed.

In system-automated tours, a significant challenge arises from the potential for sensory mismatch, where the visual cues associated with movement do not match the expectations of the vestibular system. This discrepancy can lead to cybersickness (CS), a prevalent issue in VR locomotion. Although aligning the visual optic flow closely with real walking experiences can enhance realism in VR, achieving such realistic visual stimuli in system-automated tours without inducing discomfort remains a complex challenge. Various mitigation strategies have been employed, including adjusting the field of view (FOV), modifying the depth of field (DOF), and implementing vibrotactile feedback. Notably, the use of vibrotactile feedback has emerged as an unobtrusive method for alleviating sickness. However, the application of instantaneous walking speed in conjunction with vibrotactile feedback has yet to be explored. Additionally, a positive correlation has been observed between the duration of the VR tour and the severity of sickness, yet no studies have specifically addressed reducing the duration

of system-automated tours in the literature.

Challenges in the user-controlled tours

While user-controlled VR tours enhance user agency, engagement, and physical interaction, they also have the challenge of CS, particularly when the user remains stationary. To reduce the impact of CS, it is crucial to predict its likelihood. Various methods for predicting CS have been researched. Notably, studies have established a connection between a user's emotional state and the likelihood of experiencing CS. This link underscores the importance of identifying and addressing emotional reactions within VR experiences to reduce the risk of CS. Additionally, recognizing emotions can aid in personalizing the VR tour, thereby enhancing the overall user experience. However, research focused on recognizing emotional states using sensors readily available in commercial Head-Mounted Displays (HMDs) has been limited.

In user-controlled tours, particularly those employing steering-based locomotion methods that involve continuous movement to navigate within the VR environment, there's a notable induction of CS [42]. Due to this adverse effect, teleportation is often favored over steering-based locomotion as a means to mitigate CS. However, while teleportation reduces the likelihood of CS, it can also create a disorienting effect that disrupts the user's sense of presence within the VR environment. Among the various strategies to lessen the impact of CS, controlled diaphragmatic breathing (CDB) has been identified as an effective technique to reduce sickness [43]. Despite its effectiveness in mitigating CS, the potential of using controlled breathing as an input mechanism for locomotion in VR has not yet been explored, indicating a gap in research in this direction.

1.2.2 Challenges in the off-screen POI visualization in an AR system

As mentioned earlier, off-screen POI visualization refers to the depiction of POIs that are situated beyond the current field of view in an XR environment. Various techniques have been explored to address the visualization of off-screen POIs [33, 34, 35]. These methods aim to represent these POIs by utilizing different visual cues. Metaphors like wedges [44], halos [45], and arrows [46] have been employed to indicate the presence of POIs in the surrounding environment. The objective is to provide users with distance and direction

awareness, guiding them to orient their devices and align the POIs within their field of view, thus enhancing their overall awareness of the surrounding POIs.

The visualization techniques are applied across different platforms, including both HMD and handheld devices. Techniques like EyeSee360 [47], 3D Wedge+ [48], and CompassbAR [49] have been evaluated for visualizing POIs in the HMD context. Conversely, techniques such as the 3D Bezier Curve [36], 3D Arrows [32], and AroundPlot [50] have been assessed for use on handheld devices.

However, due to the limited size of the smartphone screen, the use of excessive on-screen cues or large-sized visualizations on handheld devices can lead to visual clutter, making it difficult for users to distinguish between different POIs [44]. Additionally, within the existing literature, there is a noted trend of adopting metaphors like 3D wedges and 3D arrows, which are often positioned in the central region of the display screen. This approach, while making the cues prominent, carries the risk of occluding the main content displayed on the smartphone. Furthermore, when the POIs vary significantly in altitude relative to the user's position, the visualization technique needs to also indicate these vertical differences. This is crucial for users to comprehend the spatial arrangement of the POIs without needing to move their devices excessively. This situation is commonly encountered in AR applications that need to showcase different off-screen POIs, such as restaurants located on various floors of a building or when a user is at a dense market intersection point. Researchers have introduced different techniques for visualizing off-screen POIs, such as displaying a halo on the screen based on the position of the off-screen POIs [45], showing the POIs in a mini-map view [34], and using 2D arrows to point towards the POIs [32]. However, these techniques have not adequately addressed the challenge of distinguishing the vertical position of off-screen POIs, which remains a critical aspect for comprehensive and effective visualization in AR environments.

1.2.3 Objective

The main objective of the thesis is to come up with different methods to enhance the users' navigational experience. The thesis focuses on enhancing realism, addressing CS and contribute on wayfinding to improve navigation in XR. Based on the discussed challenges of navigation, the specific goals of the thesis are outlined below.

- **Enhancing realism in a system automated VR tour:** A system automated tour lacks realism due to the limited involvement of user's interactions. We aim to enhance the realism by building a model to predict the instantaneous walking speed of the user and implementing it during the system-automated VR tour.

- **Mitigating CS in system-automated VR tours without compromising realistic walking experience:** We intend to explore the proposed walking speed prediction model while vibrotactile feedback based CS mitigation strategy is used. Moreover, lengthy exposure to VR can result in severe CS issues and in a system automated tour, the duration of the interaction time and the path of the VR environment can be manipulated by the system itself. Therefore, we aim to come up with methods for finding the optimal path in a VR environment in minimum duration and also allowing users to visit maximum places in the tour in a given time.

- **Recognition of affective state in user-controlled VR tours without using any additional sensors and equipment:** Recognizing the affective state in user-controlled VR tours can significantly enhance the personalization of the VR tours. Furthermore, the user's emotional state has a considerable impact on the likelihood of inducing CS. Our objective is to detect the user's emotional state by leveraging the built-in sensors found in consumer-grade HMDs. This will allow the developers to foresee potential occurrences of CS during user-controlled VR tours and also personalise the content based on the users' emotional state, thereby improving the VR experience.

- **Mitigating CS in user controlled VR tours:** Studies have found relationship of CDB in mitigating the onset of CS. We intend to incorporate breathing-based walking locomotion to reduce CS in user controlled VR tours.

- **Designing clutter-free off-screen POI visualization system for handheld AR system:** Visualizing a large number of off-screen POIs in the small screen of a smartphone is challenging. We aim to propose a handheld AR-based off-screen POI visualizer designed explicitly for high-density environments characterized by a larger number of POIs.

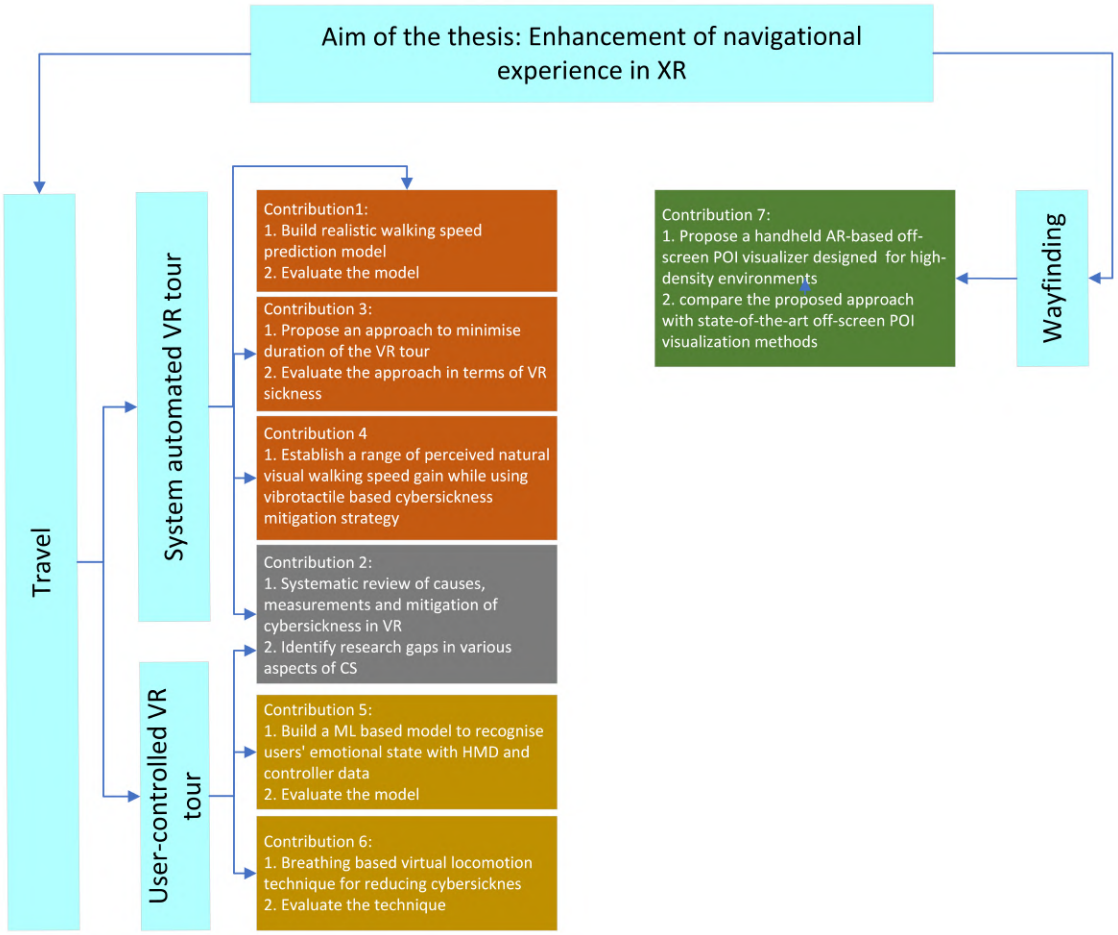


Figure 1.2: Visual Representation of the Thesis Contributions

1.3 Contributions of the Thesis

The research objectives related to navigation in XR systems, as outlined in the previous section, are addressed in the remainder of this thesis. We offer a comprehensive synopsis of the research contributions that tackle these challenges, as follows.

1.3.1 Modelling realistic walking speed for system-automated VR tour

This contribution lies in proposing a model for the prediction of instantaneous walking speed for users experiencing a system-automated VR tour, without the need for continuous user input. This model addresses the challenge of providing a realistic walking experience in system-automated VR tours, where the users' travel is controlled by the system with no continuous input required from the user. Whether exploring historical sites, museums, or virtual real estate tours, realistic walking speeds can enhance the sense of presence, making users feel like they are truly exploring these spaces in person. In order to build a model that can predict the instantaneous walking speed of a person, we started with collecting walking speed and frequency data from 40 participants. After plotting a graph depicting distance traveled versus frequency, we observed a pattern which mirrors characteristics akin to a pulse wave commonly utilized in digital signal processing, resembling a square pulse wave and this type of pulse wave can be well approximated by taking the first few terms of the corresponding Fourier expansion. With this idea, we modeled the instantaneous walking frequency. In a user study involving 34 participants, we carried out a comparison to assess the perceived realism between our proposed speed prediction model and the traditional method of using a constant walking speed during a system-automated VR tour. The results of this comparison demonstrated that participants showed a significant preference for the instantaneous speed model, perceiving it as significantly more realistic than the constant speed approach. This finding underscores our model's potential to enhance the quality of system-automated VR tours by more accurately simulating the natural experience of walking. However, we recognize the potential risk of inducing CS when continuous visual motion is presented during the tour. This awareness has motivated us to delve deeper into the subject of CS, which we explore in our subsequent contributions.

1.3.2 Systematic review of causes, measurements and mitigation of CS in VR

This contribution involves a thorough analysis of 223 relevant studies, focusing on critical aspects including the causes, measurement techniques, and strategies for mitigation. This comprehensive review, often overlooked in similar studies, provided crucial insights into various dimensions of CS. Specifically, our analysis of 81 relevant sources enabled the development of a new taxonomy encompassing diverse measurement methodologies used in CS research. This taxonomy serves as a comprehensive overview, assisting researchers in making well-informed decisions regarding CS measurement approaches. Additionally, our synthesis of insights from 78 studies on CS mitigation strategies highlighted the absence of comprehensive guidelines. Hence, we formulated a robust framework applicable to both developers and users, integrating strategies discussed in the literature from both perspectives. By scrutinizing 87 research articles, we created a comprehensive taxonomy categorizing factors that induce CS, facilitating a more organized understanding of these variables. Our systematic review also revealed significant research gaps across various dimensions of CS, thus contributing substantially to the evolving body of knowledge on CS and laying groundwork for future investigations.

1.3.3 Minimising the duration of a system-automated VR tour

As discussed earlier, the duration of a VR tour significantly influences user experience, with extended exposure potentially leading to CS and inducing fatigue due to the weight of the HMD. Our work aims to address these concerns by proposing a novel approach to minimize the duration of system-automated VR tours within a given VR environment. Our strategy involves optimizing the path connecting various sites in the VR setting, theoretically computing the optimal time and path to cover all locations. We also cover another variation of the problem, which is to cover maximum locations within a stipulated time frame. To achieve these, we conceptualize the problem akin to the renowned Vehicle Routing Problem with Profit (VRPP). We discuss the use case by implementing our strategy to build an application offering an immersive VR tour of *Majuli*, the largest river island in Assam, India. To evaluate the efficacy of our approach, we conducted a between-subject study involving 40 participants, comparing the induction of CS and discomfort while navigating the

optimal path prescribed by our method against a tour following an arbitrary user-selected path. Our findings underscore the substantial reduction in VR sickness and discomfort by adhering to the optimized path, demonstrating the potential of our approach in enhancing user experience by minimizing VR tour durations and associated health concerns. We also discuss how the walking speed prediction model proposed in section 1.3.1 can be merged while using this duration minimisation strategy.

1.3.4 Establishing a range of visual gain for perceived realistic walking speed while using vibrotactile-based CS mitigation strategy

In our efforts to enhance the realism of system-automated VR tours by implementing a realistic walking speed model, we encountered a challenge associated with potential CS due to continuous motion. Existing research has explored the use of vibrotactile feedback behind the ear to mitigate CS during these tours [40]. However, the application of vibrotactile feedback impacts the visual gain factor in the proposed walking speed model described in Section 1.3.1, making it inapplicable as the visual gain varies with different locomotion techniques. To address this limitation, we conducted a study focusing on perceived natural walking speeds in stationary settings while utilizing vibrotactile feedback to prevent CS. Participants were exposed to visual gain ranging from 1.0 to 3.0, revealing a new range of visual gain for perceived naturalness spanning from 1.40 to 1.78. This study emphasizes the integration of the walking speed prediction model in conjunction with vibrotactile feedback as a mitigation strategy within stationary VLTs. By exploring perceived natural walking speeds and considering the impact of vibrotactile feedback, our work seeks to refine the application of the walking speed prediction model, enhancing user experiences in stationary VLTs amidst efforts to reduce CS during system-automated VR tours.

1.3.5 Recognising users' emotional state from HMD and handheld controller data during a VR tour

Overcoming the limitations posed by the absence of integrated biosensors and eye-tracking capabilities in many commercial HMDs, we introduce a Sequential Recurrent Neural Network (RNN) based model to recognise the emotional state of a user while navigating in a VR tour assuming the VR controller as an input mechanism. The main intention behind this work

was to check if we can recognise the emotional state of a person while s/he is navigating in a neutral environment. Although there are several datasets available for emotion recognition, we did not find any dataset where data from the VR controller is included. Therefore, we created a dataset of 17 participants while navigating in the VR environment. We first induced them in different emotional states by exposing them to different evident visual stimuli. Once they were in the intended emotional state, we asked them to navigate in a neutral VR environment using a joystick and the trigger of a VR controller. We collected joystick pressure (multiple samples for an event), trigger pressure(multiple samples for an event), controller velocity, controller quaternion rotation and HMD quaternion rotation while pressing the controller buttons (joystick/trigger). We built and validated the model to predict the emotional state in four classes based on the Circumplex model of affect, namely, High-Positive, High-Negative, Low-Positive and Low-Negative.

1.3.6 Breathing-based VLT for reducing CS in VR

Inspired by the positive impact of CDB on CS [43] and acknowledging the merits of breath as an interaction modality [51], in this contribution, we introduce a new VLT named "BreathWalk." In this technique, users walk in a VR environment by utilizing their breath, captured through a nose-attached microphone. Adjusting their breathing frequency, from low to high, allows users to modulate their walking speed, while steering direction is achieved by directing their torso accordingly. To track the torso, we attached one VR controller (the non-dominant hand) which was mounted on the lower chest of the user. Leveraging breath as an input mechanism not only aims to mitigate CS but also seeks to enhance immersion within the virtual space. In our study, we conducted a comparative analysis between BreathWalk, and joystick-based steering regarding CS, realism, task performance and preference. We found BreathWalk to be significantly better in terms of CS, realism and preference. Conversely, the joystick-controlled method yielded better in task performance.

1.3.7 Off-Screen POI visualization for handheld AR in vertically dense regions to reduce clutter

This contribution is on the wayfinding part of the navigation in XR. Here, we've addressed the challenge of visualizing off-screen POIs in 3D environments viewed through a smartphone,

particularly in high-density settings (more than 15 POIs). Traditional methods tend to clutter the screen, impacting visualization efficiency. To tackle this, we propose a novel method using 3D arrows strategically distributed over the edges of the screen rather than at the center. We divide the edges into multiple cells and each cell contains a 3D arrow, directing to a cluster of POIs. Additionally, we introduced a model called the cluttering threshold to determine the number of POIs, a cell should contain given the average locate time. If the number of POIs exceeds the cluttering threshold, we filter it out on a priority basis. To filter out POIs on the basis of radial distance, we use a distance filter. Through a study involving 16 participants, our approach demonstrated higher efficiency and accuracy compared to existing state-of-the-art methods like 3D Wedge+ and Halo3D, especially in high-density POI environments. Our work aims to facilitate precise off-screen POI location and distance interpretation in smartphone-based AR, ensuring effective user spatial awareness even amidst larger POI volumes, thereby enhancing the navigational experience.

1.4 Organization of the Thesis

The rest of the thesis is organized as follows:

Chapter 2 forms the backbone of this dissertation. It lays the groundwork by conducting a critical examination of the literature concerning several challenges including the enhancement of realism in system-automated VR tours, the causes, measurement/detection, and mitigation strategies for CS (second contribution), as well as an extensive survey of current off-screen POIs visualization techniques.

Chapter 3 provides an detailed description of the first thesis contribution. It elaborates on the process involving the development and validation of our proposed model designed to predict the instantaneous speed within a system-automated VR tour.

Chapter 4 delves into the fourth contribution of the thesis, addressing the incorporation of visual gain into the proposed speed prediction model while implementing a vibrotactile-based CS mitigation strategy. After that we discuss the third thesis contribution, outlining the methodology aimed at reducing the duration of a system-automated VR tour.

Chapter 5 elucidates the fifth thesis contribution, detailing the process of recognizing emotional states through HMD and handheld controller data.

Chapter 6 presents the sixth thesis contribution, introducing an breathing-controlled VLT designed specifically for user-controlled VR tours with a focus on mitigating CS. Additionally, this chapter encompasses a comparative analysis between this new technique and the joystick-based steering VLT, evaluating their effectiveness in terms of CS mitigation, realism, task performance, and user preference.

Chapter 7 elaborates on the seventh thesis contribution, detailing an off-screen POIs visualization technique aimed at diminishing on-screen clutter. Additionally, this chapter conducts a comparative analysis of the task performance between this new technique and established methods, namely 3D Wedge+ and Halo 3D, evaluating their efficiency and accuracy.

Finally, **Chapter 8** gives a summary of the contributions of this thesis and discuss the potential avenues for future research exploration.

CHAPTER

2

Related Works

In this chapter, we start with a brief discussion of various types of VLTs to navigate in a VR environment. After that, we discuss various works related to providing realistic walking experience in VR. We then highlight the importance of providing a realistic walking experience in system-automated VR tours. After that, we present a systematic literature review on the causes, measurement, and mitigation of CS. We also discuss various gaps in CS research. Subsequently, we explore the concept of off-screen POIs for enhancing wayfinding strategies within XR, with a focus on methods to improve off-screen POI visibility on AR handheld devices.

2.1 Traveling techniques in VR

Travelling in VR can be achieved through various VLTs, as categorized by Zayer et al. [29], into walking-based, steering-based, selection-based, manipulation-based, and multiscale VLTs. Walking-based VLTs simulate walking within VR, enhancing presence and naturalism through methods like Real Walking [10] and redirection techniques [52] to mirror human movement closely, improving spatial orientation and reducing VR sickness. Steering-based VLTs [42] allow users to control direction and speed, using body movements or vehicular props to navigate intuitively within VR environments, enhancing user engagement and navigation. Selection-based VLTs automate movement towards a chosen destination, simplifying navigation by using target selection and route planning [53, 5] techniques to

minimize cognitive load, with teleportation and map dragging as examples. Manipulation-based VLTs employ user gestures for navigating or manipulating the virtual world, like the "Eyeball-in-hand" metaphor[54], but may be less efficient in speed and comfort. Multiscale VLTs [55, 56] cater to navigation across different scales and vertical movements, offering both active and automatic scaling methods to adjust the virtual scale, improving usability and comfort. Nilsson et al.'s [57] study further categorizes VLTs based on user mobility and interaction metaphors, distinguishing between mobile and stationary techniques, mundane and magical forms of movement, and vehicular versus body-centric techniques, aiming to enhance immersion and address the challenges of navigating virtual spaces within physical limitations.

As previously mentioned, interactions within a VR tour can be categorized as either user-controlled or system-automated. Among the VLTs discussed in this section, VR tours utilizing user input for movement within the VR environment — such as those employing walking-based techniques, steering-based techniques, manipulation-based techniques, and active scaling multiscale VLTs can be classified under user-controlled VR tours. On the other hand, tours that rely on selection-based VLTs (particularly for navigating from a source to a destination automatically), cinematic VR navigation scenarios [27], and automatic scaling VLTs [55] can be categorized under system-automated VR tours.

2.2 Realistic walking experience in VR

Walking is a crucial mode of travel in real life. Providing a realistic walking experience is helpful in many VR applications [20]. For instance, to implement an experience of walking through a cultural heritage site or walking through the street side of a city, providing a realistic walking experience can be beneficial. In order to improve the realistic walking experience, researchers have come up with many VLTs. Obviously, the more closely a VLT can mimic human walking, the more realistic it becomes. In real walking, the walker gets the surrounding information from several sources. It involves external (visual, auditory and cutaneous), internal(vestibular and proprioceptive) and efferent (efference copy and attention allocation) information [58, 59, 40]. Researchers have used many techniques to improve the experience of walking by manipulating the internal and efferent information. Some works [60, 61, 62] presented the use of haptic and auditory feedback in enhancing

the real walking experience. Vision, however, is known to be the most dominant sensory organ providing spatial information. Some works [40, 62] applied visual head oscillation to provide a more realistic walking experience. One of the major visual factors affecting the experience of walking is the walking speed [38]. In the prior work, several efforts [63, 64] were made to make WIP more natural, with the primary intention of generating speed profiles that are closer to real walking. To measure the walking speed of a user, researchers have mapped the user input with the walking speed [63, 39]. For instance, some works used external devices such as wristband [65, 66], handheld devices [67], lower body-mounted Inertial Measurement Unit (IMU) [68] and head-mounted IMU [69] to predict the walking speed. Yan and allison [64] calibrated their WIP implementation to deliver more realistic speeds by using data obtained from the back of individuals' legs while physically walking. These works are, however, suitable for user-controlled virtual tours.

Though walking is a multisensory experience, in the case of system-automated tours, due to the low interaction of the user with the system (the user is not moving their body parts), the primary area of improvement is convincing the visual sense. As mentioned earlier, providing visual optic flow close to real walking can enhance the walking experience in VR. However, unlike a user-controlled tour, a system-automated tour does not allow continuous input while travelling from one point to another. In literature, we found very less focus on implementing realistic walking experiences for system-automated tours. The probable reason for this could be that adding continuous camera motion to non-interactive or semi-interactive locomotion can cause a sensory mismatch, thereby causing VR sickness [70]. The sensory mismatch is a widely accepted theory for the cause of VR sickness, which states that if there is a mismatch between the visual information related to motion and orientation with the vestibular information, it can create a sense of nausea [71] (We discuss about CS in detail in the next section). However, prior works [40, 72] have confirmed in their study that the use of vibrotactile feedback in the back of the two ears could significantly reduce the chances of VR sickness while using stationary VLTs . These works open up an opportunity to use natural walking speed as the visual speed for the system-automated tours.

We found that researchers have been using constant speed in a system-automated virtual tour. Few works have proposed biomechanically inspired models for predicting the

average walking speed of a user. Inman et al. [73] expressed human walking speed ($|V|$) as a product of step frequency (f) and step length (l):

$$|V| = f \times l \tag{2.1}$$

Furthermore, Dean et al. [74] stated that the step length and height of a person are positively correlated. Using these relationships, they expressed walking speed ($|V|$) as:

$$|V| = \left(\frac{f}{0.157} \times \frac{h}{1.72}\right)^2 \tag{2.2}$$

Where $|V|$ (in cm/sec), f and h represent a person's walking speed, walking frequency, and height, respectively. Researchers [75, 39] used this relation in order to estimate the actual walking speed in their work. However, the walking speed of a person is not constant throughout the walk [41]. The instantaneous walking speed of a user throughout the travel can mimic the actual walking speed better than the constant average speed.

2.3 CS in VR

CS is an inevitable phenomenon in VR. The term "cybersickness" was first coined in the early 1990s, around the same time that VR technology was first becoming widely available. In the realm of CS, significant literature reviews have been conducted, including the systematic review by Davis et al. [76] focusing on CS measurement, which highlighted the need for updated analysis due to advancements in measurement techniques and mitigation strategies. Rebenitsch and Owen's review [77] further delved into factors contributing to CS, such as field of view and navigation, indicating the necessity for more comprehensive research. Recent reviews by Chang et al. [78], Caserman et al. [79], and Saredakis et al. [80] have predominantly explored the causes of CS, with lesser emphasis on its measurement and mitigation. Additionally, Yildirim and Caglar reviewed deep learning (DL) approaches for CS classification using EEG signals [81], Yang et al. examined machine learning (ML) methods for CS detection/prediction [82], and Li et al. looked into electrostimulation methods for CS mitigation [83], alongside reviews focusing on gender [84], latency [85], and user susceptibility to CS [86], although these lacked a holistic perspective on CS's causes,

measurement, and mitigation strategies. To fill this gap and have a holistic understanding of CS, we performed a systematic literature review providing a comprehensive analysis of CS's causes, measurement, and mitigation, offering a holistic understanding of the condition, its potential consequences, and strategies for measurement and mitigation, alongside CS mitigation guidelines for users and developers and identifying critical research gaps to inform future studies, making it a valuable resource for VR system developers and users.

2.3.1 Systematic review methodology

Our review targets recent developments in CS, encompassing causes, measurement techniques, and mitigation strategies. Leveraging databases like ACM DL, IEEE, PubMed, and Web of Science known for VR research, we aimed to dissect current research and uncover unexplored areas. This comprehensive review, drawing on search strings shown in table 2.1, covered papers from 2014 to February 2023, reflecting the surge in VR research. From 1005 records found plus an additional 80 from Google Scholar, we removed 180 duplicates, 22 short poster/workshop papers, and 4 non-English or inaccessible papers, following the PRISMA protocol [87, 88]. A panel of three raters, achieving a 97.3% agreement after a calibration exercise on 150 papers for consistent evaluation, conducted the initial screening to identify relevant studies for a deeper understanding of CS and the enhancement of prevention and mitigation strategies. During this initial screening phase, publications were labeled according to the following criteria:

- Relevant, wherein the search terms referred on the title, abstract, and/or conclusion are relevant (providing cause, measurement, or mitigation for CS in VR).

- Not Relevant, wherein the search terms referred on the title, abstract, and/or conclusion are not relevant (Does not provide any information towards causes, measurement, or mitigation for CS in VR).

Out of 879 remaining records, we found 223 relevant articles. We categorized the relevant papers into causes, measurement, and mitigation of CS. We found 21 articles that belonged to both cause and measurement or measurement and mitigation or all three of them. Lastly, forward and backward citation searching was used to incorporate highly relevant and/or cited publications that were not captured by our search [89].

Database	Search String	Total results
ACM digital library	[[Abstract: "VR"] OR [Abstract: or "vr"]] AND [Abstract: "*sickness"] AND [E-Publication Date: (01/01/2014 TO 02/28/2023)]	110
IEEE Xplore Digital Library	("Abstract":"VR" OR "Abstract":"VR") AND ("Abstract":"*sickness") Filters Applied: 2014 - 2023	295
PubMed Central (PMC)	(("VR [Title/Abstract] OR "VR" [Title/Abstract]) AND ("*sickness" [Title/Abstract])) AND (("2014/01/01"[Date - Publication] : "3000"[Date - Publication]))	251
Web Of Science	(AB=("VR" OR "VR")) AND AB=("sickness") Timespan: 2014-01-01 to 2023-02-28 (Index Date)	349

Table 2.1: Databases and search strings used

2.3.2 Causes of CS

CS is a complex phenomenon that is yet to be fully understood. Several theories have attempted to explain its causes, including the sensory conflict theory [90, 91, 92, 93, 94], postural instability theory [95], and the evolutionary theory, often referred to as the poison theory [96]. Various studies have suggested that vection might be responsible for the occurrence of CS [97, 98, 99], although there is contradictory evidence showing that CS can occur without eliciting vection [100]. Prothero et al.'s rest frame hypothesis [101] suggests that motion sickness is caused by conflicting stationary frames of reference rather than conflicting orientation and motion cues. These theories propose that conflicting cues from different senses, an unstable posture, and an evolutionary response to perceived poisoning can all contribute to CS. However, none of these theories fully explain the subjectivity involved in CS or predict its severity in different virtual environments. We identified several key factors that can contribute to the onset of CS. To better organize and understand these factors, we have developed a comprehensive taxonomy that breaks them down into specific categories. By examining each of these categories in detail, we hope to shed some light

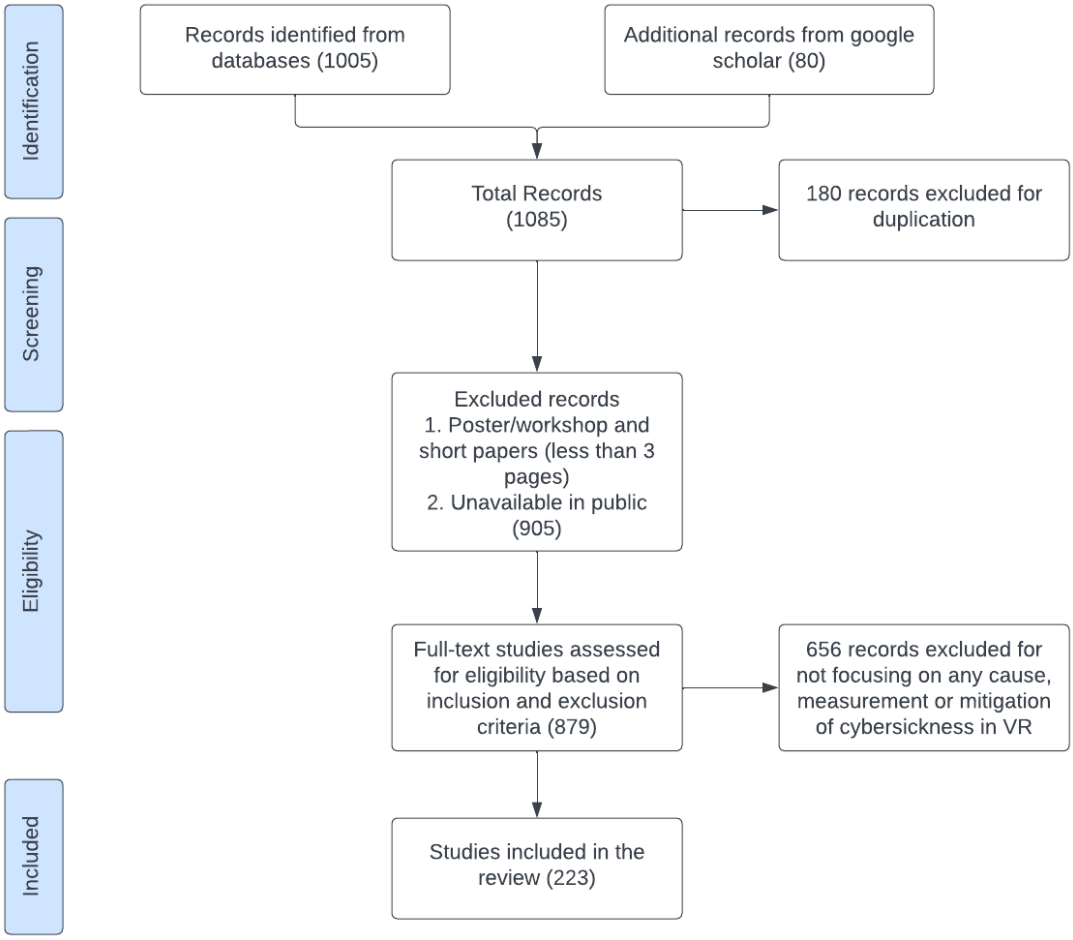


Figure 2.1: Prisma Flow Chart

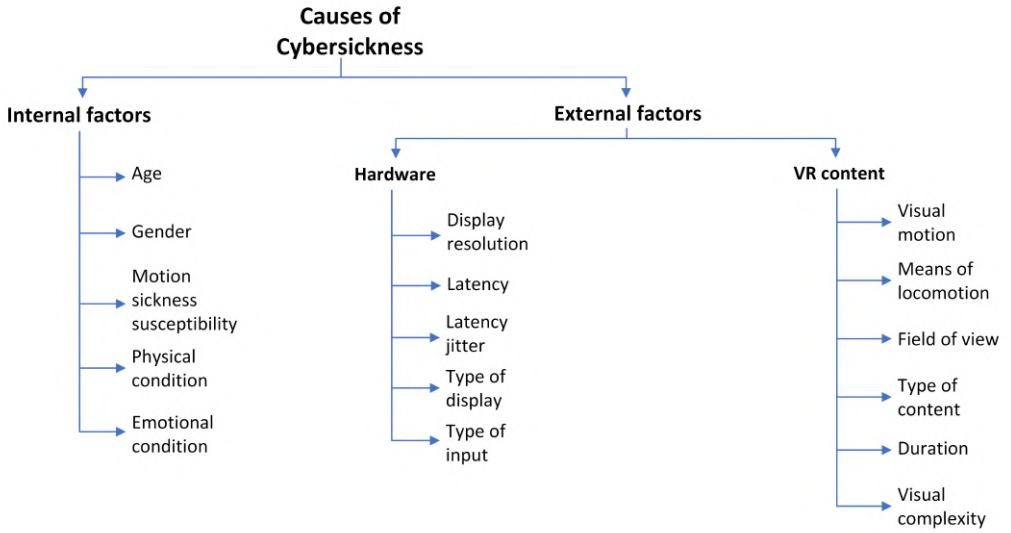


Figure 2.2: Taxonomy of various causes of CS

on the complex array of factors that can contribute to CS and provide a more nuanced understanding of this phenomenon.

Internal Factors

When we refer to "internal factors" in the context of CS, we are specifically talking about a set of individual characteristics and conditions that can contribute to an increased likelihood of experiencing CS. These factors are inherent to the user and can include a variety of different elements.

Age Research indicates that susceptibility to CS increases from ages 2 to 12, then decreases from 12 to 21 [102]. Older individuals tend to experience higher levels of CS [103], possibly due to age-related changes in the vestibular system. Contrarily, CS is rare in very young children [104], potentially because their frequent carrying and movement might impact their vestibular response. Yet, a study [105] suggests that while general VR content's impact on CS is age-independent, age influences CS levels with controllable VR content. This highlights the importance of considering age in developing VR experiences to mitigate CS risks.

Gender The role of gender in CS is complex, with studies showing mixed results. Some research indicates that women may experience more severe CS, while other studies find no significant gender differences in CS severity. A notable study [106] found an association between the proportion of female participants using HMDs and increased simulator sickness, highlighting the need for diverse sampling in VR research. The interpupillary distance (IPD) issue, where females often struggle to find a suitable fit, suggests a need for VR displays with a broader IPD adjustment range [107]. The impact of gender on CS, influenced by factors like IPD, underscores the necessity for further investigation to clarify these findings.

Motion sickness susceptibility Research links motion sickness susceptibility to increased risk of CS. Motion sickness, caused by conflicting sensory signals, leads to symptoms like nausea, similar to those induced by VR's realistic movements. Studies [108, 105, 109, 110] show a positive correlation between motion sickness susceptibility and CS risk. For instance, Gavgani et al. [108] found that individuals with higher motion sickness susceptibility experienced more severe CS symptoms in VR. Similarly, Katsigiannis et al. [111] noted that participants with greater motion sickness symptoms were more prone to CS using a VR exercise bike, suggesting a higher CS risk for those more susceptible to motion sickness. Additionally, prior VR experience influences CS likelihood. Individuals with previous VR exposure tend to have lower CS risk [103, 112]; Garrido et al. [103] observed that more VR exposure correlates with fewer CS symptoms. Repeated VR exposure also reduces CS symptoms over time [113], indicating the role of familiarity and adaptation in mitigating CS.

Physical conditions Body awareness, the perception of bodily sensations, has been linked to increased susceptibility to CS as per Mittelstädt et al. [114]. Altena et al. [115] found that poor sleep quality could induce CS, indicating that sleep improvement might mitigate its effects. Kim and Kim [116] reported that individuals with greater retinal eccentricity are more prone to CS. Furthermore, those with neurological conditions like multiple sclerosis tend to experience more severe CS symptoms [117]. Xu et al. [118] demonstrated that the user's physical position affects CS susceptibility, with standing participants experiencing more symptoms than seated ones. These studies highlight the significance of addressing physical conditions and user posture in CS mitigation during VR usage.

Emotional conditions Some studies have explored the relationship between CS and the emotional state of the user. Studies [119, 102, 120] have found that anxiety can contribute to the experience of CS in immersive VR environments. One study [121] found that negative emotions, especially anxiety and frustration, were associated with an increased likelihood of experiencing VIMS. The study suggested that regulating emotions could be a useful strategy for reducing CS. Additionally, it has been found that individual susceptibility to CS is also associated with negative emotional responses to VR experiences, and personality traits such as neuroticism increased the likelihood of experiencing CS [122]. A study by Wang et al. [102] used fuzzy logic to predict CS during VR navigation and found that individual differences in emotional responses, particularly anxiety, and stress, were crucial factors in predicting CS. These studies suggest that understanding individual differences in emotional regulation may be important for reducing the likelihood of CS during VR experiences.

Various methods, including physiological and behavioral measures, are crucial for emotion recognition in VR. The majority of research employs heart rate variability (HRV) and electrodermal activity (EDA) for measuring emotional arousal, with advanced analyses like HRV frequency and non-linear domain analyses and partitioning EDA signals into tonic and phasic components, offering deeper insights. Additionally, central nervous system measures like electroencephalography (EEG) have become significant, providing unique insights into emotional responses. Behavioral measures, such as eye-tracking, head movement, and gait patterns, reveal emotional states through observable behaviors, with some studies exploring the correlation between gamepad pressure and player arousal and others predicting affective states from smartphone interactions. Despite these advancements, there's a lack of techniques for detecting affective states using sensors in consumer-level VR devices like Oculus Quest and HTC Vive.

External Factors

Hardware The quality of the VR experience of the user is significantly affected by the hardware used. Previous studies initially associated CS symptoms with inferior equipment, assuming that the advancement of technology would alleviate user discomfort [123]. Despite numerous hardware enhancements, CS persists, as some aspects of the device are believed to be contributing to the occurrence of VIMS.

Display resolution "Display resolution" refers to the count of pixels on a screen, indicating clarity and detail level through pixel rows and columns. In VR, higher display resolutions are pivotal for crafting realistic, immersive experiences, potentially diminishing CS by enhancing environmental seamlessness. Wang et al. [124] demonstrated that increasing render resolution improves gameplay experience and mitigates CS to a certain threshold without affecting performance. Their research showed significant CS reduction when moving from 1k to 2k resolutions but found no notable differences in CS between 2k, 3k, and 4k resolutions, highlighting the diminishing returns on higher resolutions beyond 2k for reducing CS.

Latency Latency, the delay between a user's action and the virtual environment's response, is a key factor in CS. Studies, including Palmisano et al. [125], have shown that lower latency reduces CS by aligning visual and vestibular cues more closely. This research also indicated that discrepancies in head orientation between virtual and physical realms heighten CS, underscoring the conflict between sensory systems. Further research [126, 127, 128] corroborates that simulator sickness intensifies with increased latency, leading to heightened nausea, disorientation, and eye strain. Kuwamura et al. [127] observed that higher latency affects stability, causing more sway in users. These insights underline the importance of minimizing latency in VR design to improve the user experience and mitigate CS.

Latency Jitter Jitter, the variability in latency, leads to irregular data delivery timings, causing user discomfort and disorientation. Stauffert et al. [129] demonstrated that latency jitter significantly increases CS incidence, with higher jitter levels correlating with more severe symptoms. Minimizing latency jitter is crucial for enhancing VR experiences and reducing simulator sickness.

Type of display The display type also significantly affects CS severity. Mittelstaedt et al. [130] found that users of HMDs reported more CS than those using conventional monitors in a virtual bike simulator study. Similarly, Choy et al. [131] observed that VR headset users experienced more CS compared to users of flat or panoramic screens when viewing stereoscopic 3D videos. Additionally, higher-end HMDs were associated with reduced CS compared to lower-end models [132]. These findings underscore the importance of selecting appropriate display types for VR systems to mitigate CS and enhance user experience.

Type of input Input methods in VR, like hand controllers, treadmills, and motion sensors, influence CS levels. Mittelstaedt et al. [130] observed in a virtual bike simulator study that motion controls, such as steering with handlebars, induced more CS than non-motion controls like joysticks. Monteiro et al. [133] found in a study on first-person shooter games that gamepad controllers caused less CS than hand controllers. These findings highlight that the choice of input method significantly affects CS, necessitating further research to elucidate the relationship between various input methods and CS across different VR settings.

VR content The design of the virtual environment can have a significant impact on the likelihood and severity of CS. The virtual content-related factors are as follows:

Visual motion Visual motion stimulus, the movement perceived in a virtual environment, is a key cause of CS [134]. Studies have shown that higher speeds of visual motion elevate CS risk [135, 136]. The content's movement within virtual environments significantly impacts CS levels, as found by Tran et al. [137] in their evaluation of 360-degree videos' effects on user experience. They discovered that content movement within these videos influenced CS, unlike the device used for rendering. Keshavraj et al. [134] identified a strong link between vection (sensation of self-motion) and visual motion, especially when peripheral stimuli were involved. CS severity escalates with increased height and speed in the virtual environment [136], although brief exposure to turbulent visual motion alone did not induce CS [138]. Additionally, angular velocity increase leads to higher CS intensity, particularly when rotation occurs around the X and Y axes rather than the Z axis [139].

Means of Locomotion Research on locomotion methods in virtual environments reveals their impact on CS due to sensory mismatches between virtual and physical movements. Control over motion significantly influences CS; Venkatakrishnan et al. [140] reported that control directly affects presence and indirectly affects CS through presence, with a greater sense of presence potentially reducing CS. Contrarily, another study [141] indicated that participants with motion control experienced more CS than those without, particularly noting that hand-based controls caused less CS than whole-body controls. This implies the mechanism and extent of control in VR can alter CS susceptibility.

The transition method within VR also plays a critical role; teleportation is identified as less inducing of CS compared to other techniques, with studies [42, 142, 143, 144] noting its

effectiveness in reducing CS over naturalistic movements by eliminating visual transitions and thus sensory conflict. Nevertheless, walking-based locomotion methods are preferred for their realism when physical exertion is not a concern [145], indicating a balance between user comfort and immersion needs to be managed in VR design.

Field of view FOV significantly impacts CS in VR, representing how much of the visual field is encompassed by the VR headset's display [146, 147]. A wider FOV enhances immersion by making more of the virtual environment visible to the user. Adjustability in HMDs allows for varying FOV to user preference. Research indicates that decreasing FOV can mitigate CS by lessening the visual-physical FOV disparity [148, 149, 150, 151, 152, 153]. However, a too-narrow FOV can diminish immersion and provoke more head movements, potentially worsening CS [154]. Therefore, optimizing FOV is essential to balance minimizing CS while preserving immersion in VR experiences.

Type of the content Research has explored how different virtual environment types affect CS [155, 156, 157]. Environments with action content, fast movements, and abrupt changes tend to trigger more CS symptoms than static or neutral scenes [158]. Interestingly, enthusiasts of adrenaline sports are generally less affected by CS. The genre significantly influences CS, with horror environments causing more discomfort than pleasant ones [156]. The user's perspective within VR also matters; third-person views are less likely to cause CS than first-person perspectives [159]. Furthermore, the virtual terrain's geometry impacts CS levels, with irregular terrains and bumps increasing discomfort compared to flat surfaces [157]. Weather conditions in VR may also modify CS severity [160]. These insights are valuable for VR developers aiming to design environments that minimize CS, though more research is needed to understand the precise relationships between different virtual settings and CS.

Duration Recent research underscores the critical role of exposure duration in CS severity. Studies by Risi and Palmisano [113], Aldaba and Moussavi [161], and others have consistently shown that longer VR sessions, particularly with HMDs and during tasks requiring navigation or motion control, significantly elevate CS risk. Cao et al. [162] found higher CS levels from extended use of HMDs compared to desktop displays in VR driving simulations. Similarly, Treleaven et al. [109] noted increased CS with longer exposure during neck motion-controlled VR. Porcino et al. [163] utilized ML to confirm prolonged exposure

as a CS factor in VR gaming, while Zhang et al. [112] observed escalating CS with more extended viewing of live-action 180-degree videos. These findings collectively highlight the necessity of managing exposure times in VR settings to mitigate CS risk.

Visual complexity Visual complexity in VR content, or the quantity of visual information presented to users, has been identified as a contributing factor to CS. Research indicates that higher visual complexity elevates the risk of CS. A study by Kuosmanen et al. [164] demonstrated that participants exposed to VR content with high visual complexity experienced more severe CS symptoms than those exposed to content with low visual complexity. Pouke et al. [165] also found a positive correlation between visual complexity and CS symptoms. These results suggest that minimizing VR content's visual complexity could lessen both the occurrence and intensity of CS, achievable through the simplification of visual elements or optimization of graphics settings to present less visual information to the user.

2.3.3 Measurement of CS

The VR industry's growth necessitates measuring CS to enhance user experiences and identify at-risk individuals. Measuring CS is crucial for developing more comfortable VR environments, understanding its prevalence, severity, and guiding research on its causes, effects, and mitigation. With VR's expanding role in therapeutic areas like pain management, phobia treatment, and post-traumatic stress disorder intervention, assessing CS ensures patient safety and therapy efficacy. Recent studies have emphasized the importance of measuring CS, leading to the development of the CS Evaluation System (CES). CES utilizes inputs ranging from questionnaires to physiological signals like heart rate (HR) and brain activity, processing these to evaluate a user's CS state. The discussion on CS measurement encompasses various methodologies based on a comprehensive literature analysis.

Inputs of CS evaluation

In computer science, the principle of "garbage in, garbage out" (GIGO) highlights the dependency of output quality on input quality; incorrect data input leads to undesired outputs. Our survey classified inputs for quantifying CS levels into four categories: questionnaire, physiological, behavioral, and content types.

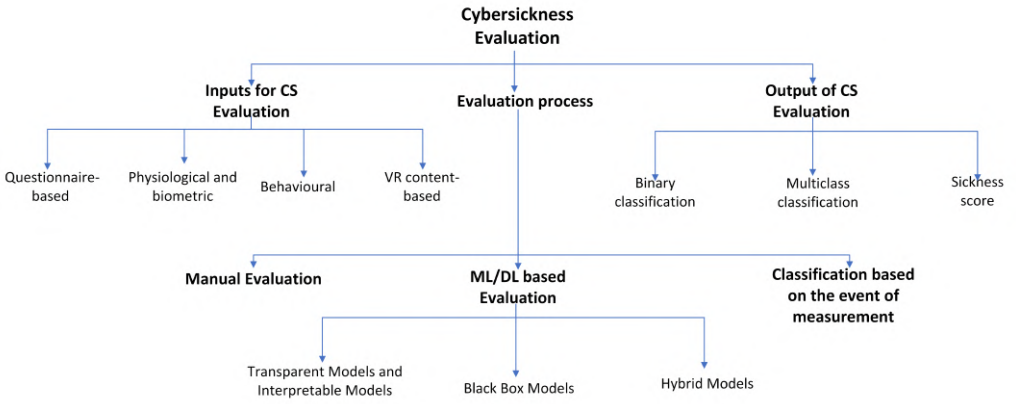


Figure 2.3: Taxonomy of various measurement techniques of CS

Questionnaire-based input Subjective measurement of CS utilizes self-report questionnaires like the Simulator Sickness Questionnaire (SSQ) to rate discomfort and disorientation. Administered before, during, or after VR experiences, these questionnaires are crucial for assessing CS, allowing for the analysis and generalization of results. The SSQ, divided into oculomotor, disorientation, and nausea subscales, is particularly noted for its standardized approach, comprehensive symptom assessment, ease of administration, and quantifiable results, making it the primary tool for CS evaluation [166]. However, the SSQ's reliance on self-reporting and its origin in flight simulation pose challenges, including potential bias and a lack of VR specificity. To overcome these limitations, modifications and alternative questionnaires have been developed. Bouchard et al. [167] proposed adjustments for better VR suitability, while Kim et al. [168] introduced the VRSQ, focusing on efficiency and VR relevance. Other adaptations, like the CSQ [169], the Misery Scale (MISC) [170, 171, 99], and the Fast Motion Sickness Scale (FMS) [172, 173, 174], address specific VR conditions or offer improved psychometric properties. Additionally, tools like the Motion History Questionnaire (MHQ) [175, 176], Motion Sickness Susceptibility Questionnaire (MSSQ) [177, 178], Visually Induced Motion Sickness Susceptibility Questionnaire (VIMSSQ) [179], Nausea Scale [180, 181, 182], and the Motion Sickness Assessment Questionnaire (MSAQ) [183] further enrich CS assessment by measuring susceptibility and overall discomfort. These diverse tools highlight the importance of using a combination of measures for a comprehensive understanding of CS's effects in VR.

Physiological and biometric input Subjective measurement of CS heavily relies on communication from users, which may not always be reliable.However, advancements in technology and the availability of sensors have enabled the use of physiological and biometric inputs to detect CS. Inputs like HR, galvanic skin response (GSR), and EEG, show significant changes during CS episodes [184, 185, 186, 187]. These objective measures, combined with ML and DL models, offer non-verbal ways to detect CS. Event-related potentials (ERP) are also impacted by CS, suggesting ERP analysis could indicate CS levels [188]. EEG, particularly, is extensively used to detect CS, with research indicating specific brain pattern changes and network disconnections post-CS exposure [127]. DL applied to EEG data enhances CS detection [189], analyzing power spectrum and waveform. Heart rate variability (HRV) from electrocardiograms (ECG) [190], along with other signals like electrogastrogram (EGG), electrooculography (EOG), and photoplethysmography (PPG), contributes to CS quantification. Variables derived from these signals, such as EEG power bands [191, 192, 193, 187, 194, 189, 195, 196, 197], EOG potentials [193, 198, 184], and GSR [198, 184, 185, 199, 200], are used to quantify CS. Despite some inconsistencies, the diversity of physiological and biometric data presents new directions for CS detection and understanding, highlighting the need for further research to refine prevention and intervention methods.

Behavioural input Behavioral inputs like postural sway are key in assessing CS, with axial movements or center of pressure (COP) changes during VR exposure serving as indicators [201, 99, 202, 203, 204]. Studies have linked these metrics with CS prediction [205, 206, 207] and found positive correlations with subjective self-reports [208], affirming the relationship between postural instability and CS severity [116, 209, 210, 211, 212]. Eye-related factors, including eyeblinks [193, 198, 184, 213], distance heterophoria [214], eye position changes [215], and vergence/accommodative responses [216], are also studied as CS indicators. Jeong et al. [217] and Chang et al. [218] explored eye-tracking data's predictive value for CS, with Wang et al. [219] and Lopes et al. [220] investigating eye movement and blink rate for real-time CS detection. Additionally, weight shifts [193, 221] and head/waist movements [193, 222] from body sensors have been analyzed as potential CS measures, underscoring the significance of blink rate, head/body movement, and postural sway in CS evaluation.

VR content-based input The complexity of visual environments in VR is known to influence CS. Studies have utilized VR content features for CS assessment. Padmanaban et al. [223] demonstrated a correlation between vection, CS, and relative motion depth, employing FlowNet [224], a convolutional neural network-based optical flow algorithm, to analyze motion depth and predict CS. Optical flow and depth cues, critical for self-motion perception and presence in VR, can trigger CS when altered, serving as potential indicators for CS episodes. Identifying conflicts in these cues can aid in pinpointing CS contributors, allowing for design and technological mitigations. Kim et al. [225] highlighted exceptional motion velocity in VR content as a significant CS factor, using a deep convolutional autoencoder network in their research to measure CS.

Evaluation Process Upon receipt of the input, whether objective or subjective, it is imperative to process or interpret the data to assess CS. This evaluation process can be broadly categorized into two distinct methods - manual evaluation and ML-based evaluation. The former entails human intervention and cognitive processing, while the latter relies on automated algorithms and artificial intelligence. The choice of evaluation method is typically dependent on the specific needs and context of the analysis.

Manual Evaluation This category encompasses the evaluation of CS through the intervention of human cognition and processing. Typically, any assessment conducted by means of questionnaires, such as the SSQ, VRSQ, falls under this category [226, 227, 167, 228, 229, 179, 168].

Machine-learning Based Evaluation The categorization of ML models based on their explainability has garnered significant attention in recent years. The classification framework includes four categories - Transparent Models, Interpretable Models, Black Box Models, and Hybrid Models. All types of models come with their own advantages and disadvantages. It is important to consider various aspects of the models while selecting the appropriate model for a particular application.

Transparent models and Interpretable models Transparent models offer insights into their decision-making process, making their inner workings easily understandable, while interpretable models, although similar, allow for greater customization and detail in under-

standing input-output relationships. Objective inputs like EEG, motion-to-photon latency, and postural stability have been explored for predicting CS onset, with many researchers [230, 231, 232, 233, 234, 235, 236, 237, 238] employing transparent model approaches to analyze these inputs' relationship with CS. For example, changes in EEG power spectra in theta (4-8 Hz) and alpha (8-13 Hz) bands are associated with CS [239]. Dennison et al. [184] used stepwise regression analysis with physiological measures to predict CS. Interpretable ML algorithms like CN2 rule induction [192], Bagged Decision tree [193], Naive Bayes [195], KNN [222], LDA [193], and LSTM regression analysis [199] have been applied to measure CS. Kundu et al. [240] combined ML techniques, including the Explainable Boosting Machine (EBM), decision tree, and logistic regression, to elucidate CS from bio-physiological and subjective data. Interpretable models enhance model trust by providing clear explanations of their predictions, thus improving transparency and accountability in ML systems.

Black Box Models Black box models, characterized by their opaque decision-making processes, utilize ML and DL algorithms for CS detection, analyzing both objective and subjective data. Jin et al. [241] tested Long short-term memory (LSTM), Convolutional Neural Network (CNN), and Support Vector Regression (SVM) classifiers for CS discomfort levels, finding LSTM most effective. Lee et al. [242] enhanced CS detection accuracy by integrating optical flow, disparity, and saliency features using a 3D-CNN and a multi-modal deep fusion method, outperforming previous methods. Jin et al. [241] and Kundu et al. [243] respectively utilized CNN and LSTM for cognitive state estimation from brain signals and Kalman filtering techniques with LSTM for CS, highlighting the effectiveness of black box models. Despite their accuracy, these models demand extensive data and computational resources, and their complexity makes them difficult to interpret.

Hybrid Models Hybrid models merge multiple ML models to enhance accuracy and decision-making transparency. Lee and Alamaniotis [244] introduced a Deep Embedded Self-Organizing Map (DESOM) combined with an EEGNET-based Auto-encoder, outperforming KNN and SOM with high purity (0.96875) and NMI (0.42561). Li et al. [222] utilized a voting classifier incorporating Random Forest, KNN Classifier, Logistic Regression, and Multilayer Perceptron Neural Network, achieving notable accuracy (0.911) and kappa (0.80) for CS classification among 20 subjects, with binary classification accuracy at 0.763 and three-level classification sensitivity values of 0.791, 0.504, 0.867, and kappa of 0.51.

Hybrid models blend black box and transparent model benefits, offering predictive accuracy alongside interpretability—crucial in fields like medical diagnosis. They provide design and optimization flexibility by combining different models' strengths, such as using deep neural networks for feature extraction and decision trees for classification. This approach enables efficient resource use and quicker training. Hybrid models stand as a robust tool in ML, facilitating the integration of diverse model advantages while addressing their limitations.

Classification based on the event of measurement CS detection involves identifying data linked to an ongoing CS event, while prediction focuses on forecasting future CS based on preceding data. Most studies [184, 192, 193, 195, 196, 197] have concentrated on CS detection using physiological signals. However, a smaller number of studies [184, 185, 199] have aimed at predicting CS with physiological signals. Additionally, research utilizing video features for CS prediction includes works by [245, 246, 247, 248, 249, 225, 250, 237, 251, 252]. There's a noted gap in studies predicting CS incidents through pre-VR immersion baseline measurements and in ML research predicting CS susceptibility or imminent episodes using an individual's typical physiological state.

Output of CS evaluation The output from CS self-report questionnaires is scores or ratings reflecting users' subjective symptoms like dizziness, nausea, and disorientation [167, 168, 179, 190, 229, 227, 228], used to gauge CS severity. ML-based CS detection and prediction models provide outputs as binary or multiclass classification labels or a calculated sickness score from physiological signals or other data. These outputs indicate CS likelihood or severity based on model inputs. For example, Islam et al. [253] categorized sickness levels as Low, Moderate, or Acute, while Khoirunnisaa et al. [196] distinguished between sickness presence or absence. Dennison et al. [184] used a regression model to assign a quantifiable sickness score, measuring CS severity experienced by users.

2.3.4 Work on CS mitigation

CS poses significant challenges within VR, necessitating efforts to reduce its effects to enhance user experience, safeguard health and safety, and boost VR adoption for the industry's growth. Mitigation requires actions from both users and developers, although strategies often trade off CS reduction with the sense of presence in VR. This review consolidates

various CS mitigation strategies from existing literature, addressing both developer and user perspectives, aiming to develop effective solutions without detracting from the VR experience.

Mitigation strategies from developers' perspective

Developers have devised multiple strategies to mitigate CS in VR, ranging from virtual environment modifications to hardware adjustments. Altering the FOV is a proven method to reduce CS by limiting peripheral optical flow [148, 149, 150, 151, 152, 153, 254, 255, 256], though it may affect the sense of presence and task performance [257, 258]. Dynamic FOV adjustments, utilized in commercial VR games [259, 150], and foveated FOV restrictors [260] aim to balance immersion and nausea reduction. Adaptive FOV restrictors [261], side restrictors [149], and ground-visible restrictors [262] offer innovative approaches to minimize CS while maintaining user immersion.

Blur effects like depth of field (DOF) and peripheral blur [263, 264, 265] have been explored to reduce CS by simulating natural visual focus, with dynamic DOF blur [266, 267] showing effectiveness in alleviating CS symptoms. "Rest frames" such as clouds [268], grids [269, 270, 271], and a virtual nose [272, 273] have also been used to reduce CS with varying implementations like wireframe models [274], reticles [275], and dynamic rest frames [276] showing potential in CS mitigation.

Optical flow manipulation [277, 278, 279, 280, 281] and hardware alterations, including sparse peripheral displays [282, 283] and novel camera setups [284, 285], address CS by enhancing visual congruency or reducing motion blur. Galvanic vestibular stimulation (GVS) [286, 287, 288, 289], haptic cues [72, 290, 291, 40, 292], and acceleration matching [293] are physical interventions aimed at aligning sensory inputs to reduce CS.

Different locomotion techniques, like teleportation and dynamic speed adjustments [294, 295, 296], have been proposed to lessen CS by minimizing visual-vestibular conflicts. Novel VR navigation solutions, such as invisible ramps for virtual stairs [297], aim to improve user comfort. Additionally, innovative VR camera controls [298] and rendering techniques [299] like Dynamic Mono-Stereoscopic Rendering System (DMSRS) have been explored to diminish CS while preserving user experience. These diverse approaches reflect ongoing efforts to address CS from multiple angles, highlighting the complexity of developing effective

CS mitigation strategies in VR.

Mitigation strategies from users' perspective

To mitigate CS, user-initiated countermeasures are crucial due to CS's subjective nature and varying susceptibility. Limiting VR exposure to $55 - 70$ minutes can prevent severe symptoms [300], and repeated short to longer exposures can foster habituation, reducing symptom intensity over time [301]. Slow diaphragmatic breathing has been shown to alleviate CS by enhancing the parasympathetic nervous system's tone [302, 43]. Acustimulation presents a nonintrusive, cost-effective potential solution [303]. Over-the-counter anti-nausea medications, used before VR exposure, can effectively combat CS, and alcohol at specific concentrations has been shown to reduce CS symptoms, though further research is needed for safety validation [304].

Home remedies like chewing gum and ginger root offer accessible CS relief [305, 306], despite some conflicting studies [307]. Olfactory stimuli, such as peppermint aroma, have been effective in reducing CS, albeit with limitations regarding nausea [308]. Pleasant music also reduces CS, offering a low-cost, easy-to-use intervention [174]. External airflow, suggesting a tactile "white noise" effect, has been shown to lessen CS symptoms without specialized equipment [309]. Positive framing could mitigate nocebo effects associated with CS, advocating for further research on its broader applications [310]. Training to enhance visuospatial skills may also diminish CS effects [311], highlighting a multifaceted approach to managing CS in VR users.

2.3.5 Research gaps in CS

Despite being a well-known issue in the VR community, there are still gaps in the research on CS. After critically analyzing a carefully curated list of 223 highly relevant research papers on various aspects of CS, we have identified several research gaps in the field (table 2.2). Addressing these research gaps will enable us to advance our understanding of CS and develop effective countermeasures for it.

Research topic	Title	Remarks
Causes of CS	Role of the physical environment in Inducing CS	There is a lack of research on how physical environment elements like room size, temperature, ventilation, and lighting plays a role in inducing CS. Knowing how these elements play a role in inducing CS will enable us to come up with strategies to mitigate CS by altering the physical environment elements accordingly.
	Role of ethnicity in causing CS	Although some studies [102] have looked into how ethnicity affects CS, there is still a requirement for more study. Ethnicity may be able to explain some degree of subjectivity associated with CS.
	Role of gender in inducing CS	Although some studies [107, 312, 313] found that gender plays a role in inducing CS, there have been some contradictory studies suggesting that gender does not play a crucial role in causing CS. Further research is required on this front. Also, to remove any gender biases in the study, there is a need to conduct gender-neutral studies.
	Role of visual disability in inducing CS	It is reasonable to assume that users of VR technology with impaired binocular function caused by issues with convergence would experience greater oculomotor side effects than those with normal vision. Further research is required to establish vision impairment as a potential factor in causing CS.
	Role of the genre of VR content in inducing CS	There is a lack of uniformity in genre classification systems and the need for more qualitative research. More research is needed to understand the relationship between the genre of the content and CS.
Measurement of CS	Lack of standardized metric	There is a lack of standardized metrics present in the literature that can objectively compare the effectiveness of different measurement techniques. The lack of a standardized metric makes it difficult to compare different studies.

Continued on next page

Table 2.2 – continued from previous page

Research topic	Title	Remarks
	Unavailability of large, quality datasets	Lack of data in large quantity is a bottleneck in measuring CS accurately. Without acquiring data in larger quantities, we would not be able to leverage cutting-edge DL methods to measure CS. Also, a lack of data will be a roadblock in developing a model that is able to generalize well.
	Lack of explainability of ML/DL models	Although Kundu et al. [240] proposed predicting and detecting CS using explainable ML (xML) models, there is a lack of research regarding using xML to measure CS. The trustworthiness of ML/DL models can be greatly increased by having inherent interpretability, which can give an understanding as to why a certain conclusion was reached by the model. This will also help researchers understand which factors contributed more towards CS, and accordingly, mitigation strategies can be proposed.
	Lack of predictive models for CS	Majority of the study we reviewed used models to detect CS from physiological signals. We observed that there is a scarcity of studies in the literature concerning the prediction of CS using physiological signals. Predicting CS symptoms before it actually happens would allow both the developers as well as the users to take adequate preventive measures.
	Investigation of various eye-tracking data as features to measure CS	As there is a number of eye-tracking features available, e.g., pupil diameter, pupil position, gaze direction, gaze origin, eye openness, etc., there is a need to study how different combination of eye features can be used to measure CS. Further studies are required to understand the significance of each component of the eye-tracking data.

Continued on next page

Table 2.2 – continued from previous page

Research topic	Title	Remarks
	Using only sensor data of commercially available HMD to measure CS	Research by Islam et al. [253] suggested that it is possible to measure CS using only the data available from sensors present in commercially available HMDs. Further research is required on this front. One suggestion for the future is to integrate controller data along with HMD sensor data to measure CS. Further research can also be done on the real-time prediction of CS by integrating VR video content data along with sensor data.
	Use of incremental (online) learning	ML models would greatly benefit if we incorporate incremental (online) learning, making the models more robust.
	Prediction of emotional state using only sensor data of commercially available HMD to anticipate the induction of CS	We found in literature that the emotional state of a person is co-related to the induction of CS [105, 119, 120]. However, we didnot found any work that recognise emotion from the the sensors readily available in a consumer-grade HMD.
Mitigation of CS	Lack of CS-specific medication	Over-the-counter medications often prescribed for CS, are primarily for MS. Since there is a significant difference between MS and CS, it is suggested to develop CS-specific medications. To test the efficacy of novel medications and their delivery methods, further investigations are required.
	Controlled breathing as mitigation technique of CS	Controlled breathing has been proven to be helpful in reducing CS symptoms [302, 43]. More research is required to understand how controlled breathing helps in the reduction of CS and further optimize the technique to increase its effectiveness.
	Role of soothing environment elements	More research is required on the usage of music and pleasant aroma as mitigation techniques of CS.

Continued on next page

Table 2.2 – continued from previous page

Research topic	Title	Remarks
	Adaptive mitigation based on the severity and the cause of the sickness	Mitigation techniques of CS cannot be generalised. Applying the proper CS mitigation strategy with required level of intensity by identifying the exact reason and the intensity of CS can provide more specific solution to CS. More research needs to be done in this direction.

Table 2.2: Research gaps

2.3.6 Guidelines for mitigating CS

Most Immersive VR experiences come with some sort of undesired symptoms of CS. In a study by Regan and Price, [314], more than 60% of the 146 participants reported symptoms of CS at some point throughout a twenty-minute immersion and ten-minute post-immersion phase. This statistic suggests that it's highly likely that a user will experience some sort of undesired symptoms of CS during longer exposure in the virtual environment. Researchers have still not found a way to stop CS completely. In fact, as of now, we are not even close to completely uprooting the problem of CS. Hence, it's extremely important to reduce the impact of CS as much as possible. Otherwise, users will not adopt VR technology, owing to the terrible user experience due to CS. Hence, there is a need for preventive measures to tackle the problem of CS. It's crucial for the developers to adhere to a few rules when producing VR content, as well as for the users when consuming that VR content. After a thorough review of the literature, we are proposing a set of mitigation guidelines for both the developers as well as the users. These set of guidelines, when followed, will be able to reduce the impact of CS, as backed by several studies found across the literature.

Developer guidelines

1. Developers should allow for adjustable IPD in HMDs as studies show smaller IPD in females lead to poor HMD fit and increased susceptibility to CS. [107].
2. Reduce application to photon latency, also known as motion-to-photon delay (MPD).
3. Repetitive patterns (like gratings) and high spatial frequency content (like stripes,

tiny textures, etc.) should be minimized or avoided as they might cause discomfort and feelings of vection. Flatter textures should be preferred to patterned ones, such as those that are solid colours.

4. Limit or eliminate the usage of elements that create vertical acceleration (such as stairs), as horizontal optical flow when climbing (vection) and vertical acceleration can both cause symptoms. Ramps might be a better alternative, but they should only be utilised occasionally.

5. To lessen the symptoms of CS, rest-frame graphical cues can be displayed in the foreground as well as panorama elements in the background.

6. Implied acceleration and deceleration in the optic flow should be kept to a minimum.

7. As hyperrealistic graphic is known to induce CS, developers should provide the functionality of dynamically downgrading the rendering quality of the VR content in order to mitigate CS.

8. Depth cues that try to mimic real-world cues can be used to mitigate CS. For user tasks, blurring less significant areas of the image may also be helpful in reducing the impact of CS.

9. Maintain a stable frame rate for smooth movement in VR, which can reduce the risk of CS.

10. Provide comfort settings in the VR application to give users the ability to adjust the intensity of certain effects, such as motion blur, and subsequently minimize the risk of CS.

11. Inform the user if the VR content will invoke CS symptoms or not. This way, users can make a choice to participate and adequately prepare themselves for the VR experience.

12. Notify the VR user about their exposure time in the virtual environment and accordingly give them suggestions about avoiding CS symptoms.

13. Design shorter duration VR content so that users can enjoy them in one sitting and reduce the risk of developing severe CS symptoms.

User guidelines

1. User should take short breaks in between VR sessions. A single VR session should not be very long (typically $15 - 20$ minutes).

2. New users of VR technology are advised to start slowly and gradually build up their exposure. They should begin with shorter sessions and gradually increase the time spent in VR.

3. While trying out VR technology for the first few times, it is advised to have someone by their side so that in case of emergency, they can take the necessary steps.

4. Find a comfortable HMD that fits the user properly and maintains proper hygiene of the HMD.

5. Avoid sudden, rapid head movements while using VR as that can cause discomfort and contribute to CS.

6. Users should avoid indulging in VR technology if they are carrying some physical illness.

7. Users should take preventive medicines, like anti-nausea and anti-vertigo medicines, when they start to feel symptoms of CS. These medicines take time to work. Hence, preventive measures should be taken as soon as the user feels uncomfortable in the virtual environment.

8. Certain VR content may induce CS symptoms in only a specific group of VR users. As CS is extremely subjective, it's important to understand your own preferences and accordingly decide whether to avoid or indulge in certain VR content.

2.4 Off-screen POI visualization for enhancing Wayfinding in XR

XR navigation, encompassing VR, AR, and MR, employs a variety of innovative wayfinding techniques and strategies to enhance user experience in navigating both virtual and augmented spaces. One fundamental approach is the use of visual cues, such as arrows, paths, or

highlighted objects, which guide users towards their destinations by drawing on the intuitive understanding of physical world navigation [47, 315]. Auditory cues [316, 317], including spatial audio and verbal instructions, complement visual aids by providing directional guidance and contextual information through sounds that appear to emanate from specific locations in the user's environment. Haptic feedback, through wearable devices like gloves or vests, offers another layer of navigational aid by simulating the sensation of touch or resistance, thereby providing intuitive cues for direction or the presence of virtual obstacles [318].

Another sophisticated strategy involves adaptive interfaces that dynamically adjust to the user's needs, preferences, or physical movements, thus offering personalized navigation assistance. For instance, a system might display guidance based on the user's current focus of attention or the complexity of the environment [319]. Cognitive mapping is encouraged through environmental landmarks and spatial layouts designed to be memorable and easily interpreted, aiding users in forming mental maps of the virtual space for better orientation and recall [320]. Additionally, collaborative navigation techniques in multiplayer XR environments enable users to guide each other, sharing waypoints or paths, thereby leveraging social interaction as a means of wayfinding [321, 322].

A persistent challenge in XR wayfinding is managing off-screen POIs, which are crucial destinations or objects that lie outside the user's current field of view. Addressing this challenge enhances wayfinding significantly by ensuring that users are aware of important locations or items that are not immediately visible [323]. Techniques such as off-screen indicators, which can take the form of arrows or icons on the display, gently guide the user's attention towards these POIs by suggesting the direction in which to turn or move. Furthermore, augmented reality can overlay contextual information or virtual pathways onto the real world, seamlessly integrating off-screen POIs into the user's navigational context. This integration not only solves the problem of off-screen navigation but also enriches the user's spatial understanding and engagement with the environment, making wayfinding in XR a more intuitive and effective experience [324]. However, the use of excessive on-screen cues or large-sized visualizations on handheld devices can lead to clutter and make it challenging for users to distinguish between different POIs [325, 44].

In the literature, several approaches have been suggested for displaying huge information

spaces on the constrained screens of mobile devices. When possible, the information space is restructured using content analysis to create sections of relevant content that fit on the display screen. Techniques to present information on devices that exist beyond the scope of the device's screen are classified into Focus+Context (F+C), Overview+Detail (O+D), and contextual view techniques [326]. These are the representative techniques for navigating in a place that does not fit entirely on the screen. In O+D techniques, an overview is present in a separate view with a detailed view showing the user the currently focused object(s). The overview may be a top-down view of the whole workspace attached to the detail view. The detailed and overview views may be tightly coupled together. However, To absorb the two views, users must continually stare at them. One of the examples of O+D visualization is miniature map visualization [327]. F+C approaches simultaneously display the information space at various levels of detail without dividing it into various viewpoints but distorting the view of the environment. For instance, fisheye views [328] presents both the detailed map (Focus) and the wide-area map (Context), distorting the information landscape. The contextual view makes use of proxies to simulate relevant data (the off-screen objects) in the environment. One example of a contextual view technique is visualizing off-screen objects with a wedge [44]. These types of visualization techniques do not distort the information landscape and do not demand extra mental effort like O+D techniques. There is no single answer to which method is better for off-screen POI visualization as it depends on various factors, such as the user's task, the specific AR application, and the user's preference. However, the contextual view method is helpful for small screens because it saves space while allowing us to see the bigger picture. It can also be combined with other methods [34, 329].

The 2D counterpart of the problem inspires many works in AR off-screen POI visualization. However, using proxies in 2D screen space comes with limitations due to their anchorage in a device's 2D screen space, limiting the discrimination to front-rear and above-below. Extra visual or auditory cues must be used to determine an object's 3D distance or enable front-or-rear discrimination. For instance, textual annotations were added to the arrows in SidebARs [37] handheld AR visualization to indicate the distance to the items that were out of view. Baudisch et al. [45] proposed Halo2D, which visualizes the off-screen POIs via arcs of circles centred at POIs with radii slightly greater than the

distance between the POI and the closet border of the screen. The point where the arc of the circle comes together shows which way to go towards the POI. Gustafson et al. [44] introduced an improvement over Halo2D by using isosceles triangles whose one end lies on the POI and the other two ends inside the device's screen space. The triangles, termed as Wedges, can re-orient themselves to reduce cluttering. However, proposed 2D visualizations do not effectively communicate the location of things in 3D contexts [330]. While these techniques have shown improvements, they still have limitations in providing information about front-or-rear and above-or-below discrimination.

2D visualizations are not always adequate to represent the target's 3D position and distance. To address this issue, many researchers have introduced out-of-view object visualizations that rely on 3D proxies, and these visualizations have been shown to have several advantages. Shapes such as 3D versions of arcs [331], arrows [315, 332], and wedges [48] have been commonly studied as proxies for off-screen objects. In the early days of AR, Feiner et al. [333] used rubber band-like visualization to guide the user to an off-screen POI while highlighting the object itself. Attention funnel [334] uses a series of rectangles beginning at the user and ending at the object to guide the user. These approaches suffer because the visualization cues are relatively big to the screen size and can overlap, which leads to screen clutter. Halo3D [10], an extension to Halo2D, first projects the POI existing in 3D space to the device's plane and uses the distance of the projected point to the closest border to draw an arc protruding into the screen. The projection technique does not allow front or rear disambiguation of POIs at all, as POIs may project onto the same screen space point even if they differ in their directions, resulting in the same visualizations. Halo3D [331] only allows the user to get a vague sense of the POI's direction in the screen plane. The user may get some idea of the direction using the distance cue (arc length), but that seems cognitively tricky. Moreover, in high-density POI regions, there is an overlap of arcs which causes cluttering. Aggregation of arcs is used, but that fails for even greater high-density environments [44]. Difeng et al. [48] developed an off-screen POI visualization technique for HMD users called 3Dwedge+. They found this technique performs better in distance and direction awareness compared to some O+D techniques, namely, Radar and 3D Minimap. Another technique called EyeSee360 [47] was proposed specifically for head-mounted AR devices. It uses two concentric ellipses, where an inner ellipse and the boundary by the

Techniques	Research gaps
2D metaphors [37, 44, 45]	Unable to provide front-or-rear and above-or-below discrimination
	Hard to represent the 3D position and distance of a target
AroundPlot [50]	Low direction awareness
	No distance information is provided
Halo 3D [331]	Front or rear discrimination is not clear
	Creates visual clutter with a higher number of POIs
3D Arrows [32, 315, 332]	Interface becomes hard to read in large number of POIs
	Occludes the central part of the screen
EyeSee360 [47], Wedge 3D+ [48], CompassbAR [49]	Designed for HMD based AR system
	Can reduce direction and distance awareness when the number of POI is more
3D Bezier Curve [36]	Used for visualizing single POI. Will cause clutter when the number of POI increases.

Table 2.3: Research gaps in the prior work

outer ellipse represent the FoV. A mixture of size and colour is used for the representation of distance.

3D Arrows have also been used as a visualization tool. Single 3D arrows have been tested for directing the user's attention while driving vehicles [335, 336]. They have also been employed for multiple objects. 3D arrows, such as those found at [332], enable the display of the POI's direction in reference to the device's centre. The arrows are in the centre of the screen in an egocentric view. Because of using a central point for affixing the arrows in the centre, the interface becomes hard to read in high-density regions as the arrows begin to overlap. Positively, front–back disambiguation, along with verticality indication, is improved. In FlyingArrow [337], the arrow flies from the user's point-of-view towards the off-screen POI and returns a sound signal when it reaches the off-screen POI. The distance is mapped to the flight duration of the arrow. A study by Sathaporn et al. [338] found FlyingARrow with trails to be useful for devices with limited field-of-view. Another study conducted by Jo et al. [50], compared three visualization techniques: 3D arrow clusters, F+C, and O+D. Results showed that the use of 3D arrow clusters yielded a similar performance as the F+C technique, which outperformed the O+D visualization. However, when the number of items increased to 50, participants exhibited lower accuracy with the 3D arrow cluster compared

to the F+C technique. The authors concluded that this decrease in accuracy was mainly due to the occlusion of arrows, leading to confusion. They proposed a system called Around Plot, which is based on the F+C approach. Here, a transparent frame is overlaid on the device's screen, where Off-screen POIs are visualized as dots. An egocentric projection technique is used where the rotational angles needed to get an off-screen POI into view are used to project the POI into the screen's plane. The projected points are simply visualized as small quads. Though the projection method allows the user a vague sense of direction of off-screen POIs, it is not as apparent as using 3D arrows. Furthermore, in this technique, there is no way to show the distance information apart from showing texts, which may cause on-screen clutter. It is evident that techniques with lower visual clutter exhibit significantly better performance than the variant with the highest visual clutter [44, 339]. The techniques mentioned above are employed to present off-screen POIs in AR applications. Most of these techniques demonstrate the direction of POIs either directly or indirectly. However, some techniques, such as the O+D techniques, fail to differentiate between POIs with varying altitudes. Additionally, several methods offer distance perception but can create visual clutter on-screen in environments with high POI densities, which may compromise overall performance. As a result, these visualizations are not ideal for scenarios involving POI visualization while navigating through densely populated indoor/outdoor marketplaces with numerous POIs (i.e., over 15). We have presented the research gaps of the prior works in Table 2.3.

2.5 Summary

In this chapter, we explore VLTs and their significance in enhancing navigation within XR environments. The discussion begins with an overview of the challenges and importance of VLTs in creating realistic walking experiences. We delve into different strategies for delivering a realistic walking experience and highlight the importance of refining this aspect in system-automated VR tours. Furthermore, we identify a research gap in predicting a user's instantaneous walking speed during a system-automated VR tour, emphasizing the need for advancements in this area.

The chapter further delves into the systematic review of CS, addressing its causes, measurement, and mitigation strategies, highlighting the need for a holistic understanding of

this phenomenon to develop effective solutions. Various measurement techniques, including questionnaires and physiological and behavioral inputs, are evaluated for their effectiveness in detecting and predicting CS. Mitigation strategies from both developers' and users' perspectives are presented, offering guidelines for reducing the impact of CS and enhancing the VR experience. We also identified various research gaps in this field and based on our analysis provided several insights for future research.

Additionally, the chapter discusses off-screen POIs visualization in XR to improve wayfinding. Various techniques for displaying information beyond the device's screen are explored, including F+C, O+D, and contextual view techniques. The chapter discusses the gaps in the prior works on off-screen POI visualization in smartphone-based AR, highlighting the limitations of 2D metaphors and the advantages of 3D proxies for better direction and distance awareness. We point out the challenge in presenting the POIs on a smartphone screen when the number of POIs increases.

CHAPTER

3

Enhancing Realistic Walking Experience in System-automated VR Tours

3.1 Introduction

Walking in a VR environment is one of the significant interactions in VR. The speed at which a user travels plays a significant role in their virtual experience. As mentioned earlier, to simulate a realistic walking experience, it is crucial to match the virtual travelling speed with the user's actual walking pace [39]. However, determining the actual walking speed for a system-automated tour can be challenging since such tours do not allow for ongoing user input, making it impossible to adjust the virtual speed based on user movements. Previous studies have utilized a fixed walking speed to simulate movement, but this approach does not account for the natural variability in a person's speed. We propose that simulating the user's actual, instantaneous speed could more effectively enhance realism compared to a uniform speed.

This chapter seeks to answer two main questions:

- How can we predict the instantaneous walking speed of a user with no continuous user input?
- Does the usage of the instantaneous walking speed in a system-automated tour enhance the realistic walking experience than the usage of the constant speed?

In order to address these questions, we present a model for estimating users' instantaneous walking speed while moving from one point to another. This model dynamically computes users' walking speed, knowing their (actual) height and position in the virtual

world from the source and the known destination. We validated the model with the help of an empirical study. We also conducted a user study to compare the perceived realism while using the walking speed predicted by our model with the constant walking speed.

3.2 Proposed model

As discussed before, a person's walking speed depends on the step frequency and step length (that depends on the height) (eq. 2.2). The walking velocity, on the other hand, keeps on changing throughout the travel. To estimate the instantaneous natural walking speed, we first built a model that predicts the actual instantaneous walking speed by predicting the instantaneous step frequency.

3.2.1 Model building

In order to build the model, we conducted an empirical study and gathered some real walking data.

Participants We recruited 40 participants (26 males, 14 females) aged from 19 to 42 years (Mean= 29.825, SD= 6.025), with heights of 150 to 184 cm (Mean= 166.14, SD= 8.18) to take part in our study. The participants were undergraduate and postgraduate students. All of the subjects had normal or corrected vision and were free of physical deformities.

Procedure This study was intended to gather real-life walking data when a user moves from one point to another and propose a speed estimation model. Before starting the experiment, we informed each participant about the procedure. However, we did not expose the purpose of the study to the participants before doing the task. We informed them about it after they finished the task. We did it to reduce the experimental bias. We took a corridor of a building to conduct this experiment. We hung seven paintings (of seven wonders of the world) at $20, 20, 13, 16, 6.4, 19$, and 4 meter distances. The participant's task was to start walking from the starting place and stop at each painting to click a photograph of the painting. The clicking photograph task was assigned to make sure the user stops after walking a predefined distance. The floor plan of the corridor is shown in the Figure 3.1.

To calculate the step frequency, we recorded a video of the user's footstep by fixing a

smartphone(Model: Samsung M30) camera in the back of the user (shown in the Figure 3.2) focusing on their legs. We started recording the video just before the user was asked to start walking from the starting point. The videos were recorded at a 30fps rate.

While performing the task, we made sure that there were no disturbances (other people, noise) in the way. We chose weekends to collect data, as the place we chose was less populated during weekends. If we had found any disturbances while doing the task, we requested them to repeat the task.

We calculated the instantaneous frequency of steps for all the participants. It was calculated by observing the number of steps per second recorded in the video. We considered one-foot strike to another foot strike on the ground as one step [75]. The videos were observed by noting the timestamp for each foot strike. The data for each distance was separately noted. We could detect the end of each distance by observing where the user stopped for some time. Thus, we created seven files (of seven distances) for all 40 users, containing timestamps (in ms) at each step (assuming the time when s/he left the previous painting is zero). As it was a manual observation, we did the process five times and finally took the average timestamp for each foot strike. Once we got the timestamps, we calculated the instantaneous step frequencies at each step.

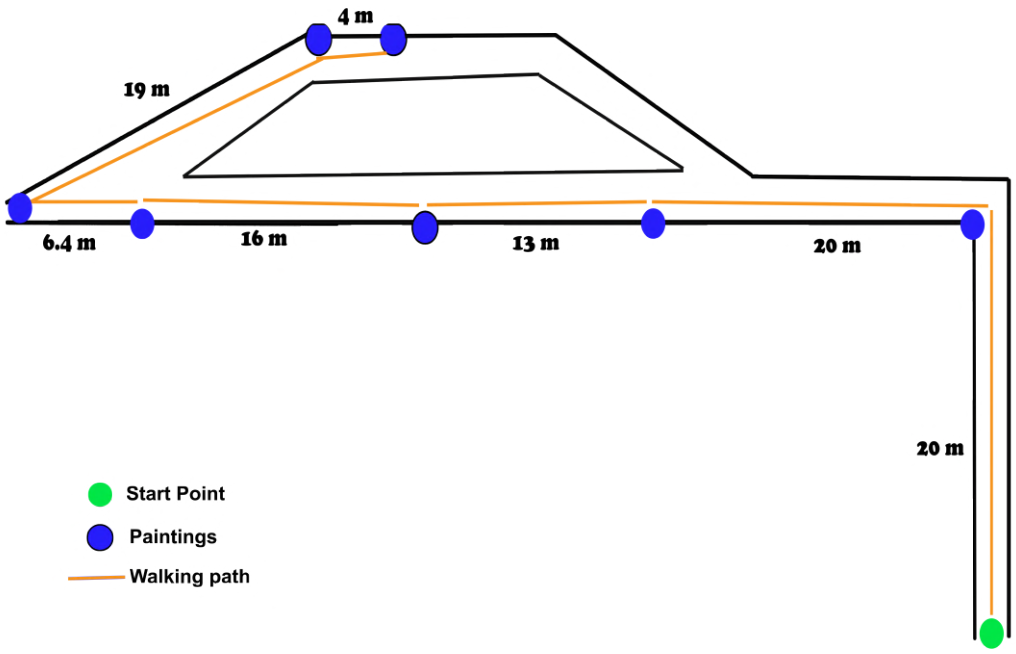


Figure 3.1: Floor plan for the empirical study

Figure 3.2: Smartphone attached in the back of the user to collect instantaneous frequency for building the model

Model formulation based on user data During the experiment, we observed that the user did not maintain constant speed during the walk. Initially, they started slow, but the velocity increased gradually, plateauing out upon reaching the maximum velocity and slowing down just before reaching the target. Later, we observed the recorded videos and found that it was the frequency of steps that varied over distance, affecting the overall pace of walking. We plotted the distance travelled from the source vs frequency(number of steps per second) graph to visualise the data (plot for 20-meter distance. we have shown the plot for three random participants in figure3.3) to clearly visualise the pattern(showing plot for all 40 participants clutters the graph). We observed the plotted data from the recorded video. The frequencies were calculated by measuring the time taken (in ms) to take each step and then multiplied 1000 ms to the inverse of this time.

As observed in this study, the step frequency grows at first to reach maximum and then it remains almost constant and then decreases near to zero. This pattern was also introduced in the work by Bowman et al. [38]. They called it Slow In Slow Out (SISO). However, they evaluated this pattern with rapid movement rather than natural walking speed. Our work intends to express it mathematically to predict the instantaneous natural

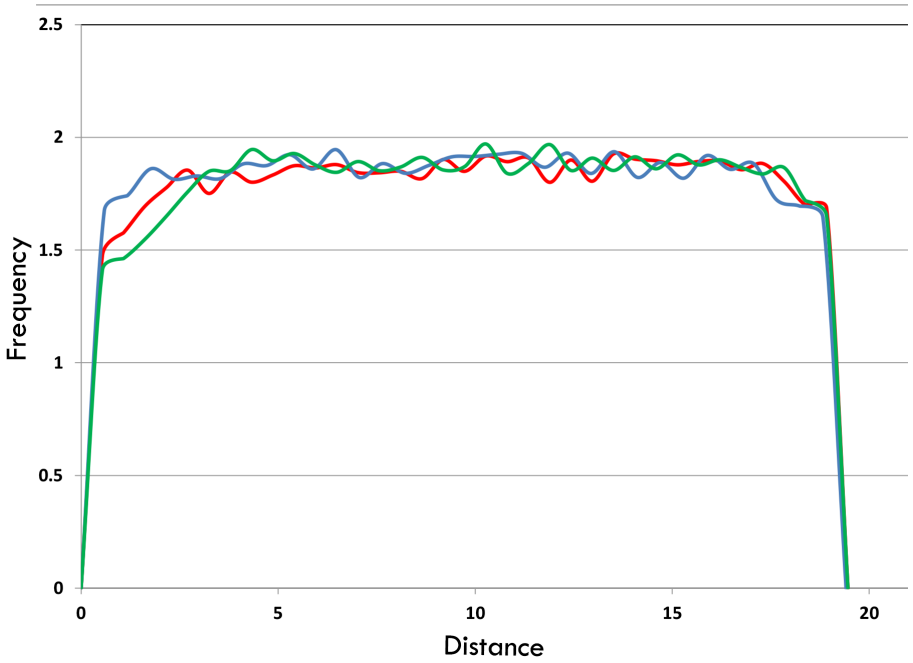


Figure 3.3: Frequency vs distance graph for 3 random participants

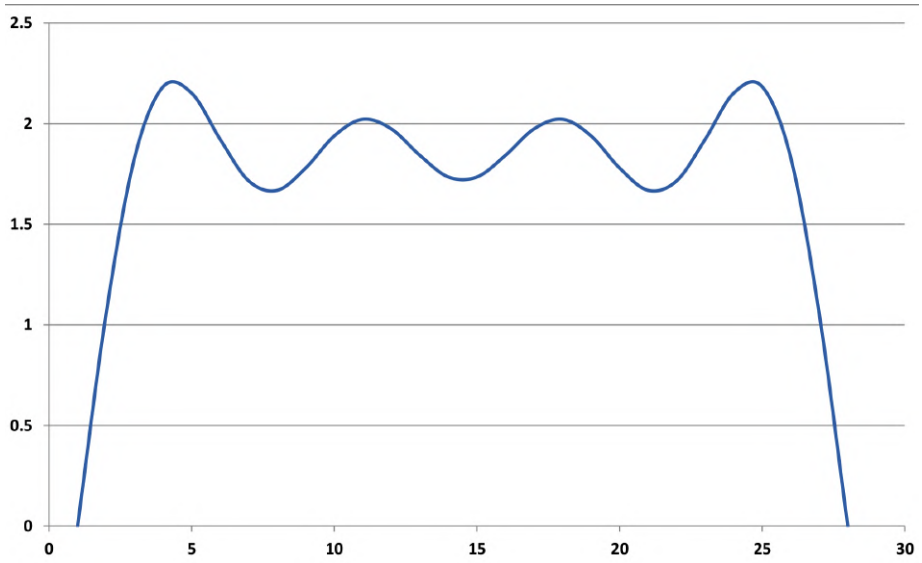


Figure 3.4: Fourier series approximation of square pulse wave

walking speed. This change in frequency pattern is nearly similar to pulse wave widely used in digital signal processing [340]. A pulse wave shows a temporary increase or decrease in the amplitude of a signal from its baseline value. The nature of the data we found from the observation closely resembles a square pulse wave, as shown in Figure 3.4. This type of pulse wave can be well approximated by taking the first few terms of the corresponding Fourier expansion [340]. We can model this with Fourier series analysis based on the idea that Fourier series with proper weights can approximate any continuous periodic function to arbitrary precision. The weights of harmonically related sinusoids can be calculated with the help of the empirical data.

However, researchers [341] argue that there is a difference between the actual speed and the perceived speed. Generally, an individual perceives the visual flow slower than real [341]. Because of this reason, showing the visual with the user's actual walking velocity is insufficient. In order to compute the desired actual visual speed (V_{vir}), we need to multiply a visual gain ($\mathcal{G}$) with the actual speed. So user's instantaneous virtual velocity can be formulated as follows

$$|V_{vir}| = \mathcal{G} \times |V| = \mathcal{G} \times (\frac{f}{0.157} \times \frac{h}{1.72})^2 \tag{3.1}$$

Here, $|V_{vir}|$ is the perceived natural speed where $|V|$ is the user's natural instantaneous speed in the real world and $\mathcal{G}$ is the visual gain. Nilsson et al. [39] defined a visual range for four VLTs: stationary, tap in place (TIP), walk-in place (WIP), and real walking. When utilising stationary, TIP, WIP, and real walking in treadmill locomotion, they obtained visual gains ranging from 1.75 to 2.45, 1.74 to 2.45, 1.58 to 2.40 and 1.65 to 2.44 respectively. Nevertheless, the gain may vary with the different locomotion techniques. In our case, we are using system automated locomotion, which falls under stationary VLT. We can multiply a gain factor of 2.10 (The average of the visual gain range for stationary VLTs found in Nilsson's study) to make the user perceive the speed as the natural walking speed. However, as we noticed the frequency to be changing over time We want to compute instantaneous V_{vir}, which can be expressed as a function of x the distance of the starting node v_1 ($V_{vir}(x)$). Since we have noticed the frequency to be changing over distance, we can modify f in

equation 3.1 as $F(x)$. Finally, the equation 3.1 can be written as followed

$$V_{vir}(x) = 2.10 \times \left(\frac{F(x)}{0.157} \times \frac{h}{1.72}\right)^2 \tag{3.2}$$

As $\mathcal{G}$ and h are constants in order to compute $V_{vir}(x)$ we only need to compute the instantaneous step frequency $F(x)$. Let the distance between two nodes v_1 and v_2 be d. Then $F(0) = F(d) = 0$ as we assume the user started walking from a stationary position and stopped after reaching the destination. Moreover, the function $F(x)$ should be continuous w.r.t x as we assume the user to be walking naturally so there won't be any abrupt changes to step frequency. The continuity of $F(x)$ ensures that $F(x)$ can be approximated to arbitrary precision in the interval $[0, d]$ using fourier series with period $2d$. Moreover, as $F(0) = F(d) = 0$ we can ignore the cosine terms of the fourier series and keep only the sine terms.

Looking at the collected data, we can simplify the Fourier series further by using the Fourier expansion of the square wave as following [340]:

$$F(x) = \frac{4}{\pi} \times F_m \times \sum_{k=0}^{n-1} \left(\frac{1}{2k+1} \sin{(2k+1)\omega x}\right) \tag{3.3}$$

$$Where, \quad \omega = 2\pi f_0, \quad and \quad f_0 = \frac{1}{2d}$$

In Equation 3.3, $F(x)$ is a function that takes the distance from the first point as an input. For instance, a user is in point A, and s/he opted to go to point B. In that case, the step frequency at A and B would be nearly zero. The output of $F(x)$ is the instantaneous step frequency at a distance x meter from point A, which varies from 0 to d. We represent $F(x)$ as a square wave with frequency $f_0 = \frac{1}{2d}$ and amplitude F_m which is equal to the maximum step frequency the user can reach. It is the Fourier expansion of square pulse wave up to n terms. However, we need to establish the value of F_m and n (the number of terms to be taken) empirically. We call this the "Step Frequency Model".

The walking speed is directly dependent on the step frequency and height of the user. Once we get the step frequency $F(x)$, we can calculate the walking speed to be used in the tour from the equation 3.4. Here, $V_{vir}(x)$ is the function that predicts the natural velocity of a user at x meter distance from point A. This function uses the instantaneous

step frequency calculated in equation 3.3 and the height (h in meter) of the user to calculate the instantaneous virtual velocity ($V_{vir}(x)$ in cm/sec). However, the original equation given by Dean et al. [74] calculates the velocity with the average step frequency and height; we modified the term step frequency to instantaneous step frequency.

We have two unknown parameters in the step frequency model (3.3): the maximum frequency F_m and the number of terms n in the Fourier expansion.

3.2.2 Model parameter estimation and validation

We used cross-validation (CV) technique to estimate and test the model parameters. We chose Leave-One-Subject-Out Cross-Validation (LOSOCV) technique for this. In this technique, we first train the model with the first n-1 row data and test it with the one left out row; then, we again train the model with the n-1 row data starting from the second row and test it with the one left out and iterate until each row is used for testing. Because the model is fitted to almost all of the training data (n-1 observations), this cross-validation technique overcomes the disadvantage of utilising small training sets, which is observed in the general validation set approach. We took the data of all 40 participants for parameter estimation. Following the principle of the least-square method, we fitted our step frequency model with the observed frequency. We calculated the mean of the maximum step frequency(F_m) of the forty participants. We found it to be 1.8 steps/sec. We also observed that the equation yields best when the number of terms n is taken almost half of the distance d. The RMSE value and the R-squared fitness mean for each subject are presented in table 3.1. The mean of the RMSE and R^2 value we found for each iteration of CV are 0.1481 and 0.9014, respectively.

After estimating the parameter, we could express the final model for predicting the instantaneous step frequency as:

$$F(x) = \frac{4}{\pi} \times 1.8 \times \sum_{k=0}^{\lceil \frac{d}{2} \rceil - 1} \left(\frac{1}{2k+1} \sin{(2k+1)\omega x} \right) \tag{3.4}$$

Where,

$$\omega = 2\pi f_0, \quad f_0 = \frac{1}{2d}$$

$$d = \textit{distance between source and destination}$$

Subject number out	RMSE value	R squared fitness mean
1	0.901	0.1362
2	0.9008	0.1296
3	0.9009	0.1364
4	0.9007	0.1316
5	0.9007	0.1264
6	0.9011	0.1377
7	0.9011	0.1358
8	0.9010	0.1328
9	0.9010	0.1374
10	0.9010	0.1381
11	0.9010	0.1386
12	0.9011	0.1403
13	0.9011	0.1428
14	0.9012	0.1441
15	0.9013	0.1459
16	0.9013	0.1463
17	0.9014	0.1470
18	0.9014	0.1486
19	0.9015	0.1485
20	0.9016	0.1506
21	0.9016	0.1502
22	0.9015	0.1483
23	0.9015	0.1470
24	0.9016	0.1492
25	0.9016	0.1497
26	0.9017	0.1527
27	0.9017	0.1541
28	0.9017	0.1542
29	0.9018	0.1564
30	0.9017	0.1571
31	0.9017	0.1568
32	0.9018	0.1600
33	0.9018	0.1589
34	0.9018	0.1588
35	0.9018	0.1611
36	0.9018	0.1619
37	0.9018	0.1621
38	0.9018	0.1632
39	0.9018	0.1642
40	0.9018	0.1639

Table 3.1: LOSO CV result

The output of this model can be fed to the equation 3.2 to get the instantaneous walking velocity to be used in the virtual camera motion in the VE.

3.3 Model evaluation

In this section, we address the second question of our chapter, which is about evaluating the realism of instantaneous walking speed as the optic flow in VR. We had to confirm whether the usage of the speed prediction model could enhance the perceived realism than simply using a constant average speed to give a realistic walking experience. The rationale for showing predicted instantaneous speed is that seeing the walking speed that matches the actual walking can enhance the experience. We compared perceived realism for instantaneous walking speed with constant walking speed. We calculated the constant walking speed from the equation proposed by (equation 2.2) Dean et al. [74], and we compared it with the speed predicted by our model. We compared the speeds with the help of an empirical study.

Participants We gathered a group of 34 participants (20 men and 14 women) ranging in age from 15 to 47 years old (Mean: 29.324, SD: 6.70) and height between 159 and 170 cm (Mean: 167.97, SD: 7.40). The participants were among the undergraduate and postgraduate students from engineering backgrounds. Eighteen of them had a prior experience in VR (Twelve of them took part in the previous experiments held in our lab). All the participants had normal or corrected to normal vision with no walking disabilities.

Procedure We used Oculus Rift Head Mounted Display (HMD) to show the environment to the participants. We tracked the position and rotations of the head through the Constellation tracking system of Oculus Rift. We developed the application in Unity 3D (2020.3.21f1 release). The software ran on a 64-bit Windows 10 Professional computer with a 3.6 GHz Quad-Core processor and a GeForce GTX 1080 4 GB graphics processing unit.

The participants were exposed to a virtual environment, and they experienced navigation with constant and instantaneous speed. In order to avoid practice effect, the sequence of speed type was selected randomly. Participants sat in a rotating chair for the study, and they were exposed to the virtual environment for 5 minutes for each speed type. We asked them to take a break of 5 minutes after each session. The navigation in the virtual environment

was system-automated. Participants could only view the surroundings by rotating their heads. The virtual environment that the users were exposed to was a 400-meter-long art exhibition hallway with some paintings hung on the wall (see Figure 3.5). The paintings in the scene were hung at different distances and near each painting the participants were stopped for 10 seconds and after that the travel was resumed towards the next painting. As there was no input from the user, there was no need for any trial to teach them how to navigate in the VE. We did not explain the nature of motion (what is instantaneous or constant speed) they are going to experience. However, they were aware that they were going to walk through a virtual art exhibition. We followed the similar approach (user rating) used by many researchers [40, 342] to measure the realism in a virtual environment. Once the participants experienced walking with both the speed types they were asked to rate the level of realism and preference rating on a scale of 1 to 10 (where 10 is highly real/preferred). We also asked for general feedback about the speed types in an informal interview. The entire procedure took approximately 20-25 minutes 4.5.

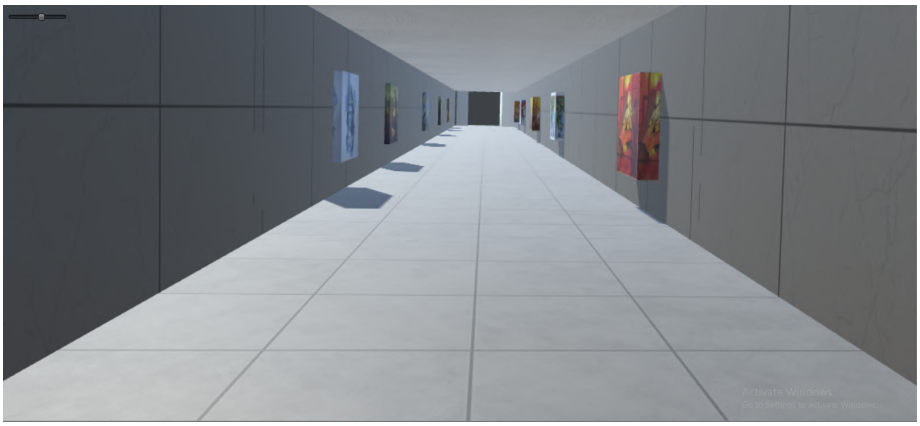


Figure 3.5: Visual stimulus for the realism comparison experiment

In this study we had two dependent variables namely realism and preference score. We formed the following null hypothesis:

- The mean difference of the realism score of instantaneous walking speed from the constant walking speed is less than equal to zero (positive for alternative hypothesis)

- The mean difference of the preference score of instantaneous walking speed from the constant walking speed is less than equal to zero (positive for alternative hypothesis)

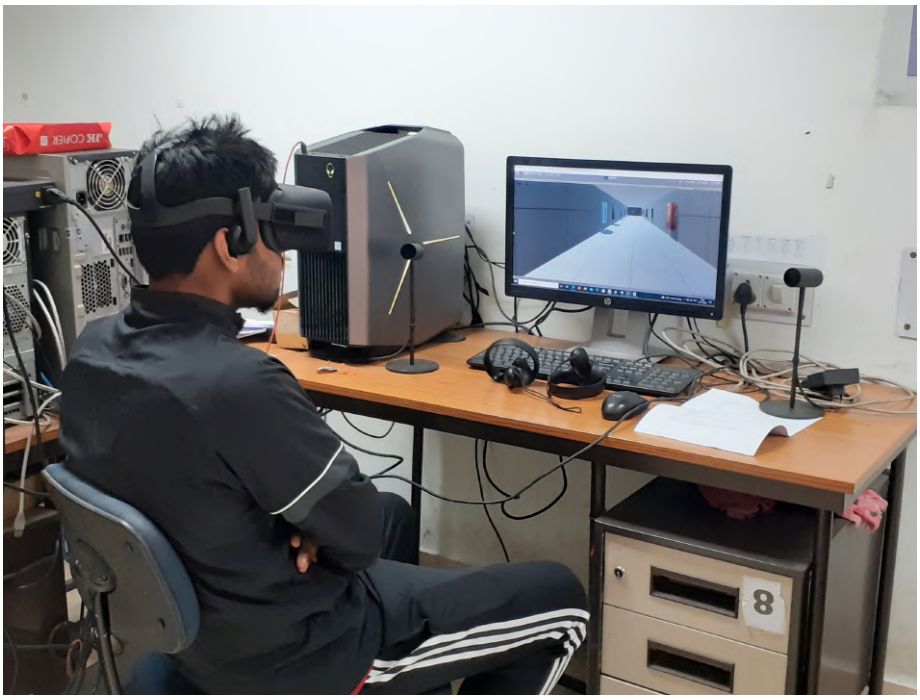


Figure 3.6: During the experiment of realism comparison

Realism and preference results In figure 3.7 (a) and figure 3.7 (b), we have shown the box representation of the ratings provided by the participants for the speed types. We ran a paired sample t-test (two tailed) on the collected score at $\alpha = 0.05$. We got t value for both the variables (t(33)=2.51 for realism and t(33)=2.74 for speed type preference) to be greater than the critical value (2.034). Thus we can reject both our null hypothesis and we can accept the alternative hypothesis (The use of instantaneous walking speed is more real and preferred than the average constant speed).

Though personal preference of the speed types varied by individual, we found 22 out of 34 participants preferred the instantaneous speed over the constant speed as optic flow. Six participants had neutral preference for both the speed types.

3.4 Discussion

We found the human step frequency to increase relatively up to a maximum and then decrease before reaching the desired target. It is possible to give an explanation for this behaviour. It is plausible that once an individual starts walking, they tend to set the maximum natural velocity by adjusting the step frequency from 0. Once they are comfortable with that

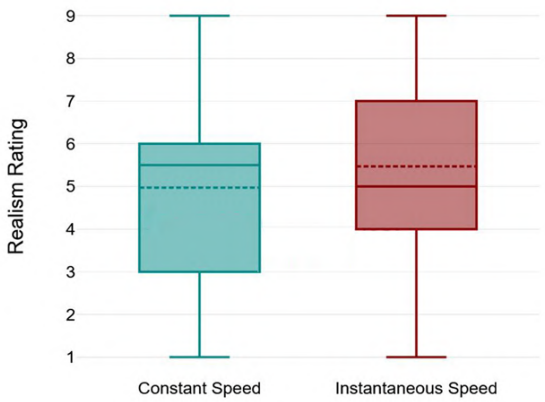


Fig. 3.7 (a): Participants' rating for realism

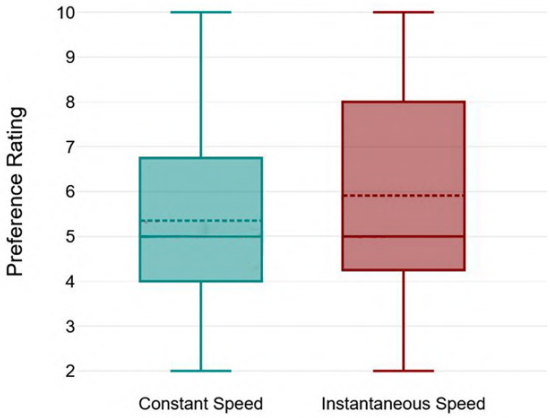


Fig. 3.7 (b): Participants' rating for preference of the speed type

velocity, they maintain that pace. However, just before reaching the target, they need to adjust the velocity by adjusting step frequency to accurately stop at the desired target. While using system automated travel, the target position is known from the beginning. So incorporating both the behaviour for system automated travel is reasonable.

In a system-automated tour, the user experiences movement in the virtual world while being stationary in the real world. As discussed in chapter 2, this may increase the chances of experiencing VR sickness, according to the sensory mismatch theory. So, the practical way of using our speed prediction model in VR would be to use vibrotactile feedback [40] or any VR sickness reduction techniques useful for system-automated travel. Our technique will be reasonable to use once the VR sickness, i.e. the fundamental problem of VR, is taken care of. Usage of any sickness prevention technique like WalkingVibe, may necessitate to modify the gain factor G in our model. For instance, as we did not use any sickness prevention technique in our study, we used the visual gain factor G to be 2.10 (established by Nilsson et al. [39]). However, the usage of vibrotactile feedback along with the visual stimuli to prevent sickness may change the visual gain significantly. We discuss it in detail in the next chapter.

In the model evaluation study, we observed that the lowest rating for both the speed types was provided by the users who had no experience with VR. Five of them reported discomfort after finishing the experiment. This observation also supports the conclusion made by previous research about the dependency of VR discomfort with the prior VR experience [301]. We observed that their preference ratings were low for the speed type

that was shown at the end. For instance, participant number 23 (P23) experienced constant walking first, followed by instantaneous walking speed. She rated the instantaneous walking lower than the constant walking, reporting discomfort during the experiment. We collected some qualitative feedback in order to clarify the results. "The gradual increase in speed was more natural to me than the sudden start of the movement" (P7, P8, P24, P32), "The imbalanced speed was more natural than the smooth movement"(P27), "The plain speed was comfortable to me, but the second one (the instantaneous walking speed) feels more realistic"(P25). Three participants (P9, P20 and P31) shared that they found the head oscillation to be missing in both the speed types.

In this study, we put one step forward towards the enhancement of realism for system-automated tours. We are keen to increase the realistic walking experience further, for example, by adding visual motions that reflect optic flow in a manner similar to natural body and head oscillations (as addressed by P9, P20 and P31) when walking. More dimensions of senses (like haptic and auditory) can be added to make virtual walking more realistic.

Moreover, this model predicts the walking speed for the straight path. We did not consider the turns in our assumption. In a virtual environment, if there is a turn between source and destination, we can break the journey into two parts, assuming the destination to be the turning point and again starting the journey to the next point. However, in real life, our walking behaviour varies with different types of twists and turns on the road.

3.5 Summary

This chapter addresses the challenge of providing a realistic walking experience in system-automated VR tours, where users do not provide continuous input for navigation. Unlike user-controlled tours, where participants actively navigate through the VR environment using various input methods, system-automated tours automate travel, potentially diminishing the realism of the walking experience due to low interaction fidelity. The primary aim of this work is to enhance walking realism in system-automated tours by predicting a user's instantaneous walking speed, thus improving the visual optic flow and creating a more immersive experience.

The chapter outlines the importance of realistic walking simulations in VR, noting

that most existing research has focused on the techniques for user-controlled VR tours that allow for dynamic user input to influence movement in the VR environment. However, system-automated tours have received less attention despite their potential for applications where user navigation input is impractical or undesirable. We argue that by improving the visual representation of walking speed (i.e. making it vary instantaneously in a way that mimics natural walking patterns rather than remaining constant), users can enjoy a more realistic walking experience even in the absence of direct navigational control.

We propose a model for estimating users' instantaneous walking speed by considering the user's actual height and the distance between their current position and a known destination in the VR environment. This model was developed and validated through an empirical study involving 40 participants, focusing on capturing real walking data to establish the model's parameters. The effectiveness of the model was then evaluated in a subsequent study with 34 participants, comparing the perceived realism of walking experiences using the predicted speed against a constant average walking speed.

The evaluation of the model demonstrated its effectiveness in enhancing the realism of the walking experience in VR. Participants in the study rated the instantaneous speed model as providing a more realistic experience compared to a constant speed model. This suggests that dynamic adjustment of visual speed to match natural walking patterns can significantly improve the sense of immersion in system-automated VR tours.

We do not claim our approach to be VR sickness-free. Due to the sensory conflict, our approach may induce VR sickness. However, researchers have come up with many approaches [40, 58] to reduce the VR sickness induced in system-automated VR tours by manipulating internal and efferent information. Yet, the proliferation of more efficient solutions for VR sickness will encourage developers to use our model while creating a system-automated tour involving a realistic walking experience. We do not even argue that our approach provides a more realistic experience than the fully interactive VR tours such as walking-based [10] and WIP [63]. However, while designing a system-automated VR tour that involves a visual sensation of walking, our approach will help to enhance the realistic walking experience.

In future, we intend to modify this model to be used for non-linear paths (paths having twists and turns). We also plan to compare the walking realism after applying the sickness

prevention mechanism along with our speed prediction model. We chose seated posture over standing while performing our study as the mentioned sickness prevention techniques are tested on a seated posture. However, the user experience of using our technique can be evaluated with different postures.

Chapter

4

Reducing CS in System-automated VR Tours

In the previous chapter, we explored enhancing the realism of walking experiences by simulating a person's instantaneous walking speed. Nevertheless, as discussed in Chapter 2, presenting continuous visual motion while remaining stationary can lead to a sensory mismatch, and prolonged exposure to a VR environment may induce cybersickness. We also reviewed various strategies to mitigate these effects from both the users' and developers' perspectives in Chapter 2. Among these strategies, vibrotactile feedback applied behind the user's ear has been designed specifically for system-automated VR tours. A study by Peng et al. [40] demonstrated that the use of vibrotactile feedback behind the ears significantly reduces cybersickness during system-automated VR walks. However, their study utilized a constant speed as the visual stimulus for walking. Our previous chapter revealed that employing our model, which predicts instantaneous walking speed, could further enhance the realism of the walking experience. Yet, the visual gain used in our model varies with different locomotion techniques, suggesting that the use of vibrotactile feedback while using our model might necessitate adjusting the visual gain than 2.1 (which we adopted from the study by Nilsson et al [39]). Since the gain 2.1 was established in the absence of vibrotactile feedback, there is a need to check whether the visual gain changes when the vibrotactile feedback is used as a CS mitigation method.

Another identified research gap in Chapter 2 concerning mitigation techniques for system-automated VR tours is the reduction of exposure duration. Exposure to a VR environment for a long duration may cause CS [70]. The chances of having cybersickness has been correlated to longer duration in VR environments, and the degree of the symptoms

rises proportionally with VR session duration [70, 343, 344]. Therefore, a shorter duration of VR exposure can always help in getting rid of these problems [70]. The duration of a VR tour is subject to the VR content. A VR environment can be small, like a room, or large, like a city, or sometimes even larger than that. In the case of a large VR environment, it is natural to have a longer duration of VR exposure. The duration of a VR tour can be dependent on the travelling time, the size of the virtual terrain and the interaction time (inspection, selection and manipulation). The travelling time can be modified by adjusting the travelling speed. It is obvious that the more travelling speed we apply, the shorter the duration will be. However, to provide a realistic walking experience, we cannot modify the speed. The size of the terrain can be manipulated by adjusting the distances between the points of interest in the terrain. Adjusting the size of the terrain can also exclude a user from getting the idea about the size of the real space [345]. The interaction time is dependent on the type of tour. If the tour is purely system-automated (i.e. user do not get any chance to interact during the tour), we can skip the interaction part. However, in this chapter let us assume the tour is semi interactive, where the travelling is automated by the system, but users get a chance to interact in between. In user-controlled travel, the interaction time varies from user to user. Therefore, the VR exposure duration can vary with the subject. Sometimes, a user may spend maximum time interacting with some non-important virtual objects. Because of the longer duration of VR exposure, s/he may experience the symptoms of VR sickness [71], leading to a lack of interest in the content, which may exclude him/her from exploring some important areas of the VR tour. By contrast, system-controlled travel can be used to limit the interaction time based on the required time. Therefore, for a system-automated tour, the reduction of duration can be done by optimising the path of the tour. However, we did not find any work that explored a way to reduce the duration of a VR tour to reduce CS.

In this chapter, we address the gaps of the two CS mitigation strategies for system-automated VR tours to maintain its realistic walking experience, namely the use of vibrotactile feedback and the reduction of duration. For the vibrotactile feedback technique, we establish a range of visual gains that can be applied to the realistic walking speed prediction model proposed in chapter 3 to make it more realistic. For the duration reduction technique, we establish a theoretical approach for finding the optimal path in a virtual world that can

be followed to visit maximum places of a large VR environment without compromising the realistic walking speed.

4.1 Finding a range of visual gain for the realistic walking model while using vibrotactile feedback as a CS mitigation strategy

In order to establish the perceptually realistic visual speed while using vibrotactile feedback, we conducted a within-subject empirical study. The details of the study are as follows.

4.1.1 Empirical study

Participants:

We recruited 21 volunteers (16 males, 5 females) aged from 25 to 40 years (Mean: 30.57, SD: 4.64), with the height of 152 to 182 cm (Mean:168.61, SD: 7.95) to take in the experiment. All the participants were postgraduate students from engineering backgrounds. Fifteen of them had a prior experience in VR. All the participants had normal or corrected to normal vision with no walking disabilities.

Procedure:

We followed the similar method used by Nilsson et al.[39] to establish the visual gain. Participants experienced 22 visuals with gain changes. There were 11 different gains, which were repeated twice randomly. We first described the tasks to the participants, and then they had to wear the HMD combined with the *walking vibe* set up (Figure 4.2). For vibrotactile feedback, we used a similar vibration motor (150 Hz, 9000 RPM, 12 mm coin style motor) used by [40], as it is claimed that using vibration with this frequency prevents the head discomfort caused by the vibrations. We made sure they were comfortable wearing those and asked them if they could quit anytime during the experiment if they felt any discomfort. However, no participant reported any discomfort during the experiment. Once they were comfortable wearing the headset, we provided vibrotactile feedback behind the participants' ears at the frequency of 1.8 steps/sec. The original walking speed of the user was estimated with the help of the equation proposed by Dean et. al.[74]. They expressed the walking

speed ($|v|$) of a person as:

$$|v| = (\frac{f}{0.157} \times \frac{h}{1.72})^2 \tag{4.1}$$

Where f and h represent the walking frequency and height of a person, respectively. As this equation takes height and frequency as input, the user's height was noted and calibrated through the Oculus Rift sensor. We asked the participants to close their eyes and imagine walking at that frequency so that they could judge the visual with their estimation. Once they were ready, we showed them a visual, and they had to choose between too slow, natural and too fast. Participants could utter their choice verbally once they were confident. After finishing the travel at a particular speed, the participants had to close their eyes and again imagine themselves walking with the frequency of the vibrotactile feedback. The visuals shown to them had gained changes, ranging from 1.0 to 3.0 in the increments of 0.2. This means the predicted normal speed is the slowest speed we showed, and the highest speed was three times greater than the normal speed.

Figure 4.1: The virtual environment shown to the participants

The visual stimulus that was shown to users was a 119-meter-long corridor of an art gallery (Figure 4.1). We used the Oculus Rift HMD with a resolution of 1080 × 1200 per eye. Two IR LED sensors (Oculus sensors) were placed at a distance of 2 meters facing the users. We could automatically calibrate the height of the user with the help of these sensors. The VR application was implemented using the Unity engine. We used the basic 3D objects from Unity and some textures downloaded from the web to build the VE. The vibrotactile

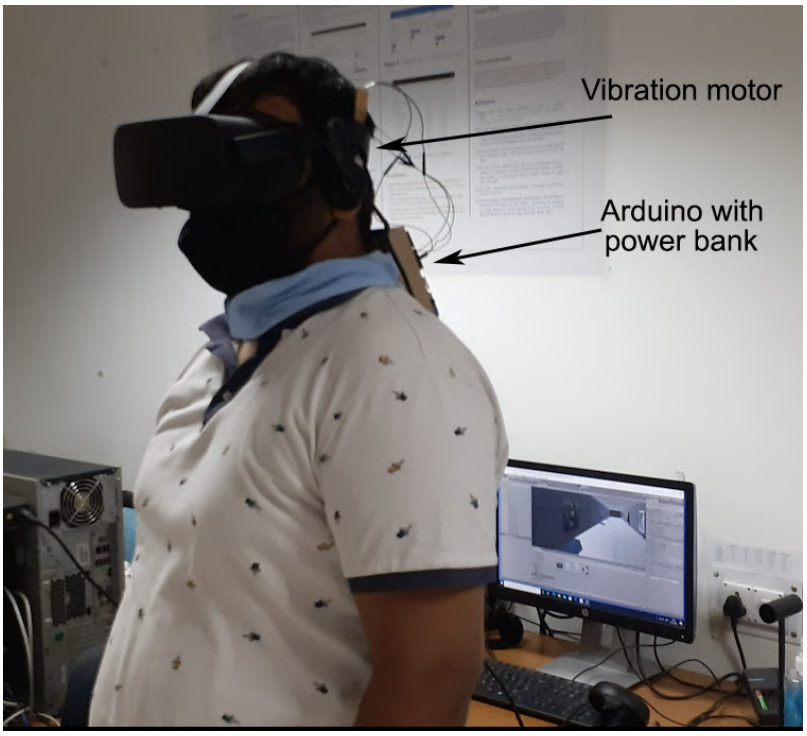


Figure 4.2: During the experiment of establishing visual gain

feedback was implemented with the help of Arduino, which was connected to a power bank and the vibration motors with the help of long connecting wires (shown in Figure 4.2).

Results & Discussions:

We took the weighted mean of all the gains that the users reported as 'natural'. We also identified the minimum and maximum of the gains that the participants stated as 'natural'. We found the mean of the reported natural gains to be 1.64 with an SD of 0.3. The average of the minimum and maximum gains rated natural is 1.40 with an SD of 0.21 and 1.78 with an SD of 0.26, respectively. We performed a one-sample t-test to confirm whether the range found by us is significantly different from the range found by Nilsson [39] for stationary VLT without the vibrotactile feedback. The test was performed for all minimum, maximum and mean values of gain found from our study and Nilsson's [39] study (taking α=0.01). We found our results to be significantly different in all three cases ($p < .00001$). We also found the mean and minimum gains to be significantly different from the normal speed.

As the range of visual gains for stationary VLT with vibrotactile feedback and without vibrotactile feedback are significantly different; we can conclude that if we use the vibrotactile

feedback to avoid motion sickness in stationary VLT, we might apply the visual gain from 1.40 to 1.78 with the speed we intend to show.

If we compare our results with the visual gains found for TIP, WIP and real walking in Nilsson's study [39], we can explain it following Barlow's subtractive model. The gain found in our study is significantly smaller than the gain found for real walking. This is because the non-visual felt speed perceived by the user would be more for real walking, which would be subtracted from the actual speed, thereby demanding more visual gain than the stationary VLT. However, we found the gain to be lesser than the gain found by nilsson for stationary VLT without using the vibrotactile feedback. This possible that use of vibrotactile feedback reduces the felt motion by stimulating the vestibular system.

The visual range we found in this study can be applied in VR in two ways. One possible way is to use the mean value of gain (i.e. 1.64), or it can be altered from 1.4 to 1.78 with the help of the user's feedback. Both ways may have different use cases. However, these values can be used when the user is using vibrotactile feedback that we have used in the study. The gain may vary if different vibration is applied.

We studied the natural walking speed, assuming the frequency to be constant. However, the step frequency of a person is not constant throughout the travel [41]. How showing the varying speed with the change in frequency (in a way it changes in real life) may affect the user experience needs to be evaluated.

4.2 Minimising the duration of a VR tour without compromising the realistic walking experience

4.2.1 Virtual environment to graph

The entire terrain where a system-automated VR tour is conducted can be represented as a weighted graph $G(\mathcal{V}, \mathcal{E})$, with the set of vertices $\mathcal{V}$ representing the sites (or points of interest) and the set of weighted edges $\mathcal{E}$ where the weight of each edge denotes the Euclidean distance between the pair of vertices it connects. We can represent the graph as a top view of the real terrain. However, the graph G does not need to be complete. In a real-life environment, there is not always a direct path from one place to another. Omitting the completeness of G allows us to formally represent terrains where at least a pair of sites

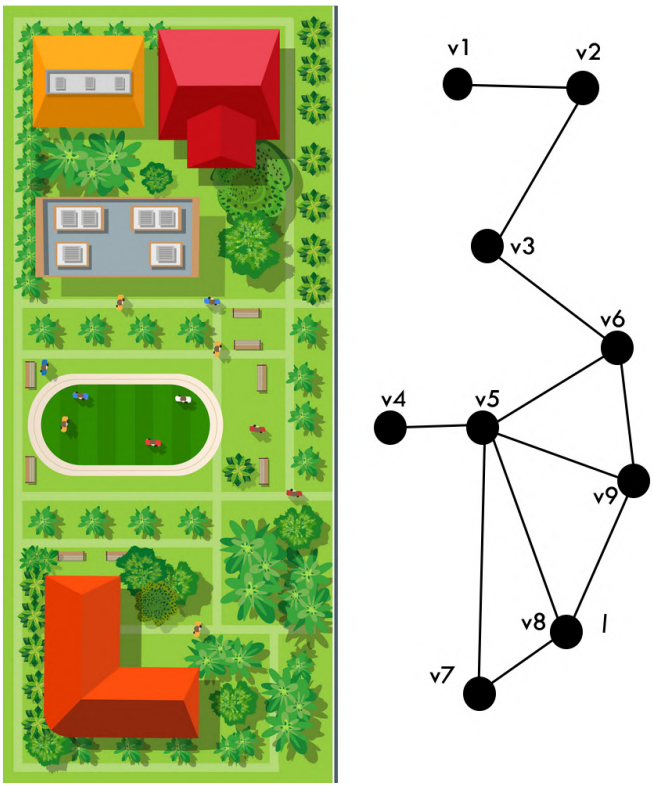


Figure 4.3: The virtual environment shown on the left hand side can be represented as the graph

are not directly reachable from one another. In Figure 4.3, we have shown the top view of a virtual environment, and based on some points of interest, we have represented it as a graph. For every pair of vertices $v_1, v_2 \in \mathcal{V}$ edge weights are given by the equation 4.2 where d denotes the Euclidean distance between a pair of nodes.

$$\mathcal{E}(v_1, v_2) = \begin{cases} d(v_1, v_2) & \text{if } v_1 \text{ and } v_2 \text{ directly reachable} \\ \infty & \text{otherwise} \end{cases} \tag{4.2}$$

Formally, we want to find a path $\mathcal{P}$ in the given graph G, which covers all vertices in $\mathcal{V}$ and finishes in the least possible time while maintaining a realistic speed throughout. In order to compute the amount of time required to visit the path $\mathcal{P}$, we first calculate the time taken to move from a node(v_1) to its adjacent node (v_2) located at $d(v_1, v_2)$ distance away. The time depends on the travelling speed of the user during the tour. Once we calculate the time taken to move from one node to the adjacent node, we can use it to compute the time needed to cover the path $\mathcal{P}$, which will be the sum of time taken to visit the nodes

present in it sequentially.

4.2.2 Estimating minimum time to complete the tour

Let us first formally estimate the time required to travel between any two directly connected nodes while maintaining natural walking speed. As seen in equation 3.4 from the previous chapter, we can break it down into two parts (As shown in equation 4.4 and 4.5). The instantaneous frequency $F(x)$ is formulated as a square wave with amplitude F_m (found to be 1.8 steps/sec in the previous chapter). Also the equation 3.2 from the previous chapter, computes the the instantaneous virtual speed of the user at a distance x from the starting node. We will compute the total time required to travel between node v_1 and v_2 which are d distance apart. Equation 3.2 can be rewritten as a differential equation (equation 4.5), and the solution to the differential equation will give us the time required to travel between any two adjacent nodes that are d distance apart.

$$Sq(x) = \frac{4}{\pi} \times \sum_{k=0}^{\lceil \frac{d}{2} \rceil - 1} \frac{sin((2k+1)\omega x)}{(2k+1)} \tag{4.3}$$

$$F(x) = F_m \times Sq(x) \tag{4.4}$$

$$\frac{dx}{dt} = V_m \times F(x)^2 \tag{4.5}$$

$$dt = \frac{1}{V_m} \times \frac{dx}{F(x)^2} \tag{4.6}$$

$$\frac{1}{F(x)^2} \approx \frac{(3 - 2Sq(x))}{F_m^2} \tag{4.7}$$

In equation 4.7, we have used the Taylor series of the function $f(z) = \frac{1}{z^2}$ at the point $z_0 = 1$ and only kept linear term while ignoring the higher order terms. At $z_0 = 1$ first two terms of the Taylor series gives us $\frac{1}{z^2} = 1 - 2(z-1) = (3 - 2z)$. Considering $z = Sq(x)$ the function $1/Sq(x)^2$ can be approximated as $(3 - 2Sq(x))$. $Sq(x)$ is Fourier expansion of the square wave with amplitude one so for $x \in (0, d)$, the value of $Sq(x)$ will be very close to point $z_0 = 1$.

Let us denote the time taken to travel between the adjacent nodes with distance d as $\mathcal{T}(d)$. Then, by solving the equation 4.6 we compute the closed expression of $\mathcal{T}(d)$. Note the integration constant C is 0 as $\mathcal{T}(0) = 0$. If we keep the first n_0 terms of infinite series of

the square wave given in equation 4.3, then the closed expression of $\mathcal{T}(d)$ is computed in equation 4.10. If $n_0 \to \infty$ the series $\sum_{k=1}^{n_0} \frac{1}{(2k-1)^2}$ converges to $\pi^2/8$ and $\mathcal{T}(d)$ converges to $d/(V_0)$. As expected, that will be the time required by the user if they maintain a realistic walking speed of V_0 throughout the travel.

$$\int_0^d \frac{1}{V_0} \times (3 - 2Sq(x))dx = \int_0^t dt \;\; where \;\; V_0 = V_m \times F_m^2 \tag{4.8}$$

$$\mathcal{T}(d) = \frac{1}{V_0}\left(3d - \frac{16d}{\pi^2}(1 + \frac{1}{3^2} + \frac{1}{5^2} + \dots)\right) + C \tag{4.9}$$

$$\mathcal{T}(d) = \frac{d}{V_0}\left(3 - \frac{16}{\pi^2} \times \sum_{k=1}^{n_0} \frac{1}{(2k-1)^2}\right) \tag{4.10}$$

Now as the time taken between two points are known, we can estimate the the minimum time to visit all the sites of the virtual environment. The entire terrain is represented as a connected weighted graph $G(\mathcal{V}, \mathcal{E})$ where edge weights are given in equation 4.2. Our goal is to find path $\mathcal{P}$ which covers all the vertices in $\mathcal{V}$ and completes in the least possible time. The time taken to complete a path is the sum of time required to visit the nodes adjacent in the path sequentially. Suppose a path P is represented as the sequence of nodes $\{v_1, v_2, \dots, v_n\}$ then time taken to complete P is defined as followed where $\mathcal{T}(\mathcal{E}(v_i, v_{i+1}))$ denotes the time taken to travel between two adjacent nodes $\mathcal{E}(v_i, v_{i+1})$ apart and $\mathcal{T}(d)$ is computed in equation 4.10 .

$$\mathcal{T}(P) = \sum_{i=1}^{n-1} \mathcal{T}(\mathcal{E}(v_i, v_{i+1})) \tag{4.11}$$

$$min_P \; \mathcal{T}(P) \quad \text{where} \;\; \forall v \in \mathcal{V} \;\; v \in P \tag{4.12}$$

Formally we want to solve the mentioned equation 4.12. It is a variation of the celebrated Travelling salesman problem(TSP), which solves the following problem: given a complete graph, what is the shortest possible path that visits each node exactly once and returns to the starting node. The shortest path is defined based on the total weight of all the edges included in the path. Note we didn't assume the graph G to be a complete graph, so there may be cases where there is no Hamiltonian path in G i.e. a path where all the nodes are covered and each nodes are visited exactly once. One popular way to solve this issue is to introduce hypothetical edges between vertices where there was no edge between the original

graph G. If there is no direct edge between two vertices v_1, v_2, then we introduce a new edge with length equal to the shortest path between v_1 and v_2 in the graph G. The shortest path between any pair of vertices is well defined as the graph G is assumed to be connected then there exists a path between any pair of vertices. We modify the edge set of the given graph $G(\mathcal{V}, \mathcal{E})$ to $G'(\mathcal{V}, \mathcal{E}')$ based on the previous idea mentioned in equation 4.13. Here $SP(v_1, v_2)$ denotes the length of the shortest path between v_1, v_2 in graph G.

$$\mathcal{E}'(v_1, v_2) = \begin{cases} \mathcal{E}(v_1, v_2) & \text{if } v_1 \text{ and } v_2 \text{ directly reachable} \\ SP(v_1, v_2) & \text{otherwise} \end{cases} \tag{4.13}$$

In order to solve for the optimal path mentioned in equation 4.12, we reduce this to the problem on the modified graph $G'(\mathcal{V}, \mathcal{E}')$ to an instance of the TSP problem. As the $G'(\mathcal{V}, \mathcal{E}')$, we only need to compute the time required to travel between any pair of nodes. We have already formulated the time required to travel between two adjacent nodes $\mathcal{T}(d)$ in equation 4.10. Let $G_{in}(\mathcal{V}_{in}, \mathcal{E}_{in})$ be the input of the corresponding TSP. Then $\mathcal{V}_{in} = \mathcal{V}$ and $\forall v_1, v_2 \in \mathcal{V}$ the weight of the edge connecting v_1 and v_2 is defined as $\mathcal{E}_{in}(v_1, v_2) = \mathcal{T}(\mathcal{E}'(v_1, v_2))$. One thing to consider is that the output of TSP will give us a path on the graph G', not on the original graph G, in other word it may have some hypothetical edges that are present in $\mathcal{E}'$ but not in the original edge set $\mathcal{E}$. Suppose the optimal path $\mathcal{P}$ as outputted by the TSP solver has an edge connecting v_1, v_2 that was not present in the original edge-set $\mathcal{E}$ then $\mathcal{E}(v_1, v_2) = \infty$. We know if $\mathcal{E}(v_1, v_2) = \infty$ then $\mathcal{E}'(v_1, v_2) = SP(v_1, v_2)$. In the optimal path $\mathcal{P}$ we can replace the edge between v_1 and v_2 by the shortest path between them, and as the shortest path between v_1, v_2 is computed w.r.t G, it only consists of edges from $\mathcal{E}$. Note replacing the edge with shortest path does not alter the time taken to complete the path $\mathcal{P}$ as $\mathcal{E}'(v_1, v_2) = SP(v_1, v_2)$ and $\mathcal{T}(\mathcal{E}'(v_1, v_2)) = \mathcal{T}(SP(v_1, v_2))$.

Though a VR tour is optimal, the time to do the entire tour may take longer time. So, one obvious variation of the aforementioned problem is to find the optimal path that completes on/before the given threshold $\mathcal{T}_0$. The threshold can be the duration we want our tour to be, which can be decided based on the type of content. For instance, we can perform an empirical study to establish the threshold for our content. In this case, it is not mandatory and not always possible to visit every single node present in $\mathcal{V}$. For this purpose,

Algorithm 1 Compute optimal time & path that covers all sites

Input: the graph $G(\mathcal{V}, \mathcal{E})$ representing the terrain.
Output: the optimal time $\mathcal{T}$ and path $\mathcal{P}$.
Compute the complete graph $G'(\mathcal{V}, \mathcal{E}')$
for $\forall v_1, v_2 \in \mathcal{V}$ **do**
    if $\mathcal{E}(v_1, v_2) < \infty$ **then**
        $\mathcal{E}'(v_1, v_2) = \mathcal{E}(v_1, v_2)$
    else
        $\mathcal{E}'(v_1, v_2) = SP(v_1, v_2)$
Compute the input graph $G_{in}(\mathcal{V}_{in}, \mathcal{E}_{in})$ to TSP solver.
$\mathcal{V}_{in} = \mathcal{V}$
for $\forall v_1, v_2 \in \mathcal{V}$ **do** ▷ Compute time required to travel between any pair of nodes
    $\mathcal{E}_{in}(v_1, v_2) = \mathcal{T}(\mathcal{E}'(v_1, v_2))$
$\mathcal{T}, \mathcal{P} = SolvTSP(G_{in}(\mathcal{V}_{in}, \mathcal{E}_{in}))$ ▷ Get optimal time and path with TSP solver
for $\forall v_1, v_2$ adjacent nodes in $\mathcal{P}$ **do**
    if $\mathcal{E}(v_1, v_2) = \infty$ **then**
        Replace hypothetical edge joining v_1, v_2
        with $SP(v_1, v_2)$

we define the importance of every node through a function $\mathcal{F} : \mathcal{V} \to \mathbb{R}^+$ which maps every node $v \in \mathcal{V}$ to a positive real value denoting its importance. The total importance of a path $P = \{v_1, \dots v_n\}$ is defined as $\mathcal{F}(P) = \sum_{i=1}^{n} \mathcal{F}(v_i)$. Formally, we want to maximize $\mathcal{F}(P)$ for paths that complete within the given time threshold $\mathcal{T}_0$, i.e. $\mathcal{T}(P) \leq \mathcal{T}_0$. The problem formulation is described in equation 4.14. As mentioned in the previous section, another aspect that needs to be considered is the waiting time, i.e. the amount of time required by the user to inspect/interact with a particular object on the node(or point of interest). We consider waiting time ($\mathcal{W}$) for each node while calculating the optimal time for the candidate paths. For this problem variant, we want to solve the problem given by the equation 4.15.

$$max_P \;\; \mathcal{F}(P) \;\; \text{where} \;\; \mathcal{T}(P) \leq \mathcal{T}_0 \tag{4.14}$$

$$max_P \;\; \mathcal{F}(P) \;\; \text{where} \;\; \mathcal{T}(P) + \mathcal{W}(P) \leq \mathcal{T}_0 \tag{4.15}$$

This problem is also well known in operation research and known to be a generalization of the Vehicle Routing Problem (VRP). In literature, this variation of the VRP is known as VRPP [346] or Team Orienteering Problem (TOP) [347] where we don't need to visit all the nodes but instead maximize the profit while respecting the time constraint. Here by profit,

we mean the priority of the node (site) in the graph (terrain). So we can use the solvers of the VRPP once we reduce the problem to VRPP. Given the graph $G(\mathcal{V}, \mathcal{E})$ computation of complete graph $G'(\mathcal{V}, \mathcal{E}')$ and the input graph $G_{in}(\mathcal{V}_{in}, \mathcal{E}_{in})$ similar to previous section. As mentioned in the previous section, the optimal path $\mathcal{P}$ may have some hypothetical edges that were not present in the original edge-set $\mathcal{E}$. For all these edges connecting node pair v_1, v_2 similar to the previous section, we must replace them with the shortest path between them $SP(v_1, v_2)$.

4.2.3 Implementation

We implemented our approach by making a VR tour of the cultural heritage of Majuli Island. Majuli is one of the largest inhabited riverine islands in the world (with 553 square kilometres of area). Located in the upper reaches of the river Brahmaputra this island has a rich cultural heritage. It is a holy place of Vaishnavite religion and culture (consists of around twenty-two Vaishnavite temples), where different races and tribes are assimilated. It is located in the upper reaches of the river Brahmaputra and within the latitude of 26°45'N-27°12'N and longitude of 93°39'E- 94°35'E. The geographical location, natural environment, richness in archaeological wealth, and biodiversity denote Majuli a unique identity. However, the existence of this cultural heritage (CH) is under serious threat due to the recurring floods and incessant bankline erosion by the Brahmaputra and its tributaries. So it is necessary to preserve the CH to pass the information to the next generations. Traditional ways to preserve these are oral-lores, museums, maintaining live sites and documentation. In the era of the internet, these modes of preservation are unable to attract the younger generations. CH also act as a means of promoting tourism, thus adding to the nation's economy. Digitizing CH is the solution for the above-raised issues. Digitization of CH uses archiving technology and immersive technology. Immersive technologies like VR and AR make the cultural content appealing and enhance the experience of the user.

There is a total of twenty-two Vaishnavite monasteries, a.k.a. satras, in Majuli. Source and place development of Vaishnavite culture of Assam, the satras are the heart of Assamese society. A satra consists of Naamghar (prayer hall), Monikut (place of idol), Haatis (hostel of the deciples), and the residence of the satradhikar (master of the satra). One of the largest, most prominent and most visited satras is Auniati satra. The campus of the satra

covers approximately 1200 sq. meters with a minimum of 25 interest points. In a real environment, a tourist may take the whole day to explore the site. However, a user is not recommended to be in a virtual environment for so long. To make the user feel present on the CH site, we chose to use mundane travel rather than magical. We used the proposed approach to minimize the duration to cover all the important places.

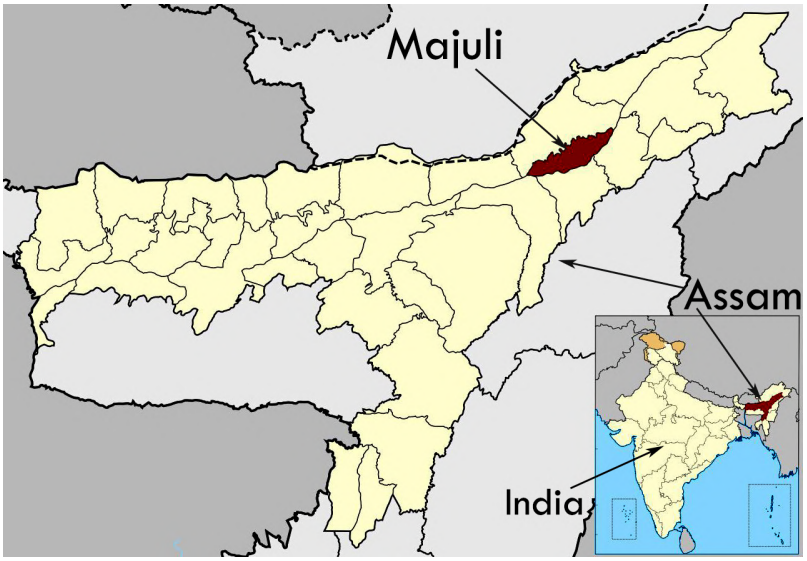


Figure 4.4: Majuli location

The virtual terrain

The virtual terrain was first mapped to a weighted graph with 25 nodes. The possible paths were connected as the edges of the graph. The weights of the edges were assigned based on the Euclidean distance between the nodes. The graph for the terrain is shown in the Figure 4.6. The 3D environment was created using Maya 3D, and the application was built using the Unity 3D game engine (version 2021.3.2f1).

Assigning node weight

By the term node weight, we mean the priority of the node (site) with respect to other nodes. We physically went to the place to observe the frequency of visits to the places. We visited the place during the festival called "Kati Bihu", which is one of the special festivals celebrated in this satra (also included in our virtual tour). We followed two ways to assign the priorities of the nodes. First, we observed people's frequency of visits to the places for

Node Number	Node Weight ($\mathcal{F}$)	Waiting time($\mathcal{W}$)
1	1	4 sec
2	1	4 sec
3	1	4 sec
4	5	8 sec
5	8	10 sec
6	8	300 sec
7	6	4 sec
8	6	10 sec
9	1	4 sec
10	6	600 sec
11	1	4 sec
12	1	4 sec
13	1	4 sec
14	5	180 sec
15	5	60 sec
16	5	10 sec
17	5	10 sec
18	5	10 sec
19	5	10 sec
20	5	10 sec
21	5	10 sec
22	5	10 sec
23	1	4 sec
24	6	900 sec
25	2	900 sec

Table 4.1: Node weight and waiting time for auniati satra

the whole day. Secondly, we talked to the local people(mostly the people that reside on that campus) and the people who are on the committee of that satra. We asked them two questions “What are the places people visit when they come to this place?" and "What are the places you suggest a tourist must visit when they come to this satra?”. Based on the observation and the interview, we assigned the priorities of the places. We also observed the objects where people want/need to interact in those places. We noted the average time for each point of interest to determine the waiting time for the nodes (Listed in table 4.1). The waiting time was determined to allow people to interact with the objects in the virtual tour.

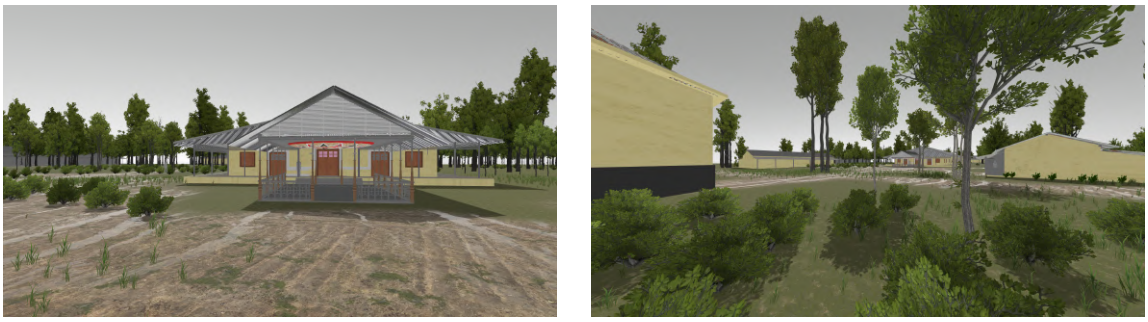


Figure 4.5: Screenshots of the virtual tour

Minimum duration and optimal path calculation

Once we have the input graph with 25 nodes depicted in Figure 4.6, we ran the algorithm described in algorithm 1. As mentioned in lines 4 — 12 in the algorithm, we first converted the graph into a complete graph by introducing new edges between the nodes that are not directly connected with a length equal to the shortest distance between them in the original graph. Once we have reduced the problem to an instance of the TSP, we used TSP solvers provided by Google's operation research tools [348] to compute the optimal path and optimal time that will be able to cover the entire path. We calculated the walking speed using the equation 4.10.

For finding the optimal path with time threshold as described in section 4.2.2 we used the time threshold of 10, 20, 30 and 40 minutes (shown in table 4.2). In this case, the weights of different sites are based on their importance, as mentioned in the previous section. For this problem, we have used the solver for VRPP with a specified time window(30 minutes in our case). We have used Google's operation research tools [348] for this purpose.

Empirical Study

In order to evaluate how our approach of minimizing the duration affects the user's experience, we conducted a between-subject study. The primary task of the participants is to explore the virtual world using the optimized tour selected by our method or using an arbitrary path selected by the user. We used the Oculus Rift CV1 VR system running on a computer with a 3.6 GHz Quad-Core processor and a GeForce GTX 1080 4 GB graphics processing unit.

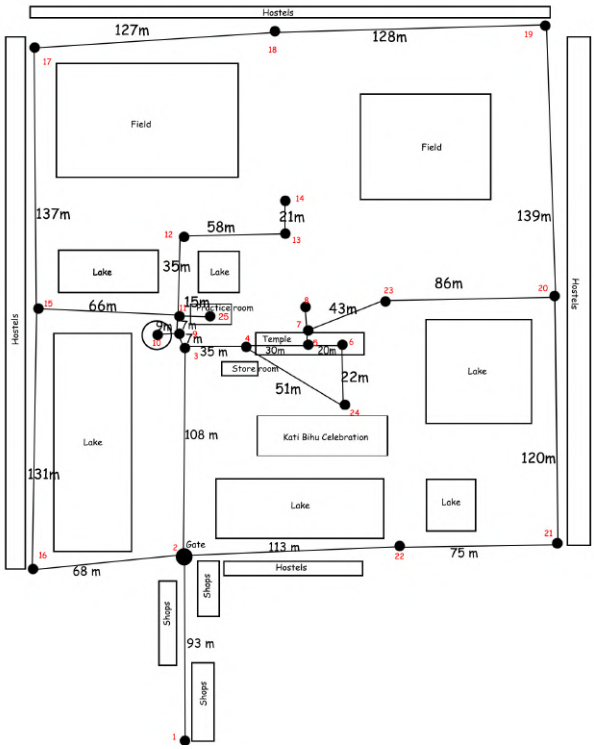


Figure 4.6: The map of auniati satra represented as graph

The two types of tour

As discussed in the previous section, we built the 3D model of the terrain. In the same environment, we implemented two types of tours, namely optimal tours and arbitrary tours. In the optimal tour, we considered the optimal path calculated from the discussed method. In the arbitrary tour, the user needs to draw a path connecting all the nodes at the beginning of the tour.

Participants

We recruited a sample size of 20 participants for each type of tour (optimal path tour and arbitrary tour). The participants (3 participants) who could not complete the experiment due to severe symptoms of VR sickness were replaced until the size of the sample was fulfilled. A total of 40 participants ranged in age from 17 to 36 years (M=26.87, SD= 4.01). All the participants had normal or corrected to normal vision. Eighteen participants had prior VR experience, and the remaining had no experience in VR. We equally divided these

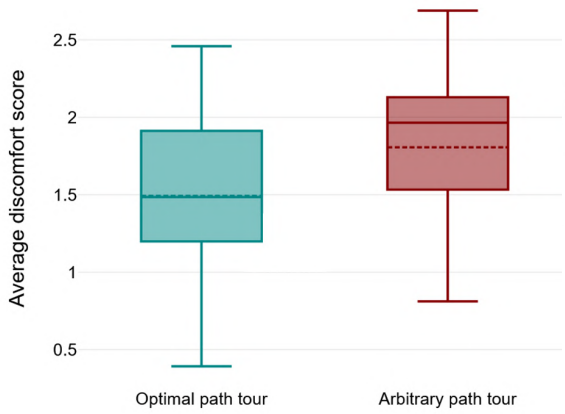


Fig. 5 (a): Participants' average rating for discomfort

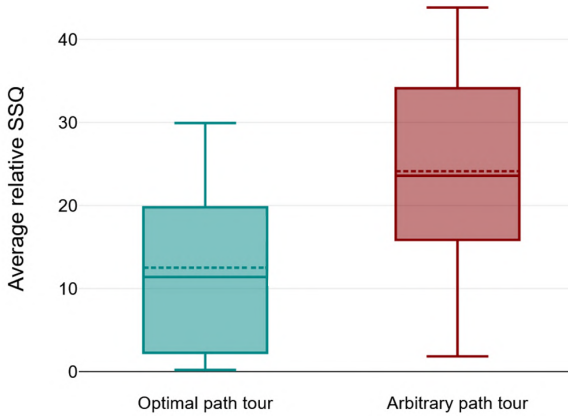


Fig. 5 (b): Participants' average relative sickness score

eighteen participants into both tours.

Procedure

We divided the 40 participants for the two types of tours equally. Before the experiment, we asked the participants to fill up a questionnaire regarding their personal information (including experience in VR). They experienced the tour in a comfortable seated position. Before starting the tour, we provided the participants with some basic information about the place Majuli and the satra for which the tour was created. We told them that they needed to visit all the checkpoints and that there would be a quiz on this tour at the end. For both tours, at each checkpoint, some time was given to the user for inspection or interaction (as shown in table 4.1). However, they could move to the next checkpoint anytime by pressing the "B" button of the Oculus controller. An alert was given to the user 5 seconds before moving to the next checkpoint automatically. This was done so that the user does not feel uncomfortable when suddenly s/he starts moving to the next checkpoint by the system. The participants experienced walking at a speed of calculated using the equation 4.10. In the case of the arbitrary tour, a map pops up before starting the tour, and the user needs to draw the route connecting all the nodes (checkpoints).

Throughout the study, we used two different questionnaires. The first questionnaire was about the measurement of discomfort [150]. This measure was taken after reaching each checkpoint. Once the user reached a checkpoint, s/he was asked to rate the level of discomfort from 0(no discomfort) to 10(very uncomfortable) verbally. Once a user rated

Time threshold	Optimal Path	Optimal Time
10min(600sec)	$1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 7 \rightarrow 8 \rightarrow 7 \rightarrow 5 \rightarrow 4 \rightarrow 3$ $\rightarrow 9 \rightarrow 11 \rightarrow 12 \rightarrow 13 \rightarrow 12 \rightarrow 11 \rightarrow 15 \rightarrow 16$ $\rightarrow 15 \rightarrow 17 \rightarrow 18 \rightarrow 19 \rightarrow 20 \rightarrow 21 \rightarrow 22$	581.25sec
20min(1200sec)	$1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 7 \rightarrow 8 \rightarrow 7 \rightarrow 5 \rightarrow 6 \rightarrow 5$ $\rightarrow 4 \rightarrow 3 \rightarrow 9 \rightarrow 11 \rightarrow 12 \rightarrow 13 \rightarrow 14 \rightarrow 13$ $\rightarrow 12 \rightarrow 11 \rightarrow 15 \rightarrow 16 \rightarrow 15 \rightarrow 17 \rightarrow 18 \rightarrow 19$ $\rightarrow 20 \rightarrow 23 \rightarrow 20 \rightarrow 21 \rightarrow 22$	1126.25sec
30min(1800sec)	$1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 7 \rightarrow 8 \rightarrow 7 \rightarrow 5 \rightarrow 6 \rightarrow 5$ $\rightarrow 4 \rightarrow 3 \rightarrow 9 \rightarrow 10 \rightarrow 9 \rightarrow 11 \rightarrow 12 \rightarrow 13 \rightarrow 14 \rightarrow 13$ $\rightarrow 12 \rightarrow 11 \rightarrow 15 \rightarrow 16 \rightarrow 15 \rightarrow 17 \rightarrow 18 \rightarrow 19$ $\rightarrow 20 \rightarrow 23 \rightarrow 20 \rightarrow 21 \rightarrow 22$	1730.58sec
40min(2400sec)	$1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 7 \rightarrow 8 \rightarrow 7 \rightarrow 5 \rightarrow 6 \rightarrow 5$ $\rightarrow 4 \rightarrow 3 \rightarrow 9 \rightarrow 10 \rightarrow 9 \rightarrow 11 \rightarrow 25 \rightarrow 11 \rightarrow 12 \rightarrow 13$ $\rightarrow 14 \rightarrow 13 \rightarrow 12 \rightarrow 11 \rightarrow 15 \rightarrow 16 \rightarrow 15 \rightarrow 17$ $\rightarrow 18 \rightarrow 19 \rightarrow 20 \rightarrow 23 \rightarrow 20 \rightarrow 21 \rightarrow 22$	1887.78sec

Table 4.2: Optimal path calculated from our approach from different time thresholds

10, we immediately discontinued the study and instructed her to take a rest. The next questionnaire was about the measurement of simulator sickness. It was conducted using the SSQ [349]. Here the users were asked several questions related to the sickness before and after experiencing the virtual tour. We calculated the relative sickness score as performed in [40].

Results

The data from the participants' responses to perceived discomfort and VR sickness were analyzed. The findings are as follows: The data from the participants' responses to perceived discomfort and VR sickness were recorded. The findings are as follows:

Discomfort Discomfort ratings were collected at each checkpoint, and we calculated the average overall discomfort rating for each participant. Our goal was to determine if there was a significant difference in discomfort levels when using the optimal path tour. Initially, we conducted a Shapiro-Wilk test to assess the normality of the data, which confirmed that the data were normally distributed. Subsequently, an F-test was performed to examine the equality of variances between groups, revealing equal variances for the

discomfort data in both groups. Therefore, we proceeded with a two-tailed t-test for both conditions. This analysis revealed a significant difference in discomfort between the tours (t= 1.90082, p= .032465 at an alpha level of 0.05), indicating that the discomfort rate was notably higher for the groups navigating the arbitrary path-based tour.

VR sickness After gathering the SSQ scores, we calculated the relative sickness score by subtracting the pre-tour score from the post-tour score for each participant. The Shapiro-Wilk test indicated that the data were normally distributed, and the F-test showed equal variances between the two groups. To assess the significance of the difference in relative sickness scores between the two groups, we conducted a two-tailed t-test. The results showed a significant difference in VR sickness between the optimal and arbitrary path-based tours (t-value = 3.33915, p-value = .000946 at an alpha level of 0.05). Participants who experienced the optimal path tour reported lower VR sickness scores compared to those who took the arbitrary path tour.

4.2.4 Discussion

According to the results of the empirical study, we found that usage of an optimal path may benefit in reducing sickness and discomfort. However, we noticed that most of the participants did not stay at a particular checkpoint for its estimated time. We believe that this can be better estimated by performing an empirical study after implementing the tour. The exact estimation would be more useful when the user does not have any choice but to wait for the system to move her to the next checkpoint. In our case, estimating a higher waiting time than the user wants to stay at a site is beneficial.

We have also shown how we can compute the path covering maximum sites at a given duration based on their importance. Our method of optimizing the duration of a virtual environment can be used while planning the path during a tour. It can be used to estimate the overall time for the virtual environment. We may not be able to cover everything, given the time threshold. However, if the user takes less amount of time in the sites than expected, the time saved in interaction can be utilized in exploring the remaining sites. If the threshold of the duration is more than the estimated time, the whole tour can be designed as it is. Nevertheless, reducing the VR exposure duration always keeps the user

on the profit side. If the threshold of the duration is less, we can make many variations based on the type of terrain and the intention of the tour. We may reduce the distance between sites of the terrain (some edges of the graph) to fit it in the given duration, or we may change the travelling speed based on the time that needs to be reduced. If we modify the virtual travel speed, we may need to compromise on the perceived realism [144]. We can apply faster travel speed wherever the realistic walking experience is not needed much. However, in real life we do not always walk around the whole terrain. Based on the terrain, we may choose different modes of travel. Sarupuri et. al. [350] used different locomotion techniques to travel in a virtual world. If the distance between two sites is too far, we may also add vehicles to make the travel faster. However, further study is required to understand the factors that need to be modified in an optimum way (depending on the virtual content) when we need to finish a tour in a given time threshold.

Sometimes, it is not important to cover all the points for all the users. Based on the user's interest, the algorithm can be modified so that the interest points are visited based on the user's preference. The user's preference can be captured before starting the tour. For some tours, it may happen that some point is not allowed to visit before visiting a particular place. For instance, if we want to create a tour of a museum, where the sections of the museum must follow a sequence for visiting the places. In that case, we should disconnect the edges joining the sections which are not in a sequence.

Our approach is limited to system-controlled travel. Because both the minimum time and optimal path are calculated assuming a fixed time to be spent for each place given to the user by the system. If the tour becomes fully user-controlled, the user may spend more time navigating or staying in one place, exceeding the ideal duration's time threshold. However, this problem can be solved by guiding/alerting the user to move to the next place whenever necessary. In that case, how it will affect user experience needs to be studied.

4.3 Summary

This chapter addresses two primary gaps in CS mitigation strategies for system-automated VR tours: enhancing realism while using vibrotactile feedback strategy and reducing exposure duration to minimize CS risks. The first part of the chapter establishes a range of visual gains

that need to be used while using a vibrotactile feedback as a CS mitigation strategy. Here the visual gain is for maintaining the realisitic walking experience, where the vibrotactile feedback is to mitigate CS in system-automated tour. Through empirical study involving 21 participants, the study identified that the use of visual gain ranging between 1.40 and 1.78 can be used to simulate realistic walking speed, while using the vibrotactile feedback behind the ears. This finding shows a significant difference from previous benchmark, which did not account for the impact of tactile feedback during the stimulation.

The chapter also tackles the challenge of reducing VR exposure duration to lower the risk of CS. It presents a novel approach that conceptualizes the virtual environment as a weighted graph, where vertices represent points of interest and edges denote the walking paths between them. The objective was to calculate the shortest path that allows a user to visit all sites of interest within the minimum time or given a time to visit all the important points of interest, thereby minimizing VR exposure without detracting from the tour's educational or entertainment value. This method was applied to develop a virtual tour of Majuli Island's cultural heritage, demonstrating its effectiveness in creating an immersive yet comfortable experience for users.

A between-subject study evaluated the user experience of the optimized tour against a non-optimized counterpart. A total of 40 participants were divided into two groups, with one experiencing the system-automated tour designed using the proposed methods and the other navigating an arbitrary path. Results indicated that the optimized path significantly reduced discomfort and symptoms of CS, highlighting the benefits of the proposed strategies in enhancing user comfort in a system-automated VR tour.

Chapter

5

Emotional State Detection while Navigating in a VR Environment

The chapters 3 and 4 were based on enhancing the walking experience in system-automated tours. This chapter is focused on enhancing the navigational experience in a user-controlled VR tour. As discussed in chapter 2, studies have unveiled a correlation between a user's emotional state and the likelihood of experiencing CS [120, 114]. This finding underscores the importance of recognizing and managing emotional responses in VR experiences to mitigate CS risks, ensuring a more comfortable and enjoyable VR tour for users. In recent advancements, researchers are actively integrating the emotional state of users into the realm of VR tours to create more immersive, personalized, and adaptive experiences [351, 352]. By considering emotions like excitement, stress, or calmness during virtual exploration, developers aim to tailor the content dynamically. Integrating emotional understanding not only enhances immersion and personalisation but also prioritizes user well-being in the evolving landscape of VR.

Previous studies have made significant strides in understanding and predicting emotional states by leveraging biosignals such as HR, EEG, and respiratory signals [351]. These physiological indicators offer valuable insights into an individual's emotional state. However, harnessing these bio-physiological signals typically demands the use of external sensors, presenting considerable challenges, especially in consumer-level HMDs like Oculus Quest, HTC Vive, Sony PlayStation VR, and others. The necessity of external sensors complicates the deployment of standalone emotional state predictors within these HMDs. Furthermore, users are often required to restrict their movements during the data collection process to

minimize noise in the bio-physiological signals. This constraint imposed by external sensors not only limits users' freedom of movement but also interferes with the natural locomotion and interaction expected in immersive VR experiences, often necessitating the cumbersome attachment of sensors to users' hands or bodies [353]. Although it might be appropriate for a system-automated tour, it may not be suitable for a user-controlled VR tour.

Bio-physiological signals have exhibited substantial efficacy in forecasting emotional states, yet alternative methodologies involving head and eye movement data also present considerable potential [354, 355, 356]. Unfortunately, most consumer-grade HMDs like *Oculus Quest 2* and *HTC Vive* lack eye-tracking facilities, limiting access to crucial emotional analysis data. This limitation has opened a new window to explore strategies for predicting emotions by harnessing available sensor data, such as head and controller movement, inherent in consumer-level HMDs.

Controller-based navigation plays a vital role as a method of travel in user-controlled VR tours, requiring ongoing input from the user [357, 11]. Particularly in the realm of consumer-level HMDs, VR controllers are the primary means by which users navigate the virtual space. Research has established a connection between the user's emotional state and how they interact with the controller's buttons [358, 359]. By utilizing data from various sources, such as head and controller movements, a significant amount of behavioral information can be gathered during navigation. This offers a chance to examine user interactions and possibly infer the user's emotional state, despite the challenges associated with integrating sensors in consumer-level HMDs.

Overcoming the limitations posed by the absence of integrated biosensors and eye-tracking capabilities, this chapter delves into devising a machine-learning model tailored to detect users' affective states while navigating through user-controlled VR tours using the VR controller as an input mechanism. This approach seeks to predict emotional states in VR by utilizing the existing sensor infrastructure, eliminating the need for extra external devices. This provides developers with opportunities to modify VR content to improve user immersion and engagement in VR environments.

5.1 User Study

Taking users' head movement and interaction data with controller during navigation, we intend classify the affective state of a user into any of the following four categories:

- Positive-High (PH)
- Positive-Low (PL)
- Negative-High (NH)
- Negative-Low (NL)

In this work, the terms 'positive' and 'negative' are used to describe the valence of emotional states, while 'high' and 'low' refer to the arousal levels associated with these states. Our selection of affective states is informed by the Circumplex model of affect [360], which organizes continuous emotional states into discrete categories based on arousal and valence. The choice of the Circumplex model was motivated by its widespread use in various applications [361, 362]. Furthermore, the application of the Circumplex model in identifying affective states is a well-established method within the field of affective computing.

5.1.1 VR Stimuli for Emotion Induction

We selected 360 degree videos from database provided by Li et al [363], which contains mean valence and arousal ratings (mean V-A ratings) from 95 subjects. Since the labels of the emotion are given in continuous valence arousal ratings, we converted them into the four affective states as mentioned. To convert valence arousal score to the four categories, a cut-off value of 5 was used, where scores below 5 are categorized as low and scores above 5 are categorized as high. The dataset contains 73 videos and we selected 16 videos (four videos per affective state). We present the list of the videos in table 5.2. We downloaded the selected contents from YouTube with 4K in resolution (38401920) pixels in equiangular cubemap projection format. We used the ffmpeg tool to convert it to equirectangular projection.

5.1.2 Emotionally Neutral Environment for Navigation

To create an emotionally neutral environment for navigation in a VR setting, the design focused on crafting a straightforward outdoor path that users can traverse by pressing the trigger button on their controller. We followed the guidelines suggested in the work by Sophia et al [364] to create the environment neutral. The path was designed to be visually unobtrusive, avoiding any elements that might evoke specific emotions. The color palette chosen for the environment consisted of subdued, natural tones reflective of a typical outdoor setting, steering clear of highly saturated colors or stark, monochrome schemes that might induce particular emotional responses. The textures along the path and its surroundings were rendered to mimic real-world materials without implying deterioration or excessive beauty, maintaining a balance that neither excites nor unnerves the user. Lighting was carefully controlled to mimic a clear, daylight setting, providing ample visibility without dramatic shadows or highlights, thus supporting easy navigation without emotionally charged visual cues. We didnot include any ambient sound in the environment. The interactive element, moving forward by pressing the controller's trigger/joystick button, was implemented to be responsive and intuitive, ensuring the user's experience is centered on simple navigation rather than emotional engagement with the environment. This approach to the VR environment's design and interactivity ensures that it serves as an emotionally neutral space, ideal for focusing on tasks at hand without the influence of underlying emotional tones.

To validate the emotional neutrality of our designed virtual reality (VR) environment, we conducted a pilot study with eight participants (3 females) to assess any changes in emotional state resulting from exposure to the environment. Participants were asked to navigate the straightforward outdoor path in the VR setting for two minutes. We employed The Self-Assessment Manikin (SAM) [365], a non-verbal pictorial assessment technique that measures emotional reaction, to gauge the participants' emotional valence and arousal levels both before and after their exposure to the VR environment. This method allowed us to capture the immediate emotional responses of participants in a nuanced and sensitive manner, without relying on potentially biasing verbal descriptors.

Upon completion of the VR navigation task, we applied the Wilcoxon Signed-Rank Test to the pre and post exposure ratings to check any statistically significant difference in

the emotional states. The results revealed no significant difference ($W = 0.0$ and $p = 0.317$ for both valence and arousal ratings) between the valence and arousal ratings before and after exposure to the VR environment. This lack of significant change in emotional states suggests that the VR environment successfully maintained an emotionally neutral stance, not swaying participants towards heightened arousal or altering their emotional valence. This outcome supports our objective of creating a VR environment that allows for focused navigation without the influence of emotionally charged elements, thereby validating the environment's design as emotionally neutral.

5.1.3 Apparatus

For our virtual simulations, we utilized an *Oculus Quest 2*, which boasts a resolution of 1920x1832 per eye and operates at a refresh rate of $90Hz$. The headset's field of view was set at 89 degrees. Additionally, we enabled the link feature to project the participant's viewpoint onto a desktop monitor for observation. The simulations ran on a computer equipped with an Intel $i7$ CPU at $3.6GHz$ (Quad-Core), $32GB$ of RAM, and a GeForce GTX 1080 with $4GB$ of video memory. The development of these virtual reality simulations was carried out using Unity 3D software, specifically employing its terrain tool for environment creation. The road texture utilized in the simulation was sourced online.

5.1.4 Participants

We asked 17 participants including 8 females (M age= 24, SD=4.8) to take part in the experiment. All the participants were recruited through word of mouth. A brief training was provided before the experiment to let the subjects understand their main task in the VE.

5.1.5 Data Collection Procedure

The procedure was designed as follows:

1. The participants received a brief training session that covered the tasks they would be undertaking and instructions on navigating using the controllers.

2. We utilized the SAM to determine participants' current emotional states, requiring

them to rate their valence and arousal on a 5-point scale.

3. These ratings were then converted to a 10-point scale and correlated with four predefined emotional states (PH, PL, NH, NL). Depending on the participant's current state, they were shown a corresponding 360-degree VR video to amplify their emotional state. For instance, if we found a participant to be in positive high state we showed them a 360 degree video that was labelled as positive high. If we found neutral state we showed them any of the four states.

4. To confirm the emotional state, participants were asked to re-rate their emotions using SAM before introducing them to a neutral VR environment.

5. Participants navigated the neutral VR setting using either a joystick or a trigger for 1 minute, with the control method assigned randomly to prevent any practice effects. Each emotional state was explored in four sessions: two with a joystick and two with a trigger. The participants were made aware that they had to walk on the VR environment for 1 minute, regardless of their speed of travel or the distance covered.

6. We again asked them to rate their emotional state using the SAM to ensure that the emotional state is not effected. We planned that if we find change in emotional state, we would not consider that data and redo the iteration. However, we didnot face such case in our experiment. Participants were also asked they could quit the experiment if they find any discomfort.

7. Steps 4 to 6 were repeated three times, yielding four sets of data per emotional state: two from joystick navigation and two from trigger navigation.

8. The participants were requested to contact us when they were in a different emotional state than the current. So that we could collect data for the remaining emotional states.

In the whole procedure for collecting data for one emotional state, the participants were exposed to the VR environment for 16 to 20 minutes. So, for collecting data for all the four emotional state took around 64 to 80 minutes for one participant. To avoid bias in the interaction behaviour, we did not reveal the type of data we are collecting untill the end of the experiment.

Figure 5.1: Participant during the experiment

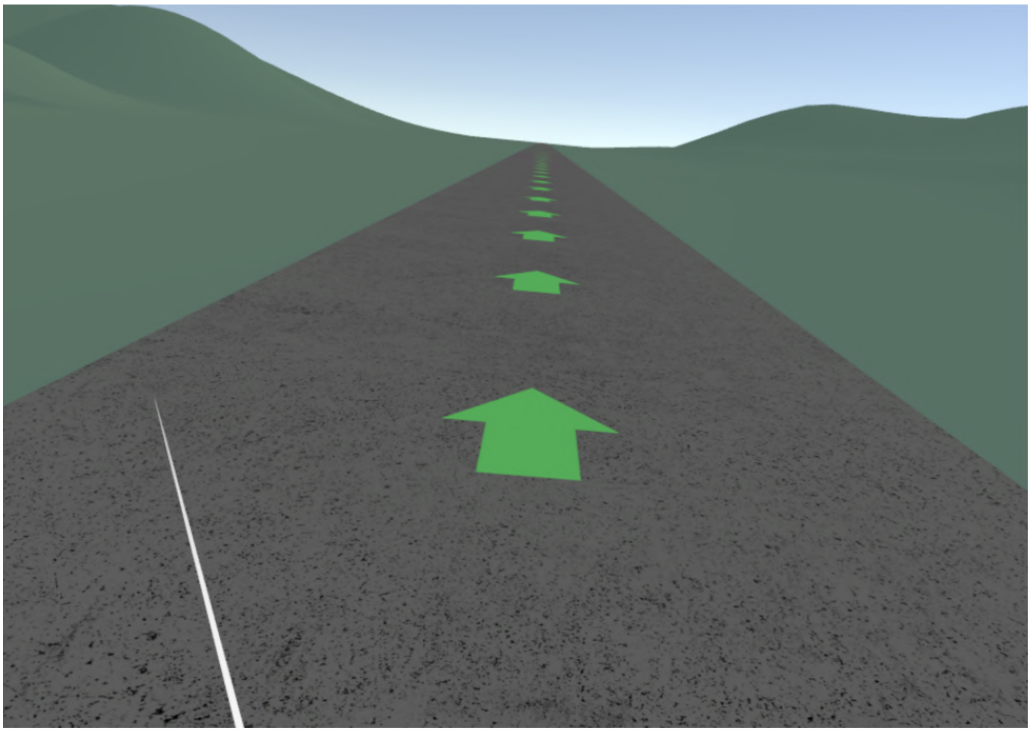


Figure 5.2: Screenshot of the neutral VR environment

Sl No	Type	(Valance, Arousal)	Name	Youtube ID
1	Negative Low	(3.33, 3.4)	Happyland 360	$lWDqeWFJdp4$
2	Negative Low	(2.53, 3.82)	War Zone	$Nxxb_7wzvJI$
3	Negative Low	(2.73, 3.8)	The Nepal Earthquake Aftermath	$5tasUGQ1898$
4	Negative Low	(3.33, 3.33)	Abandoned City	$LbxAc784608$
5	Negative High	(3.2, 5.6)	Zombie Apocalypse Horror	$pHX3U4B6BCk$
6	Negative High	(4.83, 5.25)	Kidnapped	$ywoe0obYaLU$
7	Negative High	(4.4, 6.7)	Jailbreak 360	$vNLDRSdAj1U$
8	Negative High	(4.93, 6.07)	War Knows No Nation	$CIbo0xLbNic$
9	Positive Low	(7.77, 3.92)	Great Ocean Road	$aszTdBlbfq0$
10	Positive Low	(6.13, 1.8)	Mountain Stillness	$aePXpV8Z10Y$
11	Positive Low	(6.57, 1.57)	Malaekahana Sunrise	$-bIrUYM-GjU$
12	Positive Low	(6, 2.63)	Raising Ducklings	$g6w6xkQeSHg$
13	Positive High	(6.46, 6.91)	Walk the tight rope	$JtAzMFcUQ90$
14	Positive High	(6.17, 7.17)	Mega Coaster	$-xNN-bJQ4vI$
15	Positive High	(6.75, 7.42)	Speed Flying	$g6w6xkQeSHg$
16	Positive High	(5.8, 5.4)	Tomorrowland 2014	$j81DDY4nvos$

Table 5.1: Description of 360 degree videos used in our experiment to induce emotional state

5.2 Data Processing

We recorded the data from the HMD and the controller at a frequency of $32Hz$. This data collection enabled us to accumulate 1920 samples for every minute of navigation within the virtual environment. So for each emotional state we had 130560 samples of data ($1920 \times 4 \times 17$). Given that all the recorded values ranged between -1 and $+1$, there was no necessity to normalize these values to prepare them for analysis. The navigation data was segmented based on button press events initiated by the participants. Specifically, we categorized the samples from the start of a button press event at time zero up to the time t seconds when the button was released, as a single window. These windows of data varied significantly in size, ranging from 15 to 362 samples per event.

To standardize the data for our analysis, we adjusted the window sizes to ensure uniformity. For windows containing fewer than 362 samples, we applied padding with the value -999 to each window until it reached the standard size. The choice of -999 as the padding value was deliberate, as it is a value that never occurs naturally within our dataset, thus preventing any potential confusion with genuine data points. Following this,

Data Type	Data
Head-Tracking Data	Head Rotation Quaternion (i.e. x, y, z, and w)
Controller-Tracking Data	Controller Rotation Quaternion while pressing trigger or joystick (i.e. x, y, z, and w) Controller Velocity while pressing trigger or joystick s(i.e. V_x, V_y, V_z) Joystick press data (i.e. (X, Y) in each 1/32th second) Trigger press data (i.e. T in each 1/32th second)

Table 5.2: List of Collected Data from the Oculus Quest 2 HMD & Controller

we appended the label corresponding to the emotional state induced during the event at the end of each row. This process of organizing and standardizing the data was crucial in preparing the sequence data for analysis with the LSTM model, ensuring that our approach was both systematic and consistent with the requirements for sophisticated time-series analysis.

5.3 Proposed LSTM RNN model configuration

The model architecture chosen for predicting the emotional state of a user navigating a VR environment is specifically tailored to capture the temporal dynamics of behavioral data, which includes controller and head movement data. Therefore, we chose to use sequential model. Our model begins with a Masking layer, designed to ignore any input sequence padded with a value of -999. This is crucial for handling variable-length sequences in a batch, ensuring that the model's performance is not adversely affected by padding used to standardize sequence lengths.

Following the Masking layer, the model employs a LSTM layer with a dynamically adjustable number of units. LSTM layers are adept at capturing long-term dependencies in sequential data, making them ideal for analyzing the time-series data generated by users' interactions with the VR environment. The return sequences parameter was set to be true in the first LSTM layer for maintaining the temporal sequence information, allowing the subsequent layers to further process these sequences for more nuanced insights. Dropout and recurrent dropout are integrated within the LSTM layers to mitigate the risk of overfitting by randomly omitting units from the layers during training. This randomness helps in generalizing the model better to unseen data by preventing it from relying too heavily on

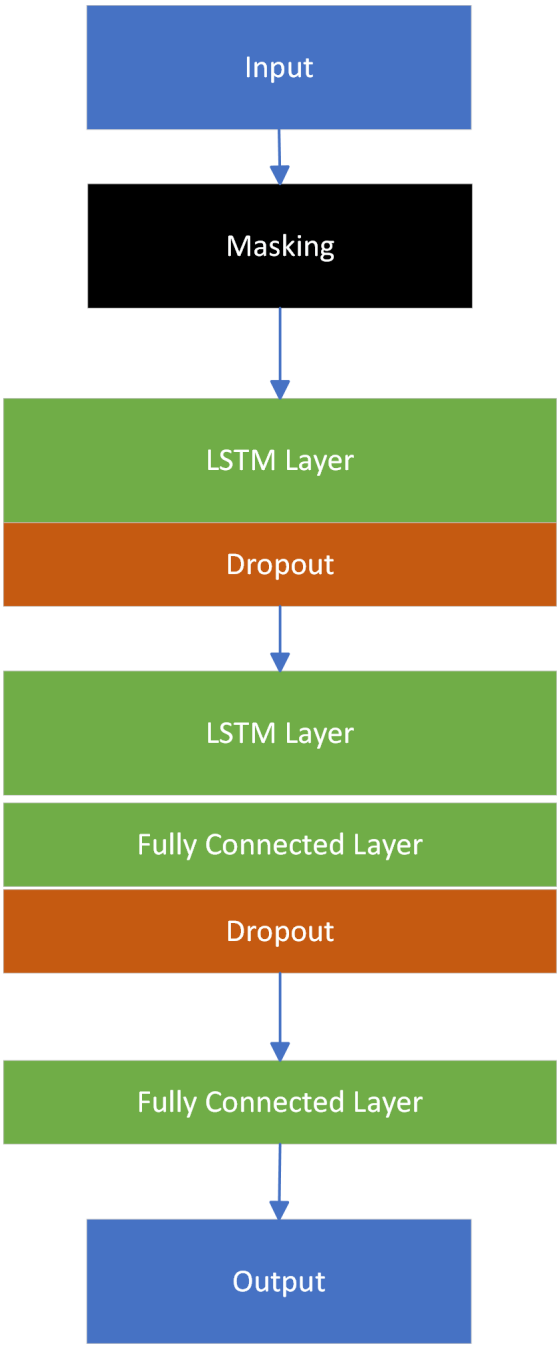


Figure 5.3: Overview of the Model Architecture for Affective State Prediction

any one feature.

A second LSTM layer, with half the units of the first, is included to further refine the model's ability to learn from the temporal data. This layer acts to compress and distil the information from the first LSTM layer, focusing on the most salient features for emotion prediction. The model transitions from LSTM layers to a Dense layer with a *relu* activation function, designed to introduce non-linearity and facilitate the combination of learned features in complex ways, enhancing the model's predictive power. Another Dropout layer follows, reinforcing the model's robustness against overfitting.

Finally, the model concludes with a Dense output layer employing a softmax activation function. This layer maps the aggregated features to the four predefined emotional states (high valence low arousal, high valence high arousal, low valence low arousal, low valence high arousal), providing a probability distribution over these classes that represents the model's predictions. We chose categorical cross-entropy as our loss function (As shown in equation 5.1).

$$L = -\sum_{i=0}^{k}(AS_i \times log(\hat{AS_i}) \tag{5.1}$$

Here, AS_i refers to the actual affective state and $\hat{AS_i}$ refers to the predicted affective state, where k is total number of classes. This loss function is particularly well-suited for multi-class classification tasks, like predicting emotional states, because it effectively measures the difference between the predicted probability distribution over the classes and the actual distribution, where the true class label is represented as a one-hot encoded vector.

This model arrangement leverages the strengths of LSTM layers for sequence data processing, dropout layers for generalization, and dense layers for classification, culminating in a comprehensive approach to predict emotional states from VR behavioral data. We used the same configuration for both the joystick and trigger model.

5.4 Model training and evaluation

We used Tensorflow for training and evaluating our model. We ran our models in an system with $16GB$ RAM with Intel $i7$ $2.6GHz$ CPU with GPU as NVidia Geforce RTX 2080, Windows 11 OS. We used 10-fold cross validation method to train and test the performance. K-fold cross-validation is a statistical technique used to assess the performance of predictive

models by dividing the data into k equal parts. The method involves repeatedly holding out one fold as the test set and using the remaining folds as the training set, thus cycling through all K folds. This approach ensures that every data point is used for both training and validation exactly once, providing a comprehensive measure of the model's predictive accuracy and helping to reduce any bias in the dataset.

In optimizing our models, we utilized the Keras Tuner library, applying its RandomSearch method for efficient hyperparameter tuning. This strategy enabled a strategic exploration of the model's hyperparameter space, focusing on key parameters such as LSTM units, dropout rates, and learning rates. RandomSearch, by randomly sampling possible configurations, offered a pragmatic balance between thorough exploration and computational efficiency. The best parameters found in this approach is presented in table 5.3. We also deployed an early-stopping strategy with a patience value of 10 while training the model, to prevent overfitting. In order to fine tune the model parameters further we used 20% of the training dataset as validation data during each fold. We used Adam as the optimiser for our model with epochs of 10 with batch size of 32.

Table 5.3: Best hyperparameters found after hyperparameter tuning

Hyperparameters	Trigger Model	Joystick Model
Units	320	224
Dropout 1	0.4	0.2
Recurrent dropout 1	0.2	0.2
Dropout 2	0.2	0.4
Dropout 3	0.3	0.2
Recurrent Dropout 3	0.4	0.2
Dropout Final	0.2	0.2
Learning rate	0.00838	0.00136

Table 5.4: Mean accuracy, precision and recall of the 5-fold cross validation on affective state classification

Model	**% Accuracy**	**% Precision**				**% Recall**			
		LN	LP	HN	HP	LN	LP	HN	HP
Trigger	85.30	0.82	0.96	0.92	0.77	0.78	0.77	0.87	0.95
Joystick	86.00	0.73	0.86	0.83	0.93	0.87	0.72	0.91	0.92

5.5 Results

The mean accuracy, precision, and recall from the 5-fold cross-validation are presented in Table 5.4. The proposed models for trigger press and joystick press achieved accuracies of 85.3% and 86%, respectively. The trigger press model demonstrated a higher precision for the class LP (0.96), whereas the recall was higher for the class HP (0.95). Overall, this model exhibited higher precision and recall, particularly for the class HN, with an F1 score of 0.89. For the joystick press model, the class HP achieved the highest precision (0.93) and recall (0.92). The F1 score for this class comes out to be 0.924

5.6 Discussion

The presented results from the 5-fold cross-validation of two the models designed for affective state classification during VR navigation—distinguished by the type of controller interaction (Trigger vs. Joystick)—shed light on several key findings and implications in the realm of emotion recognition in VR settings. The analysis focuses on four emotional states: HP, HN, LP, and LN, with metrics of accuracy, precision, and recall serving as benchmarks for evaluation. The observed accuracies of the Trigger and Joystick models, suggest a relatively similar level of proficiency in classifying emotional states during VR interactions. This similarity in performance is noteworthy, indicating that both interaction modes—despite their inherent differences in input dynamics, provide sufficient and almost equivalent behavioral cues for the models to accurately predict emotional states. The Trigger model demonstrates remarkable precision in predicting LP and HN states, with scores of 0.96 and 0.92, respectively. This high precision indicates that when the Trigger model predicts these emotional states, it is highly likely to be correct, minimizing false positives. This could be particularly beneficial in applications where incorrectly predicting these states could have negative consequences, such as in therapeutic VR environments or in training simulations where precise emotional feedback is crucial. Conversely, the Joystick model shows a good precision for predicting HP emotional states with a precision score of 0.93. This suggests that the Joystick model is particularly adept at identifying genuine instances of HP emotion, which is essential for applications aimed at enhancing user enjoyment or engagement. The lower precision in the Trigger model for HP states (0.77) indicates a

greater tendency for false positives in this emotional category, which could be a focus area for model refinement. The Trigger model shows strong recall in HN and HP states, indicating that the Trigger model is particularly effective at identifying nearly all instances of HN and HP emotions, making it valuable in applications where missing such emotional states could have significant consequences. the Joystick model exhibits strong recall for LN and HN states. This performance suggests that the Joystick model is particularly adept at detecting a broad spectrum of negative emotional states. Given the established link between negative emotional states and the exacerbation of cybersickness symptoms, the Joystick model's proficiency in recognizing these states positions it as a potentially powerful tool in predicting and, consequently, mitigating the onset of cybersickness in VR environments. By identifying users who are experiencing or are likely to experience negative emotions, VR systems can proactively adjust the virtual experience to alleviate potential triggers of cybersickness, thereby enhancing user comfort and overall experience. To the best of our knowledge, no labeled dataset currently exists for emotional state classification that combines controller interactions with HMD data. As a result, our findings cannot be directly compared with previous studies. Nevertheless, we observed a comparable level of accuracy in the study by Tikadar et al. [366], who achieved an accuracy of 86.60% in classifying four emotional states based on users' interactions with a smartphone. While the use of EEG technology can yield accuracies exceeding 95% [367], it is important to note that this method requires a cumbersome setup, which can be a significant drawback.

5.7 Summary

This contribution proposes a sequential model using LSTM RNN to predict the affective state (in four classes of circumplex model) of a person while walking in a VR environment using controller based walking, (where each press to trigger button or the joystick button moves the person one step forward). We used head rotation data and controller rotation, velocity and button (trigger/joystick) press data while pressing the controller. We had two models: one for the trigger based navigation and one for the joystick based navigation. The trigger model achieved an accuracy of 85.3% whereras the joystick model achieved an accuracy of 86%. To the best of our knowledge, this is the state-of-the-art model for predicting affective state using controller and HMD data. Since most of the commercial

HMD comes with controller, we believe that this approach can be used to predict affective state in non-lab setup to personalise the experience of user or predict the probability CS in future. Since the ground truth of the data were based on the subjective questionnaire, a more objective ground truth construction is required so that it can predict the range of the affective state. Moreover, in this work we have only considered the data while pressing the button, a more robust model can be built by considering the HMD and controller movement while not pressing the button and fusing the two models (model for while controller button is pressed and model while controller button is not pressed). Moreover, our dataset was built with only 17 number of participants. A more diverse and number of participants can created considering controller based interaction.

CHAPTER

6

Mitigating CS in User-controlled VR Tours

As discussed in chapter 2, CS is a serious concern in both user-controlled and system-automated VR tours. In chapter 4, we discussed about mitigating sickness while maintaining the realistic walking experience for a system-automated VR tour. In this chapter, we explore the same for user-controlled VR tours. Irrespective of the type of tour, the overarching objective of a VLT remains consistent: to facilitate users' navigation and movement within the VR environment in a manner that feels comfortable, and natural, and engenders a heightened sense of immersion [38]. Unlike system-automated VR tours, the prediction of realistic walking speed is not a matter of concern for a user-controlled tour since the control of speed is with the user, and the user can choose the speed with their comfort. Although achieving complete perceptual equivalence between real and virtual worlds remains a challenge, there exists a subjective threshold where interactions in VR feel inherently natural, thus fostering an illusion of reality within the VR environment [368]. A method to augment the sense of realism within VR entails the integration of natural interaction techniques [369]. Researchers have examined a variety of natural interaction techniques in the realm of VR. These techniques encompass gesture tracking [370], physiological inputs, including monitoring heart rate and galvanic skin response [371, 372, 373]. Researchers have explored indirect physiological signals such as eye gaze and muscle fixation as potential means of interaction within VR environments [372]. Furthermore, breathing has been examined as an active input method, notably by Sra et al., demonstrating its capacity to enhance user presence and serve as a natural interaction mode in VR environments. However, the use of breathing for navigation within VR has been less explored, with a few

notable exceptions that underscore the potential of this approach.

Among the categories of VLTs for user-controlled VR tours, steering-based VLTs encompass a variety of continuous input methods facilitated through bodily movements involving the head, hands, joysticks, torso, or leaning [145]. These techniques are typically implemented using sensors integrated into the HMDs or other sensing platforms. The static posture of users while employing these methods enables them to navigate substantial virtual distances within limited physical confines and with minimal physical strain. However, like system-automated tours, continuous motion in a stationary setup can induce CS due to sensory mismatch [374]. As mentioned in chapter 2, recent investigations have indicated a potential solution in CDB, demonstrating its efficacy in mitigating the onset of CS during VR exposure [43, 302]. This technique, known for its regulatory impact on physiological responses, has shown promise in alleviating the discomfort associated with exposure to VR. The integration of breath-based interaction within VR environments not only showcases its potential to diminish CS but also underscores its suitability as a harmonious and intuitive means of engaging with virtual spaces [51, 375, 376].

The integration of mapping inhalation and exhalation to movement in virtual reality was first accomplished by Davies and Harrison [377]. Here, they used breathing and balance to allow users to fly through a VR environment. Another work, called DEEP [378], utilizes VR to create an immersive underwater world where players navigate by controlling their breathing, turning deep and slow breaths into a unique method of movement within the VR environment. Similarly, [379] merged an underwater VR experience with breathing strategies to aid in stress reduction. Several works also utilize breathing mapped to vertical axis movement, which can be considered a form of navigation in virtual reality environments [380].

Inspired by the positive impact of CDB on CS and the historical use of breath for navigation and other interactions in VR, we introduce a new VLT called *"BreathWalk"*. Using this technique, users can navigate inside a VR environment by utilizing their breath, captured through a nose-attached microphone, while remaining seated on a swivel chair. Adjusting their breathing frequency, from low to high, allows users to modulate their walking speed, while steering direction is achieved by directing their torso accordingly. To track the torso, a VR controller (the non-dominant hand) is mounted on the lower chest of the user.

Our approach acknowledges and extends the foundational work in literature by utilizing breathing not just for vertical movement or abstract navigation but as a direct method for walking simulation in VR. This method maintains a relatively static body while continuously altering the optical view, akin to continuous locomotion techniques in VR. Leveraging breath as an input mechanism not only aims to mitigate CS but also seeks to enhance realism within the virtual space.

We also conducted a comparative analysis between BreathWalk and Joystick-based steering, assessing factors such as cybersickness, realism, task performance, and participant preference. Participants navigated through a one-way route within a small virtual town using both Joystick-based steering and BreathWalk. The assessment revealed that BreathWalk significantly reduced CS compared to joystick-based steering while enhancing the sense of realism, and it was predominantly preferred by the participants. Our research contributes to the field of locomotion studies by introducing a breathing-controlled locomotion technique for walking in a VR environment. Furthermore, our user study provides valuable insights into the advantages and limitations of employing breathing as an input method for locomotion within VR environments.

6.1 Design & Implementation

While navigating through a VR environment using a steering-based locomotion method, we exert control over two fundamental aspects: speed and direction. In our approach, we intricately linked the control of speed to the user's breathing rate. Simultaneously, we harnessed the natural positioning and movement of the torso to provide intuitive and responsive control over the direction of movement. This method was taken out of collaborative brainstorming sessions within our laboratory, where we evaluated various potential methods for virtual movement among ourselves. After various considerations, we settled on the following intuitive actions for facilitating walking within a virtual space:

- **Start and Stop Walking**: Traditional methods of initiating or halting movement in VR environments often rely on manual controls or gestures. However, in our case we recognized the inherent limitation in halting a continuous process like breathing to

stop walking. To address this, we proposed the *Nasal Rapid Gust* technique. This technique involves a quick, powerful burst of air expelled from the nose. This action allows users to start or stop walking within the VR environment.

- **Speed Control**: The crucial component of our approach is the development of a breathing-based input system, designed to finely tune the user's walking speed within the VR environment. We implement a direct correlation between the frequency of breathing and the user's pace of walking. The system allows for a fluid and adjustable range of speeds, varying from 1.4 m/s to 3 m/s. These speed variations correspond to the range of breathing rates, from 5 breaths per minute up to 20 breaths per minute. Although these speeds are the virtual speed, however, it is proved in the literature that the perceived speed is slower than the shown virtual speed. Considering a gain of 1.74, found in a study by Nilson et al. [39], the perceived speed for the virtual speed would be 0.8 m/s to 1.72 m/s. This range is the approximate walking speed of a person from slow to fast walking speed found in a study by Elaine et al[381]. To make sure the speed matches with the realistic walking speed we verified it with five people. Another reason for setting the maximum speed at 3 m/s was to mitigate the potential confounding effect of increased speed on the rise in CS levels, as identified in the study by Lo et al. They discovered that the severity of nausea and vection escalated when speeds increased from 3 m/s to 10 m/s, indicating a positive correlation. To map any breathing frequency in the range of 1.4 m/s to 3 m/s, we used the following linear equation:

$$y = 0.10666667x + 0.86666667 \tag{6.1}$$

  Where, y is the virtual speed and x is the breathing frequency in breath per minute.

- **Direction Control**: In order to implement a intuitive directional control mechanism within a VR environment, we utilized the natural correlation between the orientation of the torso and directional control (torso-directed steering[382]), finding it effective for our movement method. Implementation involved the placement of the VR controller (used for the non-dominant hand) in the lower chest area. The positioning of the controller is illustrated in the Figure 6.3 for reference. The walking path is determined by the torso's orientation, tracked using this controller.

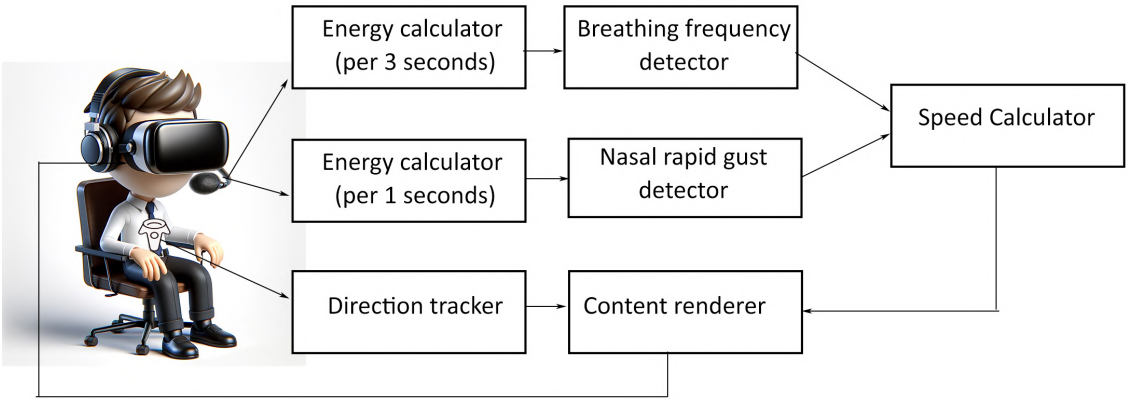


Figure 6.1: BreathWalk System Overview

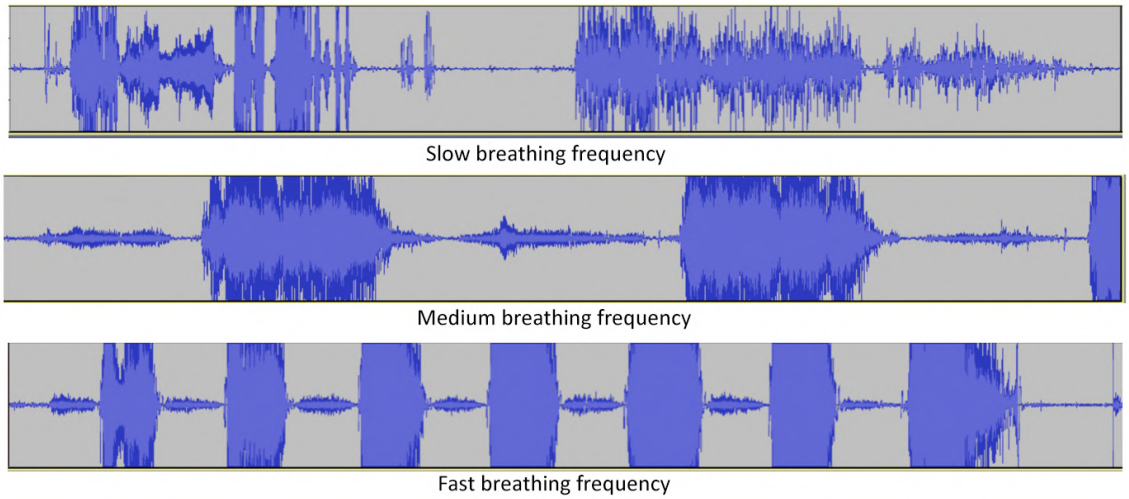


Figure 6.2: Visualisation of raw sound wave for slow, medium, and fast breathing patterns performed by one of our lab members

As depicted in the Figure 6.1, breathing sound is captured through a microphone. We captured the samples at a frequency of 48,000 Hz, aligning with the common practice in audio applications employing a sampling rate of 48 kHz. However, to synchronize with the Unity engine's update function (Unity 2022.3.3f1), which operates at a rate of 120 times per second, we performed amplitude averaging every 400 samples. To ascertain the breathing frequency, we opted to utilize the energy value of the sound signal, which entails summing the squared signal values, representing the overall energy within the signal.Various window sizes were experimented with for computing the energy value, and a minimum window size of 3 seconds was found to yield approximately 70% accuracy in detecting the breathing rate.

Given the variability in breathing intensity across individuals and microphone positioning, we implemented a calibration procedure at the start of each VR session. During calibration, users are guided to inhale and exhale at different breathing rates using a metronome, and the energy values are calculated over 3-second intervals. After collecting these energy values, a linear regression equation is derived to represent the energy values for that session. Additionally, during the calibration phase, data on the average amplitude of nasal rapid gust action is also collected per session. Unlike the breathing rate detection, we found the window of 1 second data to be good enough for detecting the nasal gust action. Detection accuracy for this action was found to be 100%. Figure 6.2 illustrates the raw waveform data recorded by the microphone, showcasing slow, medium, and fast breathing rates demonstrated by one of our team members. For torso detection, we selected to employ a VR controller attached vertically to the lower chest of the user, with the controller's front side indicating the forward direction. Users can alter their direction by rotating their torso while seated in a swivel chair. This concept draws inspiration from Zielasko et al.'s study [382]. It is essential to note that neither the technique for breathing rate detection nor the method for detecting the torso is our original contribution. Our original contribution lies in the utilization of breath-based walking locomotion as a technique.

6.2 Evaluation Study

6.2.1 Participants

Twenty individuals participated in this study, including five females, with ages ranging from 24 to 42 (mean age: 28.15, SD: 4.09). Twelve participants had prior VR experience in our lab. All were familiar with joystick steering but had no experience with breath-based mechanisms. Participants had normal or corrected vision and no nasal congestion during the experiment.

6.2.2 Apparatus

The breath input was captured by a microphone (Logitech H111 Wired On-Ear Headphones with Mic) positioned near the user's nose, as depicted in Figure 6.3. We employed the Oculus Meta Quest 2, which features a resolution of 1920 x 1832 per eye. The HMD was

Figure 6.3: Mic and controller placement for BreathWalk

connected to a PC via an enabled Oculus Link. The stimuli were rendered on a 64-bit computer (Windows 11 Professional OS), equipped with a 3.6 GHz Quad-Core processor and a GeForce GTX 1080 4 GB graphics processing unit. The participants were seated on a swivel chair during the experiment.

6.2.3 Measures

Relative Simulator Sickness Score: Before and after the study, participants were assessed for symptoms of sickness using the SSQ [349]. To calculate the Relative Sickness Score for each participant, we adhered to methodologies outlined in previous studies [40, 72]. This calculation involved deducting the total SSQ score recorded after the test from the one recorded before the test, providing an assessment of the participant's reported level of sickness.

Task Performance: We followed a similar approach used in the study by Zhao et al. [383]. This measure has three main components: Completion time, Speed and Accuracy. We tracked the duration it took participants to complete the designated path. To monitor their movement within the VR environment, we recorded their positions at each second. We also assessed the accuracy of the two locomotion techniques. This assessment involved comparing the actual travel paths of the participants with a predetermined standard path, which was centrally aligned on the road. This road was divided into five straight paths and four curved path sections. For each participant's location, denoted as (x_{user}, y_{user}), we identified the nearest point on the reference path, denoted as (x_{ref}, y_{ref}), to determine their

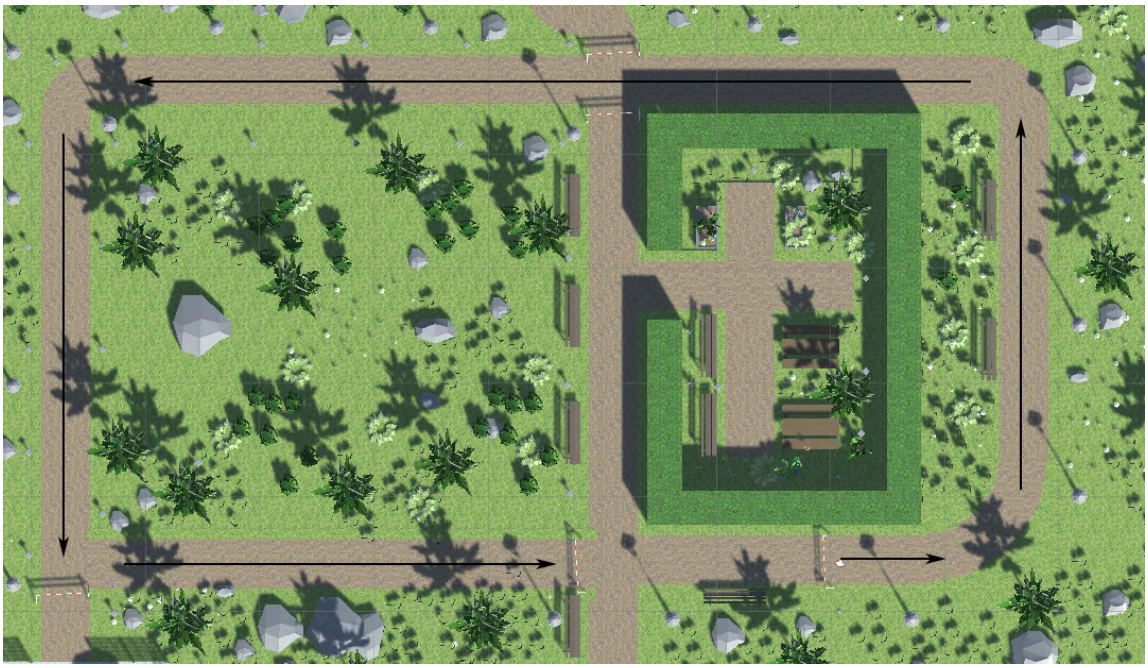


Figure 6.4: Top view of the virtual environment

deviation from the standard path. This deviation at each point was quantified using the Euclidean distance formula:

$$d = \sqrt{(x_{\text{user}} - x_{\text{ref}})^2 + (y_{\text{user}} - y_{\text{ref}})^2} \tag{6.2}$$

Subsequently, we summed these individual deviations across each segment and calculated the average deviation for each participant's journey.

Realism Another measure used in the study was perceived realism. Participants were asked to rate the similarity of the walking experience in the virtual environment to that in the real world. They rated this on a scale from 0, indicating completely unrealistic, to 10, signifying completely realistic. This method of assessment is a modification from the original Presence Questionnaire (PQ) and has been similarly employed in various VR locomotion studies. The modification of the measurement criteria was informed by observations from Pang et al.'s [40] study.

Preference We employed a preference rating scale ranging from 1 to 10 to compare participants' preferences of walking techniques between BreathWalk and the Joystick-based locomotion at the end of using both techniques. This scale allowed us to quantitatively

assess which method was favored by participants, providing a clear comparative analysis of the two locomotion techniques. We also asked them to provide a short explanation about their preference.

6.2.4 Hypotheses

Drawing upon existing literature and theoretical foundations, our formulated hypotheses are articulated as:

- H1: It is hypothesized that the implementation of BreathWalk will result in a reduction of CS experienced by users in the VR environment [43].
- H2: Utilizing BreathWalk is conjectured to enhance the perception of realism during navigation within the VR landscape [51].
- H3: BreathWalk is expected to yield a similar speed and completion time compared to joystick-based steering (Since both techniques had the same range of speed).
- H4: BreathWalk is expected to have low accuracy compared to joystick-based steering [384].
- H5: The incorporation of BreathWalk is anticipated to positively influence user preference within VR settings [51].

6.2.5 Procedure and Tasks

The experiment was of three parts: Pre-task session, task session and post-task session.

Pre-task session Before starting the task, the participants had to go through a practice session to adapt to the locomotion techniques. For the BreathWalk locomotion, we had the additional calibration step. We asked them to practice with the locomotion techniques till they did not have any doubt about it's working principle. The practice session for each technique (excluding the calibration phase) took around 3 to 4 minutes. After practising the locomotion technique, they were asked to fill the SSQ.

Task session The experimental setting consisted of a one-way route that traversed through a virtual town created in Unity (a top view presented in Figure 6.4). We imported the scene from unity asset store (City pack by 255 pixel studio). This route included four turns.

Each participant was required to use the two locomotion techniques to navigate the same route. During the session the participants were asked navigate through the road. While using BreathWalk, they were asked to use diaphragmatic breathing and they were asked to navigate as precisely as possible in their desired speed that feels natural to them.

Post-task session After completing the path using a locomotion technique the participants had to fill up SSQ and rate realism.

To mitigate any potential practice effects associated with these techniques, the order of the locomotion technique was chosen randomly. Participants were given a 30-minutes break before using the next locomotion technique to relieve possible after-effects from the previous locomotion technique. They could also rest longer if requested. At the end of trying out both locomotion techniques they were asked to rate their preference.

6.3 Results

We have analyzed the data gathered from the participants regarding VR sickness, realism, task performance, and preference from the study (as depicted in Figure 6.6). Our findings are reported as follows:

VR Sickness: We performed Shapiro-Wilk tests, which revealed that the distributions of Relative CS Score were not normally distributed ($p < 0.05$). To further compare the two techniques, we employed a one-tailed Wilcoxon Signed-Rank Test, the results of which are elaborated in Figure 6.6a. The test yielded a Z-value of -3.662 with a highly significant p-value of .00013, confirming the significance of our findings at $p < .05$. Additionally, the W-value was found to be 4, substantially below the critical value of 53 for our sample size at $p < .05$.

Realism: For the realism data, the Shapiro-Wilk test revealed non-normal distributions ($p < 0.05$). We applied a one-tailed Wilcoxon Signed-Rank Test, with the detailed results outlined in Figure 6.6b. The analysis yielded a Z-value of -3.4078 and a p-value of .00032, indicating a statistically significant difference at $p < .05$. Furthermore, the W-value recorded was 0, significantly lower than the critical value of 30 for our sample size of 15 at $p < .05$.

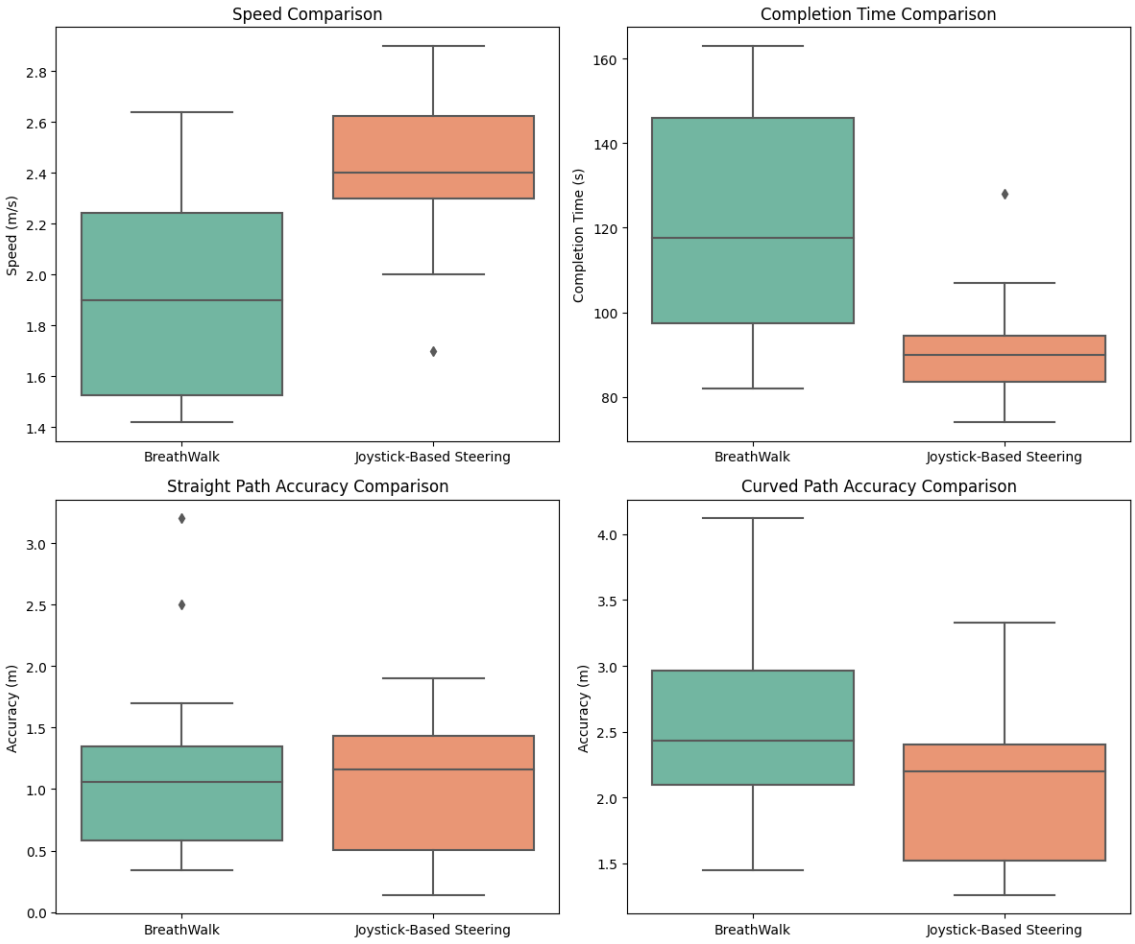
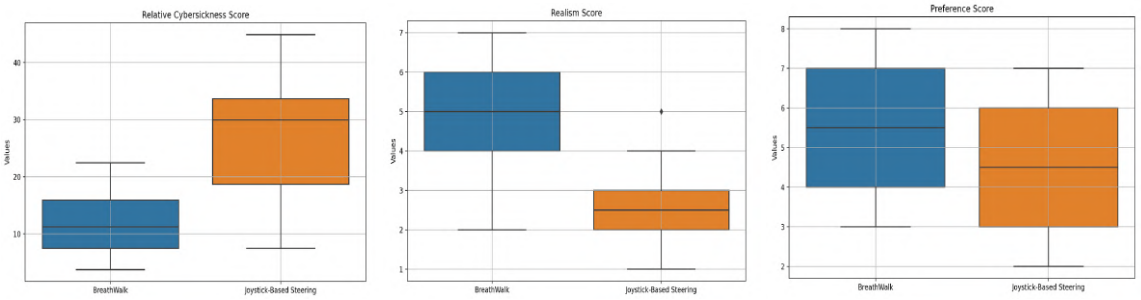


Figure 6.5: Task performance scores



(a) Relative CS score (b) Realism score (c) Preference score

Figure 6.6: Subjective scores

These results, demonstrate a notable difference in the perceived realism between the two locomotion techniques, as gauged by the participants' ratings.

Task Performance: In our study, task performance was evaluated using three metrics: Speed, Completion Time, and Control Accuracy for both straight and curved paths.

For *Speed*, the Shapiro-Wilk test showed the distribution to be not normal ($p < 0.05$). Therefore, we conducted a one-tailed Wilcoxon Signed-Rank Test. The results from the test revealed a large significant difference at $p < 0.05$ (Z-value $= -3.2479$, p-value $= .00058$, W-value $= 18$).

Regarding *Completion Time*, the Shapiro-Wilk test showed that the data were not normally distributed. The one-tailed Wilcoxon Signed-Rank Test yielded a Z-value of -3.1188 and a p-value of 0.0009, indicating a significant difference at $p < .05$. Additionally, the W-value was 17.5, below the critical value of 53, indicating a significant difference.

For *Control Accuracy*, we applied the Shapiro-Wilk test to the deviation data from both straight and curved paths. The straight path data were normally distributed ($p < 0.05$), while the curved path data were not ($p > 0.05$). Consequently, we used the Wilcoxon Signed-Rank Test for the straight path data, which showed no significant difference (Z-value $= -0.5413$, p-value $= .2946$, W-value $= 90.5$). For the curved path data, a paired-t test indicated a significant small difference between before (Mean $= 2.5$, SD $= 0.7$) and after (Mean $= 2.1$, SD $= 0.7$), $t(19) = 2.1$, $p = .046$. These results suggest that the locomotion techniques impacted task performance differently on straight versus curved paths.

Preference: Since the preference data were also not normally distributed, we used the Wilcoxon Signed-Rank Test to evaluate participants' preferences between two locomotion techniques. The test results were significant, with a Z-value of -1.742 (p-value of .04093) and a W-value of 45.5, compared to the critical value of 47 ($p < .05$), clearly indicating BreathWalk to be preferred over the joystick based locomotion.

6.4 Discussion

Our study confirms that BreathWalk significantly reduces CS compared to joystick locomotion, supporting our hypothesis (H1). This aligns with existing research on slow diaphragmatic breathing for motion sickness. While joysticks often worsened discomfort, even rapid BreathWalk (up to 20 breaths per minute) proved more advantageous. We didnot test beyond 20 breaths/minute, but future research can explore the optimal upper limit for BreathWalk frequency. Notably, a few participants reported fatigue with BreathWalk during faster movement (P2, P5, P20). Customizing the BreathWalk speed-frequency relationship to individual comfort could address this and was not tested in our experiment.

BreathWalk demonstrated a significantly enhanced sense of realism compared to joystick-based locomotion, in aspects of both user experience and immersion, aligning with our second hypothesis (H2). The increased realism experienced by participants with BreathWalk can be attributed to the seamless integration of diaphragmatic breathing with virtual movement, coupled with the intuitive use of breath for locomotion control in a virtual environment. One participant described the experience as, "The act of breathing to move felt instinctive and closely mirrored real-world walking, which amplified my sense of presence in the VE." (P11) Furthermore, the employment of controlled breathing as a means of navigation offered a more engaging and cognitively less taxing experience. This sentiment was echoed in participant feedback, with one stating, "I felt more immersed and connected in the virtual world as my movements were directly linked to my breathing, in contrast to using a joystick." (P2) This enhancement in user experience through naturalistic interaction methods has been observed in similar studies, which reported an increase in user presence and a decrease in cognitive load.

Notably, Joysticks proved significantly faster (against H3), likely due to their familiarity and established efficiency in VR navigation. Users adapted and navigated quicker with

joysticks. Conversely, BreathWalk users sometimes felt fatigued while maintaining high breath frequency for speed (although 20 breaths per minute is considered to be a normal breathing rate [385]), leading many to stick with an average pace. Regarding control accuracy, our results indicated no significant difference on straight paths between the two systems, suggesting that BreathWalk is capable of matching the precision of joystick controls in simpler navigation tasks. However, the joystick exhibited superior accuracy in navigating curved paths. Therefore our forth hypothesis remains partially correct (H4). This might be due to the more intuitive and direct control that joysticks offer for complex maneuvers, compared to the potentially less precise control dynamics of torso directed steering in such scenarios. Few participants also mentioned that there was a problem, when they tried to change the speed suddenly, due to which they had problem during taking turns (P3, P6, P18). While calculating breathing frequency we had to take a window of 3 seconds to analyse the energy value to accurately detect the breathing frequency, which caused this delay in recognising breathing frequency. However, we noticed reducing the window can cause sudden jumps in speed, which reduces the smoothness of the virtual motion. A more sophisticated breathing frequency detection method can be used to solve this issue, as done in the work by Shih et al. [386]. These findings prompt a reflection on the practical applications and potential refinements of BreathWalk. While BreathWalk may not currently match the speed and accuracy of joysticks in more complex navigational tasks, its equivalence in simpler tasks suggests potential in specific contexts. For instance, in applications where immersion and experience are prioritized over speed, BreathWalk could offer a unique and engaging alternative. Furthermore, these results highlight areas for improvement in BreathWalk, particularly in enhancing navigational precision in more complex environments.

In considering the outcomes of this study, it is also crucial to acknowledge potential confounding factors that might have influenced the findings. Participant variability in terms of prior VR experience and physical fitness could have impacted both the subjective and objective measures reported. For instance, individuals with more VR experience might adapt more quickly to a new locomotion method, potentially skewing results towards a more positive reception of BreathWalk. Similarly, differences in physical fitness could affect participants' ability to comfortably control their breathing for navigation, which might influence their performance and reported experience of CS or realism. Additionally, the familiarity of

the BreathWalk method compared to the more familiar joystick control could introduce a bias towards favoring the new experience due to its novelty rather than its inherent superiority. Future studies should aim to control for these variables by incorporating a wider demographic of participants and possibly pre-screening for VR experience and physical condition. Moreover, longitudinal studies could help discern whether the observed effects are sustained over time or diminish as the novelty wears off. Acknowledging these factors is essential for a nuanced interpretation of the results, and they highlight areas for further investigation to isolate the specific contributions of breath-based walking simulation to the VR experience.

Evaluating BreathWalk's dual-component system— torso for direction and breathing for speed—poses a challenge in determining each element's contribution to user experience in VR. While previous research supports the benefits of torso-directed interfaces in reducing CS and enhancing realism, participant feedback in our study suggests that the breathing aspect, particularly for speed control, significantly enhances immersion and reduces CS. Despite the perceived benefits, without comparing to a torso-only condition, the exact impact of each component remains speculative. Future research should isolate and combine these elements to fully understand their effects on VR navigation, informing the design of more intuitive VR locomotion methods.

6.5 Summary

In this chapter, we proposed BreathWalk, a breath-controlled VR navigation method for simulating walking in VR. We also compared BreathWalk, to joystick steering in terms of cybersickness, realism, task performance, and user preference. BreathWalk reduced CS and was preferred for its intuitive nature but had lower speed and precision (especially on curved paths). Notably, BreathWalk was evaluated in a controlled lab setting, maintaining a noise-free environment. In future research, we aim to enhance breathing frequency detection by utilizing a diaphragmatic breathing detection device like Zephyr BioHarness , instead of relying on microphone-based detection. This advancement is expected to eliminate the need for calibration steps, thereby simplifying the setup and use of BreathWalk. Additionally, we are considering the integration of audio and vibrotactile feedback to enhance users' awareness of their walking frequency. We intend to further evaluate BreathWalk by employing Cannavò

et al.'s framework [387] for a comprehensive evaluation.

Chapter

7

Visualising Off-screen POIs in Smartphone-based AR Systems

Up until the previous chapter, we have explored various models and methods aimed at enhancing the travel experience in a VR tour. In this chapter, we will address another aspect of navigation, namely wayfinding. Wayfinding plays a crucial role in both AR and VR to improve the navigation experience. However, as discussed in Chapter 2, we identified a research gap in the visualization of a large number of off-screen objects on a smartphone screen. Addressing this gap is essential for improving wayfinding in smartphone-based AR applications.

While navigating using a smartphone-based AR application, due to the limited field of view of the smartphone camera, only a portion of the surrounding POIs is visible on the smartphone screen, while other parts of the environment remain outside of the screen. As discussed in 2, existing systems developed for off-screen POI visualization have predominantly undergone testing in environments characterized by a limited number of POIs, typically ranging from 5 to 15 POIs [44, 48, 331]. In real-world scenarios, it is important to consider that there can be many more POIs than this. When this happens, we refer to it as a high-density environment. For instance, users could find themselves in crowded areas with numerous nearby shops and establishments. In such situations, the current visualization techniques struggle with visual cluttering as the number of POIs increases [50], thereby impeding the users' ability to perceive the desired information effectively. Consequently,

there is a need for an approach to address the problem of off-screen POI visualization for handheld AR in high-density settings, enabling users to precisely locate and interpret the distance information of the off-screen POIs even when confronted with a large number of POIs.

In this chapter, we propose a handheld AR-based off-screen POI visualizer designed explicitly for high-density environments characterized by a larger number of POIs. Our approach aims to provide users with both distance and direction awareness of off-screen POIs. We introduce a novel technique where 3D arrows representing the POIs are strategically positioned and distributed over the edges of the smartphone screen, thereby effectively reducing clutter in the central area of the device. Additionally, to further mitigate visual clutter, we proposed a method to aggregate the 3D arrows while preserving their individual direction information. To determine the optimal number of POIs for aggregation, we developed a model through a within-subject study. In order to enhance distance awareness, our visualizer incorporates vertically dashed lines attached to the base of the aggregated arrows, with the number of dashes varying based on the distance between the user and the represented POI. We also introduced a distance filter to disambiguate the POIs and improve distance perception. Furthermore, we conducted a comprehensive comparative analysis, comparing our proposed approach with state-of-the-art off-screen POI visualization methods, namely 3DWedge+ and Halo 3D. We found our technique to be more accurate and efficient in high-density conditions.

7.1 Proposed work

Initially, we studied both 2D and 3D POI visualization techniques. We need to visualize a set of POIs onto a 2D display of a handheld device so that the direction of the POIs (along with different altitudes) can be located quickly and accurately without having any on-screen clutter. Moreover, a reasonable estimate of the POIs' distance should be provided to the users. Formally, given a collection of POIs (P), we were required to show the elements in P, located in various altitudes and to visualize the distance information for the elements in P when it has a greater cardinality.

7.1.1 3D arrow as metaphore

As discussed in the earlier section, researchers have come up with many designs that can show the distance and the direction information of the POIs. We chose Arrow 3D in our case. Because of the 3D arrow tip, users find it easier to learn the direction of the POIs [48, 50]. In real life, we use arrows to point at something. So we found it the most intuitive metaphor to point towards a POI. However, a 3D arrow makes it challenging to know the distance when the number of arrows increases with the number of POIs [48]. We have proposed the remedy for this in the later sections.

Design of 3D arrow

We use three-dimensional arrows as a visual cue to indicate an off-screen POI. These arrows are anchored to a designated cell's centre and defined within a sphere with a radius of d. The arrow's head is represented by a cone with a height and radius of d, while the tail is represented by a cylinder with a height of d and a radius of $d/2$ to maintain symmetry within the cell. The tail and head of the arrow are distinctively coloured to enhance visibility when the direction of the arrow is either pointing away from or towards the device screen. The circular patch at the tip of the head is of the same colour as the tail, providing a clear distinction (as shown in Figure 7.2). The arrow is depicted in screen space through the use of perspective projection and is scaled to the size of the cell (d). The arrow's direction is determined using the POI's direction vector, which extends from the device to the POI. The arrow or sphere's centre is relocated to the projected POI's screen coordinate.

Positioning the 3D arrows

As handheld devices (especially smartphones) have limited screen sizes, the metaphors need to be placed so that they do not result in visual obstruction. In the literature, we found that the metaphors were placed in two ways in handheld AR: the central area of the screen and the sides of the screen. As mentioned earlier, showing 3D arrows together in a single place affects perceiving the cue. Furthermore, keeping it in the middle of the screen would occlude the central area of vision. Considering these points, we decided to use the 3D arrows on the borders of the screen (distributing them over the sides) rather than using them in the centre of the screen. Now we will discuss how we distributed the arrows around the screen.

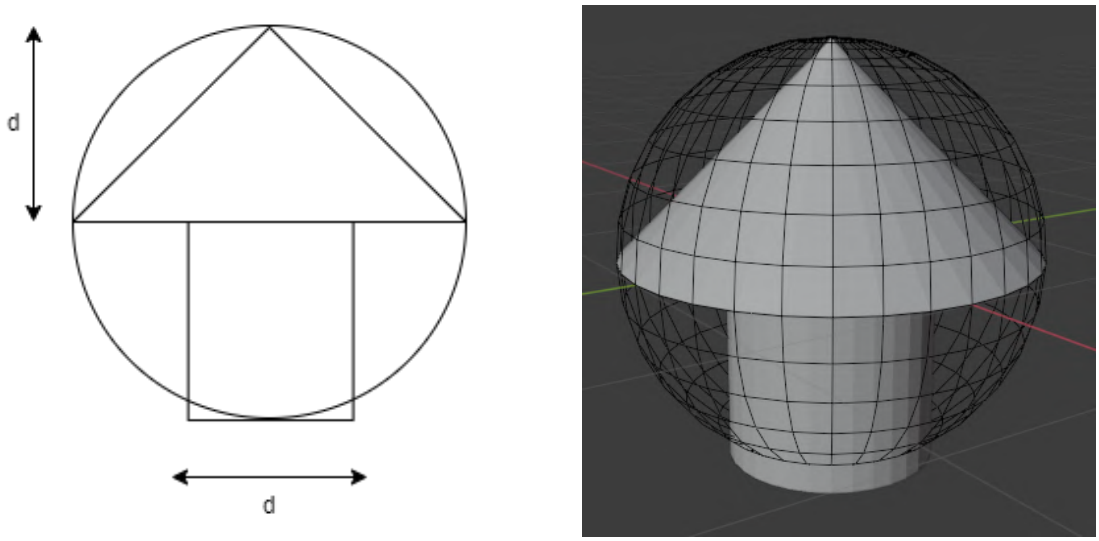


Figure 7.1: 3D Arrow design

Projection In order to effectively visualize the POIs that are located outside of the device's screen, we utilize the AroundPlot technique. This method utilizes an egocentric approach, inspired by the concept of a fish-eye lens, to project a set of POIs, denoted as P, that is situated in a three-dimensional coordinate system surrounding the user. The chosen visualization construct is then positioned at the projected point within the screen space. The position of an off-screen POI is calculated based on its distance from the device and its yaw and pitch values. These values are used to project the POI into one of three distinct zones: Up/Down, Left/Right, and Diagonal.

An illustration of the AroundPlot projection is shown in Figure 7.3. An instance is shown where the user is standing at the centre of the sphere, pointing his/her phone in the forward direction. The blue area indicates the area covered by the smartphone screen. The green and grey areas from the sphere (the off-screen area) are mapped to the smartphone's sides and corners. Note that the POIs with different distances will have spheres with different radii, the user being in the centre of the sphere.

Celling After the projection stage, the POIs are situated in the two-dimensional plane of the device. However, the arrows representing some POIs may overlap and lead to visual clutter. To solve this problem, we introduce *celling*. In this stage, the projected regions (left/right/up/down/diagonal) are divided into cells. The size of the cells depends on the left/right region width or the up/down region height of the device, denoted as d. The

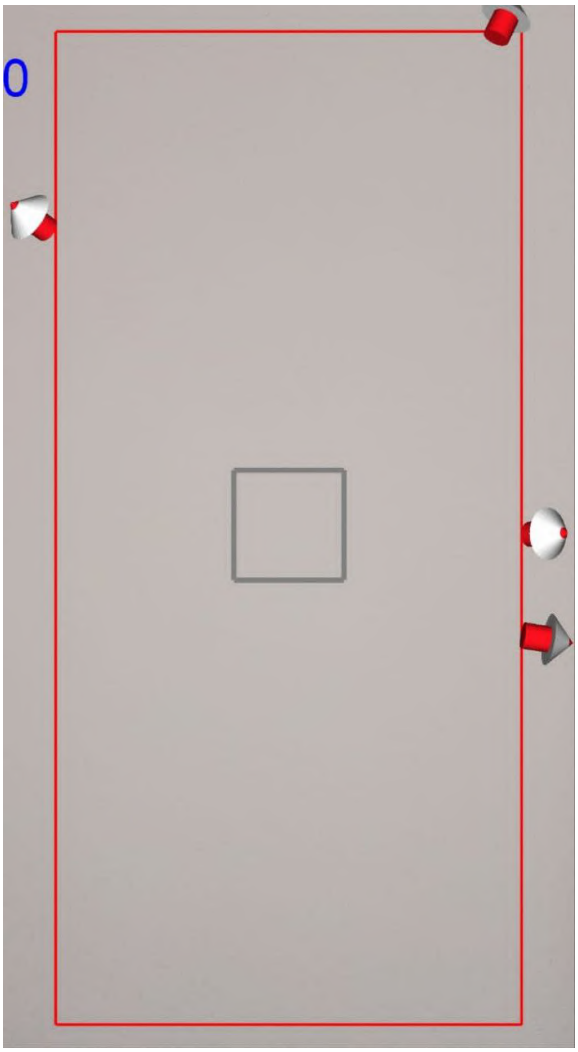


Figure 7.2: Arrow 3D pointing POIs in different directions

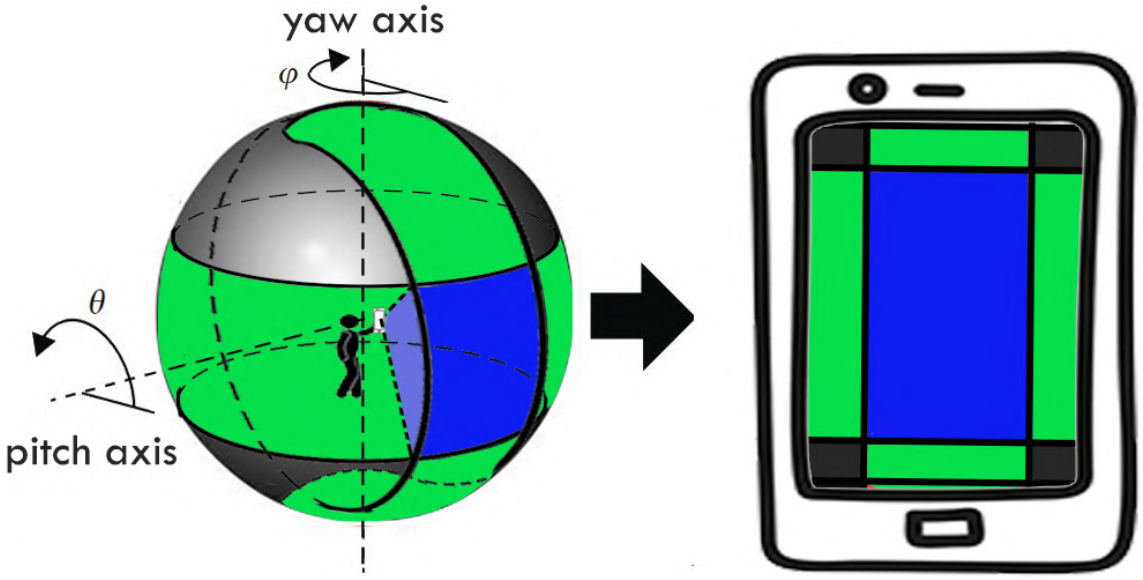


Figure 7.3: Illustration of the projection technique of AroundPlot

left/right direction of the region can be divided into cells by finding the floor value of the height (h) divided by the cell size (d) ($\lfloor \frac{h}{d} \rfloor$). The number of cells in the up/down direction can be calculated by finding the floor value of the width (w) divided by the cell size (d) ($\lfloor \frac{w}{d} \rfloor$). Diagonal directions are considered as a single cell.

We assign a projected POI to its closest cell based on the Euclidian distance from the POIs to the centre of the cell. As a result, the same screen space is used to map several projected POIs that fall within the same cell and are considered overlapped. If the parameters w and h are not multiple of d, the last cell's dimensions will differ from those of the other cells in the area. The projection is then projected to the closest diagonal region in that situation.

7.1.2 Proposed cluttering removal mechanism

A collection of 2D screen space points results from the celling stage. When multiple POIs overlap in the same cell, we need to aggregate the POIs under the same cell to eliminate the overlap. There can be two ways we can visualize the aggregated POIs

1. Using multiple arrows: Each POI is represented by a unique visual component, such as an arrow attached to a central point in the cell's centre.

2. Using combined arrow: A single 3D arrow used as a visual representation of the

centroid of the collection of POIs that are included within the same cell.

As several arrows are attached to the same location in the approach (1), this may lead to a cluttered visual appearance, especially on handheld devices with small screens where each cell only occupies a limited area of the display. Although distributing the arrows into the cells creates less clutter than displaying them together in one place, showing multiple POIs falling into the same cell using approach (1) may nullify the improvement over visual clutter. Ultimately, we chose approach (2). A single 3D arrow (aggregated arrow) is pointed towards the centroid of the projected POIs in a cell. However, the distance information for individual POI mapped into a single cell is lost when only the aggregated arrow is displayed. For instance, we can only show the average distance when two POIs with higher distances map to a single cell. We display thin lines with verticle dashes starting from the arrow's base towards the POIs (as shown in Figure 7.4) to address this issue. We adopted this idea from the 3DWedge+ technique[48]. We map the distance of the POIs with the length of the lines. The line pointing to the closest POI would be the shortest, and the highest distant POI line would be the longest. The number of dashed verticle lines is proportional to the relative distance of the POIs. The colours of the lines are chosen randomly so that the lines can be distinguished from one another. Furthermore, the opacity of the 3D arrow was reduced to 70% so that the dashed lines do not get occluded to the aggregated 3D arrow. The aggregated arrow can be used to follow the direction of the POIs, where the dashed lines are to provide distance awareness to the users. With the increasing number of overlaps in a cell, it may become harder to follow the direction of a specific POI with the aggregated arrow. Furthermore, increasing POIs within a single cell can lead to confusion, even when the dotted vertical lines are present. To mitigate this issue, we remove POIs that might not hold as much value for the user from the display.

We came up with two techniques to disambiguate the POIs from the system:

- Include the POIs in a cell up to a certain limit. When the number of projected POIs goes beyond a specific threshold, it is advisable to employ a recommendation system within that cell. We will discuss this threshold further in a later section.

- Giving control to the user for disambiguating the POIs using a distance filter

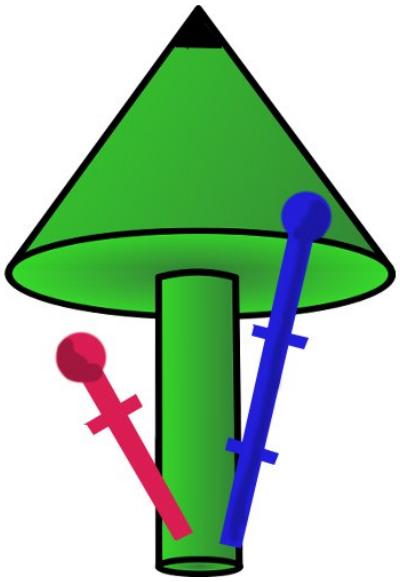


Figure 7.4: Illustration of the 3D arrow with the dashed thin lines pointing towards two POIs

Include the POIs in a cell up to a certain limit: We assume a recommended system is running in the back of the system, which takes the set of POIs as input and outputs a list of POIs in increasing priority order. The idea of using a recommended system is to pick an ideal number of POIs (we call it *Clutter Threshold*, CT) from the cell. Therefore, the recommendation system is employed when the POIs displayed in a cell surpass the threshold CT. If exceeded, top CT POIs are picked from the output of the recommendation system, and others are discarded. Note that we do not propose any recommendation system; we recommend using it to include only the most relevant POIs. However, selecting the most prior CT number of POIs would make it a *unstable system*. By the term unstable, we mean the projected on-screen POIs will appear and disappear with the change in the user's position (when the user navigates). This is due to the system's dynamic nature, as the projection of a POI to different cells can vary based on the device's current position. When a group of POIs are projected to the same cell, the POIs that are part of a *Direction Cluster* are chosen as the highest priority. A direction cluster means a collection of POIs within a certain distance, represented as l, from a direction vector. Therefore, even when the user alters the device's position, all members of a given cluster would yield a cluster of POIs on the screen that are mapped into the same cell. As long as the cluster still contains a POI, the removal of POIs from the cluster will not cause instability in the system.

Let us consider a collection of POIs P mapped to a unit sphere with the device positioned at the centre of the unit sphere. Formally a direction cluster can be defined as a collection of POIs where the largest distance between any two POIs within the group does not exceed l. Multiple clustering algorithms, including K-means[388] and DBScan[389], were evaluated but ultimately we selected a hierarchical clustering algorithm. The algorithm is presented in Algorithm 2, and the distance between clusters (C_1 and C_2) is defined in equation 7.1, with cosine similarity being utilized to define the distance between two points.

$$distance(C_1, C_2) = \max_{c_1 \in C_1, c_2 \in C_2} distance(c_1, c_2) \tag{7.1}$$

Algorithm 2 Algorithm for forming direction clusters

Require: Set of POIs $P = p_1, p_2, \ldots, p_n$ in 3D space and distance parameter l.
1: **Step 1:** Perform unit sphere projection on P and store the result in P_{proj}.
2: **Step 2:** Initialize the cluster set Z with P_{proj}.
3: **while** distance between two clusters in Z is greater than equal to l **do**
4: **Step 3:** Find the two closest clusters, (X, Y), in Z, where $distance(X, Y) \leq l$.
5: **Step 4:** Merge the clusters X and Y and add the result to Z.
6: **Step 5:** Remove clusters X and Y from Z.

The above algorithm comes with a time complexity of $O(n^3)$. However, it need not be performed for each frame. Instead, it can be run offline when there is a substantial change in the POIs' distribution. The value of the parameter l may need to be adjusted depending on the desired cluster rigidity. The value of l is determined based on the cell size on the device screen mapped to the unit sphere.

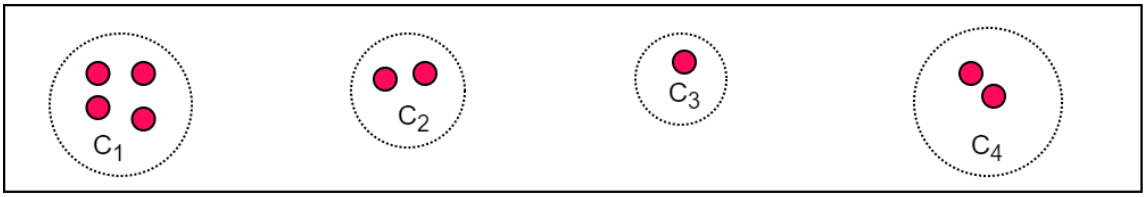


Figure 7.5: Formation of clusters

When a cell holds more POIs than the Cluttering Threshold, the top CT POIs with the highest priority are selected through the recommendation system. However, selecting only the best CT POIs can result in a system that is unstable. Take, for example, the scenario in Fig. 7.5 with four clusters, C_1, C_2, C_3, and C_4. If all POIs are mapped to the same cell in this configuration, the entire cluster C_3 would be eliminated if the POIs in C_3

have low priority. On the other hand, if low-priority POIs were removed from the same clusters C_1, C_2, and C_4, the system would remain stable. As a result, POIs within the same cluster should be removed first, followed by separate POI clusters if necessary.

Algorithm 3 Algorithm for reducing the number of POIs

Require: Set of clusters $C = C_1, C_2, C_3, C_4 \ldots, C_n$, where C_i represents a POI cluster, Cluttering Threshold (CT), and the cell

```
if count of POIs in cell ≤ CT then
    return
Determine the POIs in the cell and assign them to poiCell
Obtain recommendations for poiCell and assign to P
while count of POIs in cell > CT and P contains lower priority POIs do
    Determine the next lowest priority POI from P and assign it to p
    Find the cluster containing p and assign to c
    if count of POIs in cluster c is equal to 1 then
        continue
    Remove p from c
if the total number of POIs in the cell > CT then
    Remove POIs from the cell starting with the lowest priority until the count is equal to CT
```

In Algo. 3, The stage of recommendation might be noted as optimization. The suitable range for the threshold CT will be determined through a user study.

Use of distance filter Another approach we propose to disambiguate the number of POIs from the off-screen POIs is the use of a distance filter. It is useful when a user wants to disambiguate the POIs based on the distance. We use an exclusive distance filter that can be used to adjust the range of the distance radius the user is interested in getting the information.

Earlier, we discussed along with the arrow showing the aggregation of POIs mapped into the cell; we visualize thin dashed vertical lines towards the POIs in order to provide some distance information. This visualization gives us the relative distance information with the other off-screen POIs. However, the user does not get an idea of the distance in units. Using a distance filter solves this problem, and the users get aware of the distance in units. To increase and decrease the scale of the distance bar, '+' and '-' signs are placed, as shown in the Figure 7.8.

Our proposed system is designed to visualize both direction and distance without causing any on-screen clutter.

7.2 Modeling the range of cluttering threshold

The concept of an ideal range for CT implies that maintaining the quantity of POIs within that range in a cell will greatly improve the user's capability to reach a particular POI. We can evaluate the ideal range by checking the time the users take to select a target for different numbers of POIs in a cell. We build a model to predict the value for CT with the estimated selection time.

The exact number of POIs may have different distributions in a cell. The distribution of the POIs was designed using the best, worst, and average situations.

7.2.1 The Cases

We calculate the angle between a targeted POI situated within a group of POIs, which has been assigned to a cell and the centre point of the cluster. When we point the arrow to the centre of the cell, the direction of the arrow representing the cluster can deviate from the target POI by some angular rotation. We classify this situation into three cases:

Best Case: A distribution that results in the minimum, maximum distance between a POI and its centroid is considered the optimal scenario. In this case, all POIs are situated close to one another, making it easier to navigate to a specific POI. Mathematically, the ideal scenario would be if all POIs were pointing in exactly the same direction (but with varying distances). However, this is extremely rare in reality.

Worst Case: The worst scenario is when there is the greatest possible distance between the target POI and the centroid of the POIs (that maps under the same cell). This occurs when, of the n POIs in a cell, n-1 POIs form a cluster, the remaining POI serves as the target, and both the cluster and the target are located at opposite ends of the cell.

Average Case: The average scenario occurs when the target POI is at an intermediate distance from the centroid of the cluster, between the best and worst cases. In this scenario, the distance is neither the minimum nor the maximum but an average of both.

Figure 7.6: Different cases of POI distribution within a cell

7.2.2 Experimental Setup

Participants

The study was conducted with a sample size of 16 individuals, consisting of 8 male and 8 female participants. The average age of the participants was 25.56 years ($SD = 1.54$), and they were selected from the university campus. Although all of the participants had previous experience with handheld devices like smartphones, their familiarity with augmented reality (AR) applications varied.

Implementation details

We built our system using Android Studio and leveraged the motion sensors (magnetometer, accelerometer) of a Poco M2 smartphone device for tracking. To render the graphics, we utilized OpenGLES3.0.

Task Scenario

We asked the participants to visualize themselves as being surrounded by some POIs, like restaurants and shops. They were then tasked with finding the designated target by following the 3D arrow. Five distinct environments were used, each with a distinct CT value. The selected POI was shown in red on the screen. When the target POI was not on the screen, its cell on the map would also be coloured red.

Task

In this study, we wanted to discover the effect of the number of POIs in a cell on the target selection time, thereby developing a model that can predict the (range of) suitable number of POIs with the completion time as input and vice versa. We selected the CT value range as 1 to 8 with increments of 2. So, we used 1, 2, 4, 6, and 8 as the CT values. This range was selected as it provided a good balance between providing enough variation to gather meaningful data and not overburdening participants with too many tasks.

The users were asked to select the target three times for the same CT value. This was done to maintain the accuracy of the reading for each case. However, the repetition was limited to three times to make sure that it does not take much time for the user, which may

eventually decrease the user's level of participation in the tasks and could have potentially led to confusion or frustration. To prevent the influence of the learning effect on the results of the study, a random sequence of tasks was allocated for every CT value. Three cells were randomly picked for each CT value and filled with CT points of interest, each cell containing one target POI.

It was crucial to emphasize that the targets were not displayed all at once to the user. This could have resulted in the user accidentally discovering a target POI while searching for another. Instead, a target POI was revealed only after the user successfully located the previous one.

To familiarize participants with the system, we conducted training sessions where they were given a chance to interact with one POI using our system. The system's fundamental idea and practical application were explained, including instructions on how to choose the POI and any other relevant details. This helped to ensure that participants were comfortable with the system and understood how to use it before beginning the main study.

7.2.3 Results

The study was conducted for each participant individually, with a total of 5 tasks being assigned for each participant, one for each CT value in the set $1, 2, 4, 6, 8$. Each task featured three clusters, best, worst and average, with a target POI placed in each cluster. This led to 15 values obtained per participant (5 CT values x 3 targets). The data points obtained from all participants were averaged to generate a representative plot of the results. The graph can be seen in Figure 7.7. The visual graph presented herein offers a clear and easy-to-interpret representation of the research findings, facilitating the identification of any emerging patterns or trends.

We ran an ANOVA test for the best, average and worst cases to check the statistical significance. We found a significant difference($F = 3.19, p < 0.05$) in the completion time for the best, average and worst cases. After Tukey HSD posthoc analysis, we found a significant difference in completion time ($p < 0.05$) between the best and worst cases. We found no significant difference between the best and average cases or the average cases and the worst cases. Different CT values (number of POIs per cell) for the best and worst cases would result in different completion times. Our aim was to fit more POIs in one cell and reduce

the completion time for target selection. Therefore, we decided to come up with a model that can decide the CT for the best and the worst case, thereby establishing a range of thresholds from worst to best case.

Regression analysis was carried out on the gathered data to obtain a more comprehensive insight into the data and calculate estimates for a specific number of POIs (CT). Linear regression was chosen as the method as it provided the best fit for the data. The model's goodness of fit was determined by calculating the R^2 value, which measures how well the model fits the data. The R^2 values for the best and worst cases were found to be 0.94 and 0.95, respectively, indicating that the model fits the data well in both cases. The equations of the regression line were also determined and are presented as follows:

$$Time_{best} = 328.8 \times CT_{best} + 3470.1 \tag{7.2}$$

$$Time_{worst} = 810.48 \times CT_{worst} + 3391.7 \tag{7.3}$$

In the above equation, $time_{best}$ and $time_{worst}$ are the completion times for the best and worst cases, respectively. 328.8 and 810.48 are the slopes, and 3470.1 and 3391.7 are the intersections for the best and the worst cases. Now, if we place the mean of the completion times for the best and the worst case, we can calculate CT_{best} and CT_{worst}. However, we introduce a tolerance T, a percentage of completion time we may want to compromise over the mean completion times. It is shown in the following equations:

$$CT_{best} = \frac{3867.625 \times (T/100 + 1) - 3470.1}{328.8} \tag{7.4}$$

$$CT_{worst} = \frac{3897.375 \times (T/100 + 1) - 3391.7}{810.48} \tag{7.5}$$

We define the ideal range of CT to be $[CT_{worst}, CT_{best}]$. The range is determined by a tolerance, T, which reflects the required level of performance set by the developer for the user relative to a baseline. The tolerance is specific to the task being performed. If we assume T to be 50%, then we mean we are ready to tolerate the completion time to be 30% higher than the mean completion time of the best and the worst case. So for T= 50%, the range of CT would become $[3, 7]$ (rounded off). Assuming this range of CT will take approximately 5.8 seconds to locate a POI that falls inside the same cell. Note that the

range needs to be rounded off.

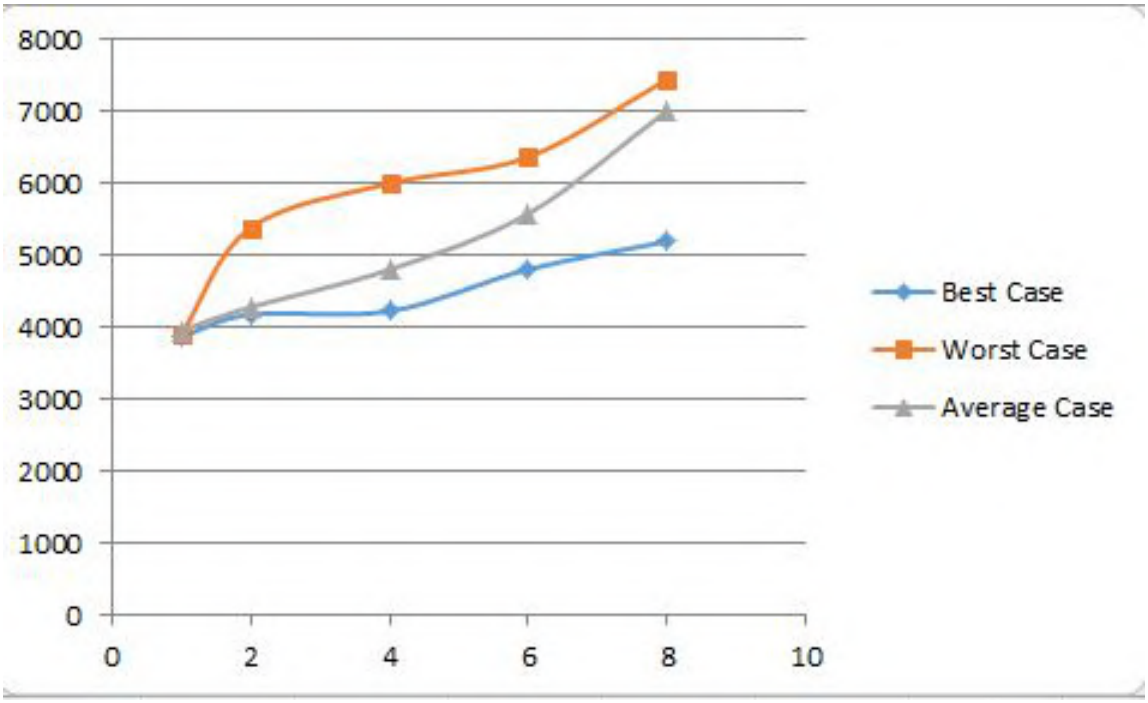


Figure 7.7: graph shows how long it takes to complete a task at different levels of clutter (X-axis is the cluttering threshold value and Y-axis is the time in milliseconds)

7.2.4 Discussion

As presented in Figure 7.7, it was found that the total selection time increased in proportion to the number of POIs per cell. We can also see a noticeable increase in the slopes of the curves representing the best, average, and worst cases. This outcome was to be expected as the direction of the aggregated arrow diverges with an increased number of POIs per cell and an increased distance between the centroid of the POI cluster and the target. This divergence makes it challenging for users to locate the target, resulting in longer selection times. The increase in selection time with the increase in the number of POIs can also be seen in the previous works[325, 390, 48]. The model built for the range of CT is most relevant when we have some recommendation system (like [391, 392]) that can be applied to a cell when the threshold calculated from the model is exceeded. Neglecting the utilization of a recommendation system capable of selecting the top CT number of elements from the POIs located within the same cell can result in the exclusion of POIs with significant interest. However, there is a trade-off between the selection time and the number of POIs

that is important to incorporate. The option for adjusting the CT range can be placed in the system settings so the user can set it to minimum or maximum. Nevertheless, if the placement of the CT value is placed on the main screen (so that it can be adjusted online), it needs to be experimentally understood. We do not recommend using it on the main screen as it may create visual obstruction and will lose the purpose of creating the clutter-free visualization. Note that the model to determine the cluttering threshold range proposed in our work is relevant for 3D arrows. This range may differ for different metaphors and must be empirically established.

7.3 Evaluating the system

The final system includes the 3D arrows distributed throughout the cells aggregated according to the CT value, the vertical dashed lines from the base of the arrows and a distance filter. However, we had to evaluate how the system performs compared to the state-of-art systems. In this section, we talk about our study, where we compared how well our system can estimate directions and distances with two other systems, namely, Halo 3D and 3DWedge+. While comparing, we did not use the recommendation system, as it does not contribute to direction and distance awareness. We were interested in evaluating our system's performance in a high-density environment. However, we chose our CT value to be 7 (Tolerance $T = 50\%$). This means the POIs were distributed in such a way that not more than 7 POIs fall onto the same cell.

7.3.1 Hypotheses

In Halo3D, the distance of a POI can be perceived by observing the curvature of the halo's arcs. The more pronounced the arc's bend, the closer the POI is to the user. Furthermore, these arcs offer a visual indication of the general direction in which off-screen POIs are positioned relative to the user's current viewpoint. However, as mentioned in section 2.1, Halo3D encounters difficulties in distinguishing between the front and rear positions and was not specifically designed to handle POIs at different altitudes. Consequently, accurately locating a POI within the three-dimensional space becomes challenging. Conversely, in the case of 3DWedge+, distance perception relies on vertical dashed lines, while the directions of POIs can be visualized through ball arrows situated at the base of the wedges. Although

a direct comparison between Halo3D and 3DWedge+ is unavailable, it is evident that Halo3D employs a two-dimensional metaphor, unlike 3DWedge. Research studies have demonstrated that three-dimensional metaphors, such as 3DWedge+, enhance direction awareness to a greater extent than their two-dimensional counterparts in handheld AR systems. Nonetheless, in low-density conditions, including our system, all three systems are perceived to provide distance information equally well due to their reduced visual complexity. However, in medium and high-density conditions, both Wedge3D+ and Halo3D have the possibility to create visual clutter. Nonetheless, our system incorporates the decluttered distribution of arrows along with a distance filter, which helps mitigate this issue by reducing visual complexity. Based on these information from the existing literature, we have formulated two hypotheses that outline the expected performance of these systems across three different densities.

- H1: Our system would take a similar selection time and accuracy in low-density conditions in the distance tasks. However, in the direction task, halo 3d would require more selection time since it was not designed for visualizing the POIs with different altitudes.

- H2: Our system will show significant accuracy on both distance and direction tasks in medium and high-density conditions. Due to the visual cluttering, 3DWedge+ and Halo 3D will be less accurate.

7.3.2 Experimental Setup

Participants

We recruited 14 volunteers (10 males, 4 females), ages between 21 to 35 (M=26.11), from our university. Eight of them took part in the previous experiments. All of them already experienced AR applications on their smartphones. However, none were familiar with the Halo 3D and 3DWedge+ off-screen visualization techniques. All the participants had normal/corrected to normal vision.

Implementation details

In this experiment, we employed the Poco M2 device, which was also utilized in our previous study. Our system was implemented using android studio, along with the OpenGLES3.0 rendering framework. As previously discussed, our design included the implementation of celling and aggregation aspects. The Halo3D feature was implemented according to the method described in their paper. The POIs were displayed on the device screen by projecting them onto the plane and presenting them as circular arcs that were shifted into the screen. Additionally, we included the 3Dwedge+ feature, which involved projecting wedges with square bases onto the bottom middle of the screen. As described in their work, we added the dashed lines with ball headers in the middle of the square bases. The height of the wedges was calculated according to the distance between the user and POIs.
In order to know the performance with the increase in density, we chose three density conditions. We took 5, 15 and 30 POIs with different distances to represent the small, medium and high-density environments, respectively. The POIs were placed at random distances (within 1 to 20 meters).

Task details

The main objective of this study was to confirm whether our technique performs better in a dense environment. In the literature, we found the off-screen POI visualizer to be focused on estimating the direction and distance of the off-screen POIs. We chose two types of tasks for direction and distance evaluation.

We paid the most attention to designing the task scenario as realistically as possible. We told the participants to imagine they were in the middle of a marketplace, and there were different places like restaurants and shops on different floors. Now, based on the task, they had to select the targets using a particular visualization technique. To choose a target, they needed to bring the POI onto the box in the middle of the screen (as shown in Figure 7.8, 7.9a,7.9b). Then, they had to tap on the screen to confirm their selection. We asked them to select the POIs as quickly and accurately as possible. The tasks with their descriptions are discussed as follows:

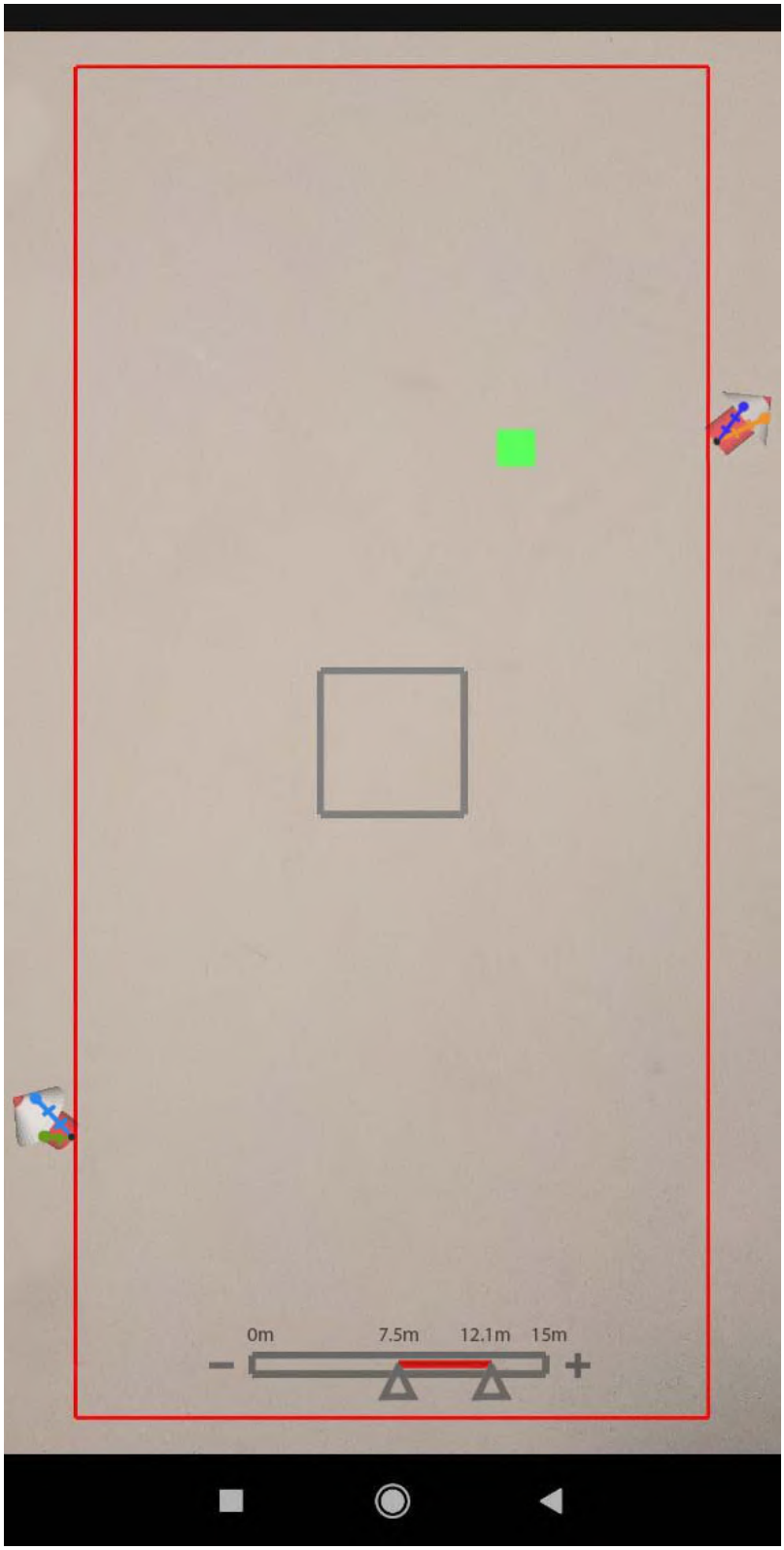
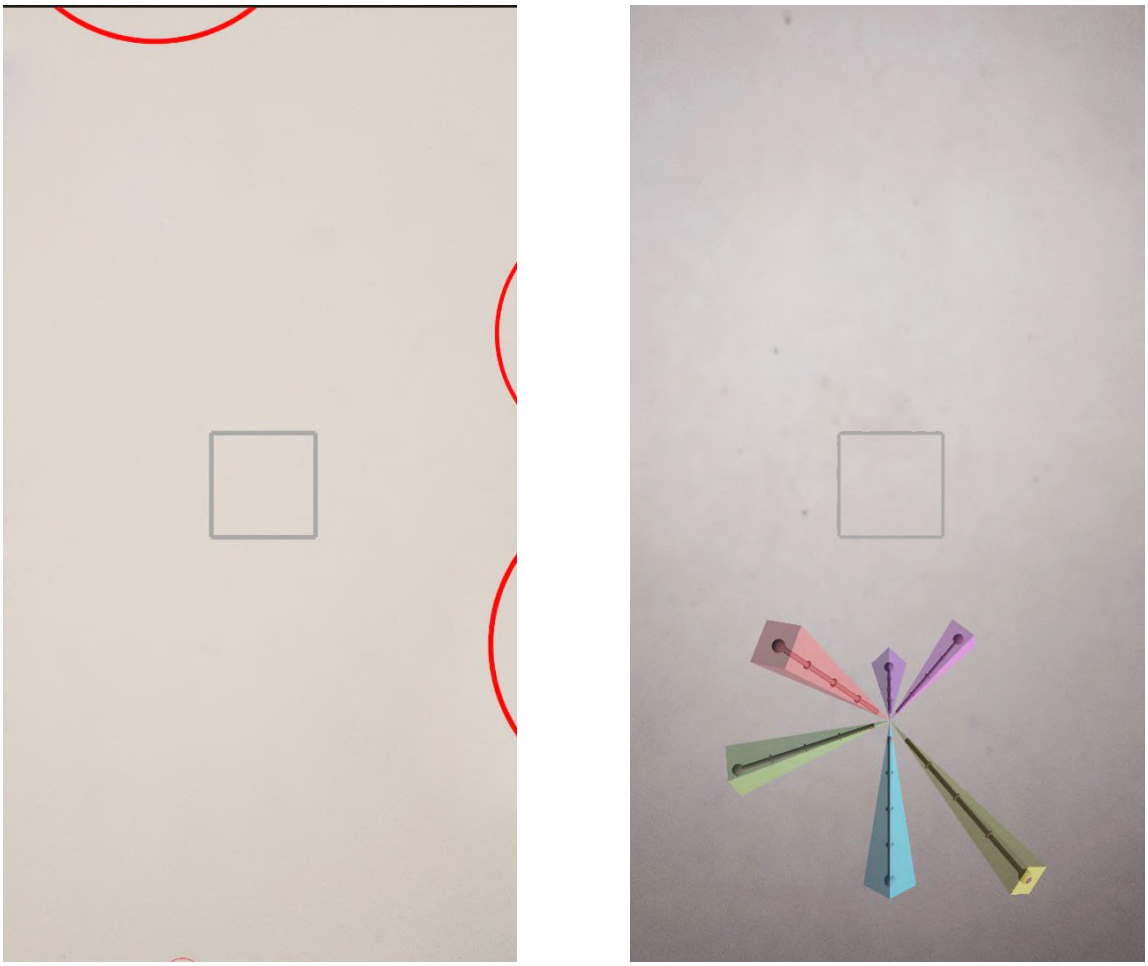


Figure 7.8: implementation of our technique



(a) Halo3D (b) 3Dwedge+

Figure 7.9: Implementation of Halo 3D and 3DWedge+

Direction Task: In this task, we randomly placed the target POIs at different angles around the device. The metaphor pointing towards the target (the dashed line along with the arrow for our technique, the aggregated halo for halo 3D and the wedge for 3DWedge+) was blinked in an interval of 500ms for 6 secs to let the participant realize whom to follow to locate the target. After 6 seconds, we start the timer and let the user locate the target. The participants had to locate the target for each density condition. In order to ensure a fair assessment of the system, the POIs were positioned off-screen at the start of each test. During the evaluation process, we recorded the following:

- Time to Completion: The amount of time, measured in milliseconds, needed to finish the task (select all the targets)
- Error: Incorrect choice of a POI.

From this task, we could determine the ability of the visualization techniques to present off-screen objects in different directions.

Distance Tasks: Tasks for the distance evaluation consists of three tasks, namely closest, estimate and order task. We adopted these tasks from [35] and [48]. These tasks were done for all three density conditions. The detail of the tasks is as follows:

- Closest: In this task, the participants had to select the closest POI from the environment. This is useful when a person wants to visit the closest POI in the marketplace. For this task, we recorded the
  - Time to Completion: The amount of time, measured in milliseconds, needed to finish the task (select all the targets)
  - Error: Incorrect choice of a POI.
- Estimate: In this task, the participant had to select the pair of targets which are closest to each other. With respect to the given scenario, the participants were asked to assume that they wanted to visit two locations in a short duration. So, choosing the closest location can minimize the time required to go from one location to another. As done in the [48], we recorded the

  - Time to Completion: The amount of time, measured in milliseconds, needed to finish the task (select all the targets)
  - Error: Incorrect choice of a Point of Interest.
  - Error rank: This measure helps assign more error weightage on bigger errors (introduced in [48]'s work). It is the rank of the selected POI from the target POI (the error rank of the target POI would be 0). For instance, if a participant selects the fifth closest POI, its error rank would be 4.

- Order: In the order task, the participant had to select five POIs in increasing order of the distance from him/her. In the considered scenario, the participants were asked to assume that they had to plan for a market tour covering five POIs from closest to farthest. With this scenario, we could evaluate if the POIs with different distances could be distinguished precisely. We recorded the
  - Time of completion: the selection time for each object by dividing the number of POIs (density) in that task (in milliseconds)
  - Error: The error was recorded when the participant selected the wrong order
  - L1 Error: To give a higher penalty to a wronger error, we computed the $L1error$ as used by [48]. It is calculated by summing up the differences between the correct and captured orders. For instance, if the correct order of selecting five objects is 2, 3, 4, 1, 5, and the user selected it as 2, 4, 1, 5, 3, the L1 error will be $|2-2| + |3-4| + |4-1| + |1-5| + |5-3| = 10$

For each volunteer, the experiment lasted about 40 minutes. A significant amount of time was given to practice the three visualization techniques. The sequence of density and the techniques were randomly generated for each participant. We received 3(techniques) X 3(density) X 14 (No of participants) = 126 attempts of target selection data for each task. We also had a post-questionnaire to know the user preference. The user had to rate the three techniques in terms of preference on five Likert scales (5 being the best).

7.3.3 Results

The following is a summary of our analysis of the participant's recorded data of the distance and direction tasks (shown in Figure 7.10).

Direction task observation

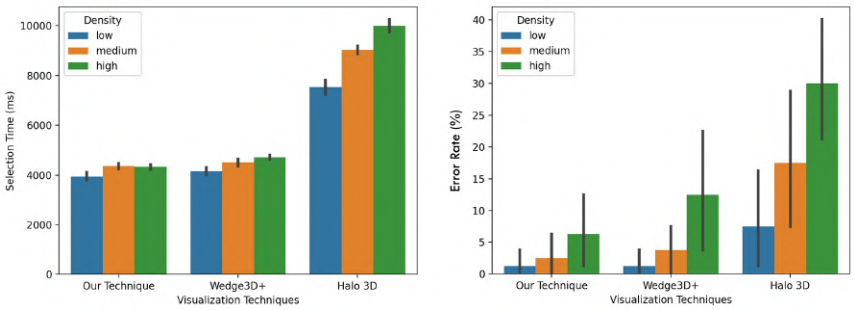
We performed ANOVA on selection time and the error with the direction task for each density condition. For low density, we found significant effects of the visualization techniques on the selection time($F_{(2,45)} = 346.21, p < 0.01$) and no significant difference in the error rate($F_{(2,45)} = 2.43, p > 0.05$) for the direction task. The post hoc Turkey HSD test revealed that our technique($p < 0.01$) and 3DWedge+($p < 0.01$) had a significantly higher selection time than the halo 3D technique. We also found a significant difference in selection time and error for the medium-density($F_{(2,45)} = 1,324.82, p < 0.01$ and $F_{(2,45)} = 5.68, p < 0.01$) and high-density($F_{(2,45)} = 1,359.29, p < 0.01$ and $F_{(2,45)} = 8.22, p < 0.01$) conditions. From the post hoc analysis, we found our technique and 3DWedge+ had significantly lower selection time($p < 0.01$,$p < 0.01$) and error($p < 0.05$,$p < 0.05$) than halo 3D in medium-density conditions. However, our technique had a significantly different selection time ($p < 0.01$) in high-density conditions than the other two techniques. Our technique ($p < 0.01$) and 3DWedge+($p < 0.05$) had significantly lower error rates than Halo 3D.

Closest task observation

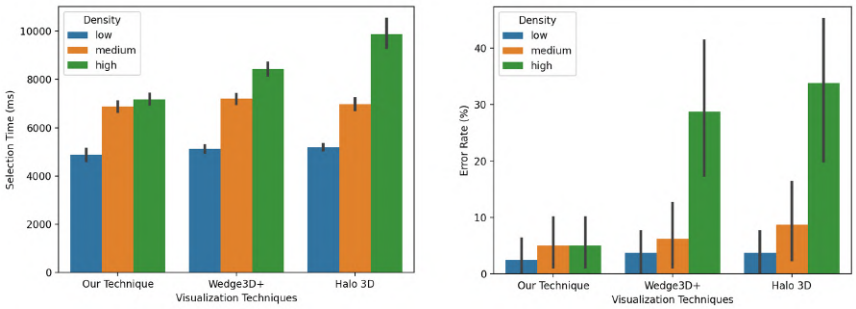
For this task, we found no significant difference in the selection time and error rate for all the visualization techniques in the case of the low-density ($F_{(2,45)} = 3.22, p > 0.05$ and $F_{(2,45)} = 0.14, p > 0.05$) and medium-density($F_{(2,45)} = 2.23, p > 0.05$ and $F_{(2,45)} = 0.40, p > 0.05$) conditions. However, for the high-density condition, we found a significant effect of techniques on the selection time ($F_{(2,45)} = 41.53, p < 0.01$) and error rate ($F_{(2,45)} = 8.43, p < 0.01$). In the post hoc analysis, we found our technique to have less selection time and error rate than 3DWedge+($p < 0.01$ and $p < 0.01$) and halo 3D ($p < 0.01$ and $p < 0.01$).

Estimate task observation

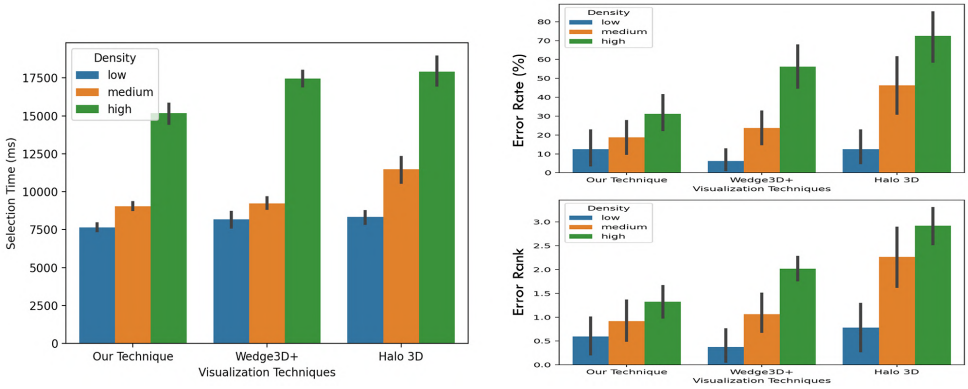
We found no significant difference in the selection time ($F_{(2,45)} = 3.15, p > 0.05$), error rate ($F_{(2,45)} = 0.71, p > 0.05$) and error rank ($F_{(2,45)} = 0.85, p > 0.05$) in the case of the low-density condition. For the medium-density and high-density conditions, we found a significant difference in selection time (($F_{(2,45)} = 23.32, p < 0.01$ and $F_{(2,45)} = 14.95, p < 0.01$)), error rate ($F_{(2,45)} = 6.35, p < 0.01$ and $F_{(2,45)} = 12.92, p < 0.01$)and error rank ($F_{(2,45)} = 3.22, p > 0.01$ and $F_{(2,45)} = 21.95, p < 0.01$). The post hoc analysis revealed that



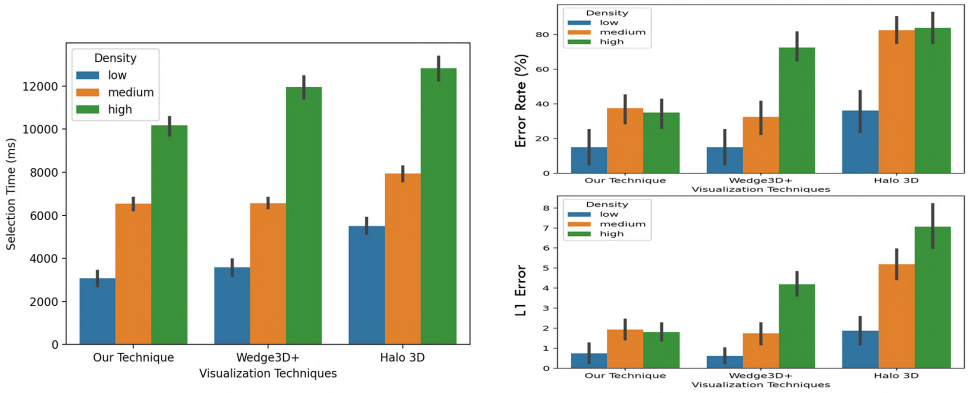
(a) Direction task results



(b) Closest task results



(c) Estimate task results



(d) Order task results

Figure 7.10: Results of the system evaluation

the 3DWedge+ and our technique perform significantly better in terms of selection time ($p < 0.01$ for both the density). Our technique and 3DWedge+ have significantly fewer error rates ($p < 0.01$ for our technique and $p < 0.05$ 3DWedge+) and error rank ($p < 0.01$ for both the techniques) than the Halo 3D in medium density. However, in the case of the high-density condition, our technique had a significantly lesser error rate (($p < 0.05$ for wedge and $p < 0.01$ for halo 3D)) and error rank ($p < 0.01$ for both the techniques) than the other two techniques. There were no significantly different selection times and error rates between the wedge3D+ and halo 3D for the high-density conditions.

Order task observation

In the order task, there was a significant effect on the selection time (Low density: $F_{(2,45)} = 46.39, p < 0.01$, Medium Density: $F_{(2,45)} = 30.49, p < 0.01$, High Density: $F_{(2,45)} = 28.30, p < 0.01$), error rank (Low density: $F_{(2,45)} = 4.97, p < 0.05$, Medium Density: $F_{(2,45)} = 38.72, p < 0.01$, High Density: $F_{(2,45)} = 33.33, p < 0.01$) and L1 error(Low density: $F_{(2,45)} = 5.83, p < 0.01$, Medium Density: $F_{(2,45)} = 36.16, p < 0.01$, High Density: $F_{(2,45)} = 44.50, p < 0.01$) for the visualization techniques in all the density conditions. After the post hoc analysis, we found that for the low-density condition, our technique and 3DWedge+ had a significantly better selection time ($p < 0.01$ and $p < 0.01$), error rank($p < 0.05$ and $p < 0.05$) and L1 error ($p < 0.05$ and $p < 0.01$) than halo 3D. Also, for the medium-density condition, we found significantly better selection time($p < 0.01$), error rank ($p < 0.01$) and L1 error ($p < 0.01$) for our technique and 3DWedge+ than halo 3D. In the high-density condition, our technique ($p < 0.01$) has a significantly lower selection time than halo 3d. However, we found that our technique has a lower error rate ($p < 0.01$) and L1 error ($p < 0.01$) than 3DWedge+ and halo 3D.

User Preference

After collecting the data, we used a post-questionnaire with a 5 Likert scale to know the user preferences. Here, the users had to rate the techniques with respect to the density conditions (5 being the best rating). We asked them the reason for every best and worst rating. We found that for low-density conditions, users rated 3DWedge+(Mean: 4.5) and our technique (Mean: 4.2) to be higher than halo 3D (Mean: 3.4). In medium density,

users gave a similar rating to both our technique (Mean: 4.2) and 3DWedge+ (Mean: 4.05) compared to halo 3D (Mean: 2.33). However, in high density, most of the participants preferred our technique (Mean: 4.8) over the 3DWedge+ (Mean: 3.26) and halo 3D (Mean: 1.66).

The users mentioned that comparing the distances with 3DWedge+ in the distance task in the low-density condition was easier than in the other two techniques. However, with the increase in density, the distance line (dashed line) was hard to be recognized in 3DWedge+(P5, P7, P8). One participant (P1) suggested using the volume button to control the distance filter. It was harder for him to control the distance filter with his fat finger. We noted this suggestion for future use.

7.3.4 Discussion

The comparative study finds some important observations. Our technique and the 3DWedge+ have significantly lower selection time and error in the low-density environment than Halo 3D in the order task. This contradicted our early assumption (H1), where we predicted that all the techniques would perform similarly in the distance tasks for the lower density condition. The participants were confused about perceiving the distance as the arc in the halo 3D gets bigger with the aggregation of POIs. The participants required a longer duration to perceive the distance in order. This took a longer time for the users to perceive the order. In a study, Perea et al.[33] found Halo 3d to be misleading the users when overlapping is more. In the case of our scenario, though it was a low-density environment, aggregated halos were more complex to perceive than the 3DWedge+ and our technique. According to our hypothesis 2 (H2), we assumed our system would perform better than halo 3d and 3DWedge+ in medium and high-density conditions. In the medium-density condition, along with our technique, 3DWedge+ performed better than halo 3D, and our technique did not perform significantly better than 3DWedge+ in all the tasks. One reason for this (as mentioned by some participants and also the authors of 3DWedge+) is that the visibility of the tip of the ball arrows in the 3DWedge+ could give comparative cues for the POIs. However, in the closest task, there was no significant difference in the selection time and accuracy in medium-density conditions. We found that users took almost the same time to perceive the closest POI with all three techniques. The time they took to use the

distance filter in our technique was almost similar to when they found the shortest wedge in 3DWedge+ and located the closest POI from the aggregated halo in Halo 3D. Nonetheless, in the high-density conditions, the visual clutter made it difficult for the participants to understand the direction and distance in 3DWedge+ and halo 3D. In the direction task, the 3DWedge+ and our technique had similar error rates due to their ability to show the direction in 3D. However, due to increased visual complexity, participants took more time to locate the target in 3DWedge+. Therefore, our hypothesis (H2) remains partially correct. The user preference test also supports our system to be preferred in a high-density environment. We found that users used the distance filter to complete tasks in medium and high-density environments. Therefore, although our system was evaluated for 30 POIs, considering it a high-density environment, it applies to more than 30 POIs. Our system is as good as the state-of-the-art POI visualization techniques for direction and distance awareness for low and medium-density conditions. However, for high-density conditions, our system can be used for accurate and efficient POI visualization.

Our findings align with previous studies suggesting that 3D arrows are an effective and popular means of visualizing off-screen objects. We contribute novel insights into the organization of arrows when dealing with a greater number of POIs. Furthermore, our results support previous research indicating that 3D Halos are less efficient, more error-prone, and less preferred compared to 3D Arrows. However, it is important to note that 3D Halos may have advantages when the goal is to highlight specific objects or areas rather than focusing on direction or location. Additionally, the placement of metaphors is a critical consideration when representing a large number of POIs. Although, prior studies have demonstrated the effectiveness of central visualizations using 3D wedges in VR head-mounted displays (HMDs), but their effectiveness is reduced when applied to the limited screen size of a smartphone.

7.4 General Discussion

To ensure the validity of both experiments, several measures were taken to address potential threats. Although the sample size of 16 participants recruited solely from the university may limit the generalizability of the findings, efforts were made to include individuals with diverse backgrounds and experiences with AR applications. Additionally, participants were

given ample practice time to familiarize themselves with the Halo 3D, 3DWedge+ and our proposed technique, mitigating potential bias arising from their lack of prior knowledge. Consistency in the implementation details, including the use of the same device and the same lab environment, enhanced the reliability of the results. Furthermore, the inclusion of varying density conditions for the POIs allowed for a comprehensive evaluation of the techniques' effectiveness across different environments. The tasks designed for the experiment aimed to simulate real-world scenarios and assess the accuracy of off-screen object presentation. The tasks were designed carefully, considering previous literature and adopting established evaluation measures. Moreover, to prevent the influence of a learning effect on the study results, a random sequence of tasks was allocated for each cluttering threshold value when establishing the threshold. This approach helped to minimize any potential bias introduced by the order in which the tasks were presented to the participants.

Our proposed method for visualizing off-screen POIs with 3D arrows and distance information has implications for the development of efficient and accurate off-screen POI visualizers (specifically for smartphones) in high-density environments. The approach of dividing arrows into cells and using aggregated arrows to represent clusters of objects can be adopted by designers to avoid visual obstruction and improve user experience. The use of a distance filter and dashed lines to display distance information can also be implemented in the off-screen POI visualizers. The findings of this study have practical implications for the design and development of AR applications in various domains like tourism, retail and advertising. For instance, our system can be useful in the field of urban tourism. Existing AR-based navigation applications often struggle with visual clutter and limited off-screen POI visibility in crowded city environments. By introducing a novel visualization technique with strategically positioned 3D arrows and aggregation, this study provides a solution to enhance the user experience in locating and interpreting off-screen POIs. Tourists can now benefit from clear and unobtrusive visual cues on their smartphones, making them aware of the nearby attractions or points of interest while minimizing distraction and information overload. This practical implication can drive the development of more effective and user-friendly AR navigation apps, improving the overall tourism experience in urban settings.

The model for estimating the range of CT was developed based on data collected from

16 individuals and was tested with a maximum of 8 POIs per cell. However, it is possible that the upward slope curve in our model may reach saturation at some point, resulting in a flat curve at extremes. To ensure the effectiveness of our model, it is necessary to conduct a further evaluation with a larger number of POIs and participants. Additionally, we have assumed that users will accurately locate the POI in a single attempt without considering the time it may take for users to correct any mistakes or missed attempts. This consideration must be taken into account when refining the model. It is essential to establish the probability of miss while locating the target POI as the number of POIs per cell increases and incorporate it into the model accordingly.

Our research demonstrates that distributing the arrows over the edges of the smartphone screen can alleviate visual clutter, which was observed in prior works utilizing arrows in the centre of the screen. Our system is more useful when the POIs are distributed over different altitudes, as the POIs will be distributed over the cells. The use of the distance filter helped the users to declutter the screen and have more distance awareness. Few participants suggested increasing the size of the dashed line. Further work should be done to improve the visualization of distance information for higher-density environments. To solve this issue, we can allow the user to tap on a particular cell and enlarge it to view it more clearly. Doing so may take extra time for the user to inspect a POI. However, a study is required to evaluate the user experience when the zooming feature is incorporated into each cell. Though this system was developed specifically for smartphones, a part of the system also can be tested for HMDs. For instance, the use of a distance filter can be evaluated for HMD-based off-screen visualizations to reduce the visual clutter. Furthermore, In our study, the evaluation of the proposed system was conducted in a controlled laboratory environment, where external distractors were minimized. As a result, there is a need for future studies to examine the effectiveness of the system in a real-world setting, where the presence of diverse environmental factors may influence its performance.

7.5 Summary

In this chapter, we have proposed a method to visualize off-screen POIs with their distance and direction in a high-density environment (more than 15 POIs). We discussed how we could use 3D arrows to visualize the POIs. To avoid visual obstruction, we divided the

3D arrows pointing towards the POIs over the edges of the screen (left, right, top, bottom and corners) into cells. We discussed how to use an aggregated arrow inside a cell to represent a cluster of objects and proposed a model to establish a range of POIs that can be used as a cluster to be efficiently directed by an aggregated 3D arrow. Furthermore, we adopted the idea of using a dashed line from 3DWedge+[48] to show the distance information. We introduce a distance filter to allow the users to disambiguate the POIs based on distance. Finally, our system was comprehensively evaluated by comparing it to the state-of-the-art off-screen POI visualizers, specifically 3DWedge+ and Halo 3D. We found our system similarly capable of showing off-screen information in low-density (5 POIs) and medium-density conditions (15 POIs) with 3DWedge+. However, our system is most efficient and accurate for visualizing off-screen POIs in high-density (30 POIs) conditions.

Future directions for research should include conducting larger-scale evaluations with a greater number of participants and POIs to validate the effectiveness of the proposed model. Furthermore, there is a need to explore techniques aimed at enhancing distance visualization in higher-density environments. We also intend to investigate the integration of our technique into head-mounted displays. Moreover, assessing the performance of the system in real-world settings with diverse environmental factors will provide valuable insights for further improvement.

Chapter

8

Conclusions and Future Works

In conclusion, this thesis presents a comprehensive investigation of navigation challenges in XR, focusing on enhancing user experience in VR and AR through novel approaches to realism, CS mitigation, affective state recognition, and off-screen POI visualization. The key contributions are as follows:

Enhancing Realism in System-Automated VR Tours We developed a model to predict instantaneous walking speed, significantly enhancing realism in system-automated VR tours by aligning the virtual movement more closely with users' natural walking pace, as validated through user studies.

Comprehensive Analysis of CS Through a systematic review, we offered new insights into the causes, measurement techniques, and mitigation strategies for CS, providing a framework for developers and highlighting gaps for future research.

Optimizing VR Tour Duration We proposed a novel approach to minimize the duration of system-automated VR tours, thereby reducing exposure-induced discomfort and enhancing user experience without compromising the realistic walking speed.

Finding Visual Gain for Vibrotactile Feedback-based CS Mitigation Our work included the application of vibrotactile feedback as a CS mitigation strategy within a realistic walking speed model, identifying a new range of visual gain that balances realism with comfort.

Recognition of Users' Affective State We introduced a model to recognize users' emotional states using data from HMDs and handheld controllers, enabling personalized

VR experiences and proactive CS mitigation.

Breathing-Based Virtual Locomotion Technique Our novel "BreathWalk" technique leverages controlled diaphragmatic breathing for navigation in VR, reducing CS and enhancing immersion and user preference compared to traditional joystick-based steering.

Off-Screen POI Visualization in Handheld AR We proposed a clutter-free method for visualizing off-screen POIs in vertically dense regions, significantly improving efficiency and accuracy in navigating high-density POI environments.

8.1 Domain of applications of our contributions

The contributions presented in this thesis spans a wide range of XR applications. For instance, in the tourism sector, the development of a realistic walking speed model for system-automated VR tours can provide tourists with lifelike explorations of distant locations, enhancing their engagement and satisfaction without the need for physical travel. This has implications for educational applications as well, where such realistic experiences can enrich learning, making historical sites accessible and engaging for students. Moreover, using our duration reduction strategy developers can make the system-automated VR tour more comfortable. The comprehensive analysis of CS, including causes, measurement techniques, and mitigation strategies, offers invaluable insights for developers and users across various XR environments. By providing a holistic overview, this review serves as a foundational resource that can guide the development of more comfortable, accessible, and engaging XR applications. Developers can utilize this information to design experiences that significantly reduce the likelihood of CS, improving user retention and satisfaction across different XR platforms. For instance, in entertainment and gaming, understanding CS can lead to the creation of more inclusive games that accommodate sensitive users, thereby expanding the market reach. In education and training, insights from the CS review can inform the design of immersive learning environments that minimize discomfort, ensuring that learners can focus on the content without distraction or discomfort. The VLT named "BreathWalk" further extends the potential of VR by leveraging controlled breathing for navigation, which could enhance the therapeutic effects of VR meditation and relaxation exercises. It also broadens the accessibility of VR for people with physical disabilities to walk around a VR environment without any discomfort. For the gaming industry, the introduction of

affective state recognition offer pathways to create more personalized and engaging gaming experiences for commercial VR HMDs. Games designed with these principles in mind can adapt in real-time to the player's emotional state and physical comfort, leading to longer, more enjoyable gaming sessions. Moreover, the advancements in clutter-free POI visualisation in smartphone-based AR systems promise to transform urban exploration and navigation, making it easier for users to find their way in densely populated cities or unfamiliar environments. This technology could be integrated into navigation apps, enhancing the user experience with intuitive, clutter-free visual cues.

8.2 Limitations

The thesis presents several contributions to XR navigation, yet it encounters few limitations across its varied studies. A primary limitation is that the model for realistic walking speeds, which is based on changes in step frequency, does not fully capture the complex factors affecting walking behavior in the real world. This is especially true in situations involving turns and varied terrains. Furthermore, our work mainly focused on the speed and neglected other aspects of movement, such as head bobbing during walking. Despite attempts to optimize navigation paths to minimize CS and discomfort and to maximize the visitation of significant sites within a limited time, variations in the participants' time spent at specific checkpoints indicate that the model's predictions may not accurately represent actual user experiences. The model's fixed assumptions struggle to accommodate the variability in user behavior, such as individuals spending less time at sites than anticipated or favoring certain locations over others due to personal interests. The strategy to reduce visit durations also fails to account for the diversity in users' travel behaviors in the real world, such as choosing different modes of transportation based on terrain, and does not consider the significance of the order in which certain places are visited, as would be necessary for tours requiring a specific sequence of stops. In developing the emotional state classification model, we were limited to identifying only four categories of emotions, likely missing the detail of the emotional intensities that users may experience during VR interactions. The use of subjective questionnaires to establish a baseline for emotions introduces possible biases and inaccuracies in identifying emotions, highlighting the need for more objective and dependable measurement methods. Additionally, data collection was exclusively focused

on instances when controller buttons were pressed, omitting other behavioral insights that could be gleaned from user interactions with the controller and HMD when no buttons are pressed. In designing the breathing-based walking VLT, the study did not explicitly examine the impact of participants' VR experience and physical fitness levels, which could affect the results. Lastly, the off-screen POI visualization technique lacks a strategy to manage to show the direction of each POI as their density per cell grows and requires further investigation for scalability. Participant feedback indicates a need for improved visualization of distances in crowded settings, and the system's performance in real-world scenarios, outside of a controlled lab environment, has yet to be evaluated.

8.3 Future Works

To address the limitations and expand upon the current work, future efforts should concentrate on several key areas. Firstly, developing models for non-linear paths in system-automated VR tours is crucial to enhance realism. This involves creating algorithms that simulate realistic walking behaviors on varied terrains, including turns and elevation changes, thereby improving the immersive quality of VR tours. Secondly, expanding the application of the emotional state prediction model is essential. By building a more extensive dataset with a larger and more diverse participant base, the accuracy and applicability of models predicting users' emotional states in real-time can be significantly improved. This advancement could lead to VR environments that dynamically adjust to users' emotional responses, enhancing engagement and satisfaction.

Additionally, evaluating BreathWalk's dual-component system necessitates a deeper investigation into the individual and combined effects of torso movement and breathing on VR navigation and user experience. Systematic studies comparing Breathwalk with different locomotion setups can offer insights into optimizing user interaction in virtual environments. Lastly, advancing XR navigation with adaptive multi-sensory interfaces is paramount. Future research should aim to develop intelligent systems that adjust navigational aids' complexity and modality based on the user's context, preferences, and cognitive load. Such adaptive interfaces could greatly enhance wayfinding in XR, ensuring a seamless and intuitive experience that intelligently adapts to both virtual and augmented environments. Collectively, these areas of focus promise significant advancements in VR and XR, offering

more realistic, personalized, and efficient user experiences.

Bibliography

[1] V. Angelov, E. Petkov, G. Shipkovenski, and T. Kalushkov, "Modern virtual reality headsets," in *2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, 2020, pp. 1–5.

[2] F. Arena, M. Collotta, G. Pau, and F. Termine, "An overview of augmented reality," *Computers*, vol. 11, no. 2, 2022. [Online]. Available: https://www.mdpi.com/2073-431X/11/2/28

[3] S. Rokhsaritalemi, A. Sadeghi-Niaraki, and S.-M. Choi, "A review on mixed reality: Current trends, challenges and prospects," *Applied Sciences*, vol. 10, no. 2, 2020. [Online]. Available: https://www.mdpi.com/2076-3417/10/2/636

[4] H. Kolivand, A. El Rhalibi, M. S. Sunar, and T. Saba, "Revitage: Realistic virtual heritage taking shadows and sky illumination into account," *Journal of Cultural Heritage*, vol. 32, pp. 166–175, 2018.

[5] B. E. Riecke, J. J. LaViola, and E. Kruijff, "3d user interfaces for virtual reality and games: 3d selection, manipulation, and spatial navigation," in *ACM SIGGRAPH 2018 Courses*, ser. SIGGRAPH '18. New York, NY, USA: Association for Computing Machinery, 2018. [Online]. Available: https://doi.org/10.1145/3214834.3214869

[6] F. Argelaguet and C. Andujar, "A survey of 3d object selection techniques for virtual environments," *Computers & Graphics*, vol. 37, no. 3, pp. 121–136, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0097849312001793

[7] D. Mendes, F. M. Caputo, A. Giachetti, A. Ferreira, and J. Jorge, "A survey on 3d virtual object manipulation: From the desktop to immersive virtual environments," *Computer Graphics Forum*, vol. 38, no. 1, pp. 21–45, 2019. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/cgf.13390

[8] M.-L. Wu and V. Popescu, "Efficient vr and ar navigation through multiperspective occlusion management," *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 12, pp. 3069–3080, 2018.

[9] D. A. Bowman, E. Kruijff, J. J. LaViola, and I. Poupyrev, *3D User Interfaces: Theory and Practice*. USA: Addison Wesley Longman Publishing Co., Inc., 2004.

[10] J. J. Yang, C. Holz, E. Ofek, and A. D. Wilson, "Dreamwalker: Substituting real-world walking experiences with a virtual reality," in *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, ser. UIST '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 1093–1107. [Online]. Available: https://doi.org/10.1145/3332165.3347875

[11] J. Frommel, S. Sonntag, and M. Weber, "Effects of controller-based locomotion on player experience in a virtual reality exploration game," in *Proceedings of the 12th International Conference on the Foundations of Digital Games*, ser. FDG '17. New York, NY, USA: Association for Computing Machinery, 2017. [Online]. Available: https://doi.org/10.1145/3102071.3102082

[12] J. C. S. Cardoso, "Comparison of gesture, gamepad, and gaze-based locomotion for vr worlds," in *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, ser. VRST '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 319–320. [Online]. Available: https://doi.org/10.1145/2993369.2996327

[13] S. H. Pyo, H. S. Lee, B. M. Phu, S. J. Park, and J. W. Yoon, "Development of an fast-omnidirectional treadmill (f-odt) for immersive locomotion interface," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, 2018, pp. 760–766.

[14] J. Chung, F. Pagnini, and E. Langer, "Mindful navigation for pedestrians: Improving engagement with augmented reality," *Technology in Society*, vol. 45, pp. 29–33, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0160791X16000142

[15] Y. Zhang and T. Nakajima, "Exploring the design of a mixed-reality 3d minimap to enhance pedestrian satisfaction in urban exploratory navigation," *Future Internet*, vol. 14, no. 11, 2022. [Online]. Available: https://www.mdpi.com/1999-5903/14/11/325

[16] A. Krekhov, S. Cmentowski, K. Emmerich, M. Masuch, and J. Krüger, "Gullivr: A walking-oriented technique for navigation in virtual reality games based on virtual body resizing," in *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*, ser. CHI PLAY '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 243–256. [Online]. Available: https://doi.org/10.1145/3242671.3242704

[17] C. Ouerghemmi, M. Ertz, N. Bouslama, and U. Tandon, "The impact of virtual reality (vr) tour experience on tourists' intention to visit," *Information*, vol. 14, no. 10, 2023. [Online]. Available: https://www.mdpi.com/2078-2489/14/10/546

[18] B.-W. Lee, H.-Y. Shih, Y.-T. Chou, and Y.-S. Chen, "Educational virtual reality implementation on english for tourism purpose using knowledge-based engineering," in *2017 International Conference on Applied System Innovation (ICASI)*, 2017, pp. 792–795.

[19] A. Ganapathy, "Virtual reality and augmented reality driven real estate world to buy properties," *Asian Journal of Humanity, Art and Literature*, vol. 3, no. 2, p. 137–146, Dec. 2016. [Online]. Available: https://iproclaim.my/journals/index.php/ajhal/article/view/567

[20] K. Rahimi, C. Banigan, and E. D. Ragan, "Scene transitions and teleportation in virtual reality and the implications for spatial awareness and sickness," *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 6, pp. 2273–2287, 2020.

[21] C. Jicol, C. H. Wan, B. Doling, C. H. Illingworth, J. Yoon, C. Headey, C. Lutteroth, M. J. Proulx, K. Petrini, and E. O'Neill, "Effects of emotion and agency on presence in virtual reality," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, ser. CHI '21. New York, NY, USA: Association for Computing Machinery, 2021. [Online]. Available: https://doi.org/10.1145/3411764.3445588

[22] C. Kyrlitsias, M. Christofi, D. Michael-Grigoriou, D. Banakou, and A. Ioannou, "Corrigendum: A virtual tour of a hardly accessible archaeological site: The effect of immersive virtual reality in user experience, learning and attitude change," *Frontiers in Computer Science*, vol. 3, 2021. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fcomp.2021.697259

[23] G. B. Petersen, G. Petkakis, and G. Makransky, "A study of how immersion and interactivity drive vr learning," *Computers & Education*, vol. 179, p. 104429, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360131521003067

[24] C. R. Guerra-Tamez, "The impact of immersion through virtual reality in the learning experiences of art and design students: The mediating effect of the flow experience," *Education Sciences*, vol. 13, no. 2, 2023. [Online]. Available: https://www.mdpi.com/2227-7102/13/2/185

[25] W. Wang, S. Baker, and A. Irlitti, "Exploring the effects of user control on social engagement in virtual reality," in *Proceedings of the 32nd Australian Conference on Human-Computer Interaction*, ser. OzCHI '20. New York, NY, USA: Association for Computing Machinery, 2021, p. 253–262. [Online]. Available: https://doi.org/10.1145/3441000.3441076

[26] H.-Y. Tsai, M.-H. Tsai, T.-H. Chiang, and C.-C. Chang, "Comparative evaluation of different following mechanisms in vr guided tour: A preliminary study," *Applied Sciences*, vol. 12, no. 19, 2022. [Online]. Available: https://www.mdpi.com/2076-3417/12/19/9630

[27] L. T. Nielsen, M. B. Møller, S. D. Hartmeyer, T. C. M. Ljung, N. C. Nilsson, R. Nordahl, and S. Serafin, "Missing the point: an exploration of how to guide users' attention during cinematic virtual reality," in *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, ser. VRST '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 229–232. [Online]. Available: https://doi.org/10.1145/2993369.2993405

[28] E. Zidianakis, N. Partarakis, S. Ntoa, A. Dimopoulos, S. Kopidaki, A. Ntagianta, E. Ntafotis, A. Xhako, Z. Pervolarakis, E. Kontaki, I. Zidianaki, A. Michelakis, M. Foukarakis, and C. Stephanidis, "The invisible museum: A user-centric platform for creating virtual 3d exhibitions with vr support," *Electronics*, vol. 10, no. 3, 2021. [Online]. Available: https://www.mdpi.com/2079-9292/10/3/363

[29] M. Al Zayer, P. MacNeilage, and E. Folmer, "Virtual locomotion: A survey," *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 6, pp. 2315–2334, 2020.

[30] R. Li, "Human wayfinding and navigation in a large-scale environment: cognitive map development and wayfinding strategies," Ph.D. dissertation, 2007.

[31] X. Zhou, "Designing navigation tool for immersive analytics in ar," in *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 2023, pp. 975–976.

[32] T. Schinke, N. Henze, and S. Boll, "Visualization of off-screen objects in mobile augmented reality," in *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*, ser. MobileHCI '10. New York, NY, USA: Association for Computing Machinery, 2010, p. 313–316. [Online]. Available: https://doi.org/10.1145/1851600.1851655

[33] P. Perea, D. Morand, and L. Nigay, "Spotlight on off-screen points of interest in handheld augmented reality: Halo-based techniques," in *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces*, ser. ISS '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 43–54. [Online]. Available: https://doi.org/10.1145/3343055.3359719

[34] M. B. Carmo, A. P. Afonso, M. Melo, B. Rocha, and V. Botan, "Augmented reality with maps for off-screen poi awareness," in *2020 24th International Conference Information Visualisation (IV)*, 2020, pp. 454–459.

[35] S. Burigat, L. Chittaro, and S. Gabrielli, "Visualizing locations of off-screen objects on mobile devices: a comparative evaluation of three approaches," in *Proceedings of the 8th Conference on Human-Computer Interaction with Mobile Devices and Services*,

ser. MobileHCI '06. New York, NY, USA: Association for Computing Machinery, 2006, p. 239–246. [Online]. Available: https://doi.org/10.1145/1152215.1152266

[36] J. Wieland, R. C. H. Garcia, H. Reiterer, and T. Feuchtner, "Arrow, bezier curve, or halos? – comparing 3d out-of-view object visualization techniques for handheld augmented reality," in *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2022, pp. 797–806.

[37] T. Siu and V. Herskovic, "Sidebars: improving awareness of off-screen elements in mobile augmented reality," in *Proceedings of the 2013 Chilean Conference on Human - Computer Interaction*, ser. ChileCHI '13. New York, NY, USA: Association for Computing Machinery, 2013, p. 36–41. [Online]. Available: https://doi.org/10.1145/2535597.2535608

[38] D. Bowman, D. Koller, and L. Hodges, "Travel in immersive virtual environments: an evaluation of viewpoint motion control techniques," in *Proceedings of IEEE 1997 Annual International Symposium on Virtual Reality*, 1997, pp. 45–52.

[39] N. C. Nilsson, S. Serafin, and R. Nordahl, "Establishing the range of perceptually natural visual walking speeds for virtual walking-in-place locomotion," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 4, pp. 569–578, 2014.

[40] Y.-H. Peng, C. Yu, S.-H. Liu, C.-W. Wang, P. Taele, N.-H. Yu, and M. Y. Chen, "Walkingvibe: Reducing virtual reality sickness and improving realism while walking in vr using unobtrusive head-mounted vibrotactile feedback," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 1–12. [Online]. Available: https://doi.org/10.1145/3313831.3376847

[41] N. Seethapathi and M. Srinivasan, "The metabolic cost of changing walking speeds is significant, implies lower optimal speeds for shorter distances, and increases daily energy estimates," *Biology letters*, vol. 11, no. 9, p. 20150486, 2015.

[42] J. Clifton and S. Palmisano, "Effects of steering locomotion and teleporting on cybersickness and presence in hmd-based virtual reality," *Virtual Reality*, vol. 24, no. 3, pp. 453–468, 2020.

[43] M. E. B. Russell, B. Hoffman, S. Stromberg, and C. R. Carlson, "Use of controlled diaphragmatic breathing for the management of motion sickness in a virtual reality environment," *Applied psychophysiology and biofeedback*, vol. 39, pp. 269–277, 2014.

[44] S. Gustafson, P. Baudisch, C. Gutwin, and P. Irani, "Wedge: clutter-free visualization of off-screen locations," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '08. New York, NY, USA: Association for Computing Machinery, 2008, p. 787–796. [Online]. Available: https://doi.org/10.1145/1357054.1357179

[45] P. Baudisch and R. Rosenholtz, "Halo: a technique for visualizing off-screen objects," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '03. New York, NY, USA: Association for Computing Machinery, 2003, p. 481–488. [Online]. Available: https://doi.org/10.1145/642611.642695

[46] S. Burigat and L. Chittaro, "Visualizing references to off-screen content on mobile devices: A comparison of arrows, wedge, and overview + detail," *Interacting with Computers*, vol. 23, no. 2, pp. 156–166, 2011.

[47] U. Gruenefeld, D. Ennenga, A. E. Ali, W. Heuten, and S. Boll, "Eyesee360: designing a visualization technique for out-of-view objects in head-mounted augmented reality," in *Proceedings of the 5th Symposium on Spatial User Interaction*, ser. SUI '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 109–118. [Online]. Available: https://doi.org/10.1145/3131277.3132175

[48] D. Yu, H.-N. Liang, K. Fan, H. Zhang, C. Fleming, and K. Papangelis, "Design and evaluation of visualization techniques of off-screen and occluded targets in virtual reality environments," *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 9, pp. 2762–2774, 2020.

[49] A. Evangelista, V. M. Manghisi, F. Laera, M. Gattullo, A. E. Uva, and M. Fiorentino, "Compassbar: A technique for visualizing out-of-view objects in a mixed reality environment," in *Design Tools and Methods in Industrial Engineering II*, C. Rizzi, F. Campana, M. Bici, F. Gherardini, T. Ingrassia, and P. Cicconi, Eds. Cham: Springer International Publishing, 2022, pp. 141–148.

[50] H. Jo, S. Hwang, H. Park, and J. hee Ryu, "Aroundplot: Focus+context interface for off-screen objects in 3d environments," *Computers & Graphics*, vol. 35, no. 4, pp. 841–853, 2011, semantic 3D Media and Content. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0097849311001087

[51] M. Sra, X. Xu, and P. Maes, "Breathvr: Leveraging breathing as a directly controlled interface for virtual reality games," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ser. CHI '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 1–12. [Online]. Available: https://doi.org/10.1145/3173574.3173914

[52] E. Langbehn and F. Steinicke, *Redirected Walking in Virtual Reality.* Cham: Springer International Publishing, 2018, pp. 1–11. [Online]. Available: https://doi.org/10.1007/978-3-319-08234-9_253-1

[53] A. Prithul, I. B. Adhanom, and E. Folmer, "Teleportation in virtual reality; a mini-review," *Frontiers in Virtual Reality*, vol. 2, 2021. [Online]. Available: https://www.frontiersin.org/articles/10.3389/frvir.2021.730792

[54] C. Ware and S. Osborne, "Exploration and virtual camera control in virtual three dimensional environments," *SIGGRAPH Comput. Graph.*, vol. 24, no. 2, p. 175–183, feb 1990. [Online]. Available: https://doi.org/10.1145/91394.91442

[55] F. Argelaguet and M. Maignant, "Giant: stereoscopic-compliant multi-scale navigation in ves," in *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, ser. VRST '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 269–277. [Online]. Available: https://doi.org/10.1145/2993369.2993391

[56] C. Wingrave, Y. Haciahmetoglu, and D. Bowman, "Overcoming world in miniature limitations by a scaled and scrolling wim," in *3D User Interfaces (3DUI'06)*, 2006, pp. 11–16.

[57] N. C. Nilsson, S. Serafin, F. Steinicke, and R. Nordahl, "Natural walking in virtual reality: A review," *Comput. Entertain.*, vol. 16, no. 2, apr 2018. [Online]. Available: https://doi.org/10.1145/3180658

[58] K. Matsumoto, K. Aoyama, T. Narumi, and H. Kuzuoka, "Redirected walking using noisy galvanic vestibular stimulation," in *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2021, pp. 498–507.

[59] M. Sra, A. Jain, and P. Maes, "Adding proprioceptive feedback to virtual reality experiences using galvanic vestibular stimulation," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ser. CHI '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 1–14. [Online]. Available: https://doi.org/10.1145/3290605.3300905

[60] R. Boian, M. Bouzit, G. Burdea, J. Lewis, and J. Deutsch, "Dual stewart platform mobility simulator," in *9th International Conference on Rehabilitation Robotics, 2005. ICORR 2005.*, 2005, pp. 550–555.

[61] M. Hoppe, J. Karolus, F. Dietz, P. W. Woźniak, A. Schmidt, and T.-K. Machulla, "Vrsneaky: Increasing presence in vr through gait-aware auditory feedback," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ser. CHI '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 1–9. [Online]. Available: https://doi.org/10.1145/3290605.3300776

[62] H. Schmidt, "Hapticwalker-a novel haptic device for walking simulation," in *Proc. of'EuroHaptics*. Citeseer, 2004.

[63] K. Fujita, "Wearable locomotion interface using walk-in-place in real space (warp) for distributed multi-user walk-through application," in *Proc. IEEE Virtual reality workshop*, 2004, pp. 29–30.

[64] L. Yan, R. S. Allison, and S. K. Rushton, "New simple virtual walking method-walking on the spot," in *Proceedings of the IPT Symposium*, 2004, pp. 1–7.

[65] B. Fasel, C. Duc, F. Dadashi, F. Bardyn, M. Savary, P.-A. Farine, and K. Aminian, "A wrist sensor and algorithm to determine instantaneous walking cadence and speed in daily life walking," *Medical & biological engineering & computing*, vol. 55, pp. 1773–1785, 2017.

[66] S. Zihajehzadeh and E. J. Park, "Regression model-based walking speed estimation using wrist-worn inertial sensor," *PloS one*, vol. 11, no. 10, p. e0165211, 2016.

[67] J.-g. Park, A. Patel, D. Curtis, S. Teller, and J. Ledlie, "Online pose classification and walking speed estimation using handheld devices," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, ser. UbiComp '12. New York, NY, USA: Association for Computing Machinery, 2012, p. 113–122. [Online]. Available: https://doi.org/10.1145/2370216.2370235

[68] S. Zihajehzadeh and E. J. Park, "Experimental evaluation of regression model-based walking speed estimation using lower body-mounted imu," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2016, pp. 243–246.

[69] ——, "A gaussian process regression model for walking speed estimation using a head-worn imu," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2017, pp. 2345–2348.

[70] R. S. Kennedy, K. M. Stanney, and W. P. Dunlap, "Duration and Exposure to Virtual Environments: Sickness Curves During and Across Sessions," *Presence: Teleoperators and Virtual Environments*, vol. 9, no. 5, pp. 463–472, 10 2000. [Online]. Available: https://doi.org/10.1162/105474600566952

[71] H. T. K. Eunhee Chang and B. Yoo, "Virtual reality sickness: A review of causes and measurements," *International Journal of Human–Computer Interaction*, vol. 36, no. 17, pp. 1658–1682, 2020. [Online]. Available: https://doi.org/10.1080/10447318.2020.1778351

[72] S.-H. Liu, N.-H. Yu, L. Chan, Y.-H. Peng, W.-Z. Sun, and M. Y. Chen, "Phantomlegs: Reducing virtual reality sickness using head-worn haptic devices," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2019, pp. 817–826.

[73] V. T. Inman, H. J. H. J. Ralston, F. Todd, and J. C. Lieberman, *Human walking*. Williams & Wilkins, 1981. [Online]. Available: https://cir.nii.ac.jp/crid/1130282270897080960

[74] G. A. DEAN, "An analysis of the energy expenditure in level and grade walking," *Ergonomics*, vol. 8, no. 1, pp. 31–47, 1965. [Online]. Available: https://doi.org/10.1080/00140136508930772

[75] J. D. Wendt, M. C. Whitton, and F. P. Brooks, "Gud wip: Gait-understanding-driven walking-in-place," in *2010 IEEE Virtual Reality Conference (VR)*, 2010, pp. 51–58.

[76] S. Davis, K. Nesbitt, and E. Nalivaiko, "A systematic review of cybersickness," in *Proceedings of the 2014 conference on interactive entertainment*, 2014, pp. 1–9.

[77] L. Rebenitsch and C. Owen, "Review on cybersickness in applications and visual displays," *Virtual Reality*, vol. 20, pp. 101–125, 2016.

[78] E. Chang, H. T. Kim, and B. Yoo, "Virtual reality sickness: a review of causes and measurements," *International Journal of Human–Computer Interaction*, vol. 36, no. 17, pp. 1658–1682, 2020.

[79] P. Caserman, A. Garcia-Agundez, A. Gámez Zerban, and S. Göbel, "Cybersickness in current-generation virtual reality head-mounted displays: systematic review and outlook," *Virtual Reality*, vol. 25, no. 4, pp. 1153–1170, 2021.

[80] D. Saredakis, A. Szpak, B. Birckhead, H. A. Keage, A. Rizzo, and T. Loetscher, "Factors associated with virtual reality sickness in head-mounted displays: a systematic review and meta-analysis," *Frontiers in human neuroscience*, vol. 14, p. 96, 2020.

[81] C. Yildirim, "A review of deep learning approaches to eeg-based classification of cybersickness in virtual reality," in *2020 IEEE international conference on artificial intelligence and virtual reality (AIVR)*. IEEE, 2020, pp. 351–357.

[82] A. H. X. Yang, N. Kasabov, and Y. O. Cakmak, "Machine learning methods for the study of cybersickness: a systematic review," *Brain Informatics*, vol. 9, no. 1, p. 24, 2022.

[83] G. Li, M. McGill, S. Brewster, and F. Pollick, "A review of electrostimulation-based cybersickness mitigations," in *2020 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*. IEEE, 2020, pp. 151–157.

[84] C. MacArthur, A. Grinberg, D. Harley, and M. Hancock, "You're making me sick: A systematic review of how virtual reality research considers gender & cybersickness," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–15.

[85] J.-P. Stauffert, F. Niebling, and M. E. Latoschik, "Latency and cybersickness: Impact, causes, and measures. a review," *Frontiers in Virtual Reality*, vol. 1, p. 582204, 2020.

[86] N. Tian, P. Lopes, and R. Boulic, "A review of cybersickness in head-mounted displays: raising attention to individual susceptibility," *Virtual Reality*, vol. 26, no. 4, pp. 1409–1441, 2022.

[87] A. Liberati, D. G. Altman, J. Tetzlaff, C. Mulrow, P. C. Gøtzsche, J. P. Ioannidis, M. Clarke, P. J. Devereaux, J. Kleijnen, and D. Moher, "The prisma statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration," *Annals of internal medicine*, vol. 151, no. 4, pp. W–65, 2009.

[88] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, S. E. Brennan *et al.*, "The prisma 2020 statement: an updated guideline for reporting systematic reviews," *International journal of surgery*, vol. 88, p. 105906, 2021.

[89] L. Crameri, I. Hettiarachchi, and S. Hanoun, "A review of individual operational cognitive readiness: theory development and future directions," *Human factors*, vol. 63, no. 1, pp. 66–87, 2021.

[90] J. T. Reason and J. J. Brand, *Motion sickness.* Academic press, 1975.

[91] J. T. Reason, "Motion sickness adaptation: a neural mismatch model," *Journal of the Royal Society of Medicine*, vol. 71, no. 11, pp. 819–829, 1978.

[92] C. M. Oman, "Motion sickness: a synthesis and evaluation of the sensory conflict theory," *Canadian journal of physiology and pharmacology*, vol. 68, no. 2, pp. 294–303, 1990.

[93] W. Bles, J. E. Bos, B. De Graaf, E. Groen, and A. H. Wertheim, "Motion sickness: only one provocative conflict?" *Brain research bulletin*, vol. 47, no. 5, pp. 481–487, 1998.

[94] J. E. Bos, W. Bles, and E. L. Groen, "A theory on visually induced motion sickness," *Displays*, vol. 29, no. 2, pp. 47–57, 2008.

[95] G. E. Riccio and T. A. Stoffregen, "An ecological theory of motion sickness and postural instability," *Ecological psychology*, vol. 3, no. 3, pp. 195–240, 1991.

[96] M. Treisman, "Motion sickness: an evolutionary hypothesis," *Science*, vol. 197, no. 4302, pp. 493–495, 1977.

[97] F. Bonato, A. Bubka, S. Palmisano, D. Phillip, and G. Moreno, "Vection change exacerbates simulator sickness in virtual environments," *Presence: Teleoperators and Virtual Environments*, vol. 17, no. 3, pp. 283–292, 2008.

[98] C.-L. Liu and S.-T. Uang, "A study of sickness induced within a 3d virtual store and combated with fuzzy control in the elderly," in *2012 9th International Conference on Fuzzy Systems and Knowledge Discovery*. IEEE, 2012, pp. 334–338.

[99] A. J. Lubeck, J. E. Bos, and J. F. Stins, "Motion in images is essential to cause motion sickness symptoms, but not to increase postural sway," *Displays*, vol. 38, pp. 55–61, 2015.

[100] J. T. Ji, R. H. So, and R. T. Cheung, "Isolating the effects of vection and optokinetic nystagmus on optokinetic rotation-induced motion sickness," *Human factors*, vol. 51, no. 5, pp. 739–751, 2009.

[101] J. D. Prothero and D. E. Parker, "A unified approach to presence and motion sickness," *Virtual and adaptive environments: Applications, implications, and human performance issues*, p. 47, 2003.

[102] Y. Wang, J.-R. Chardonnet, F. Merienne, and J. Ovtcharova, "Using fuzzy logic to involve individual differences for predicting cybersickness during vr navigation," in *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2021, pp. 373–381.

[103] L. E. Garrido, M. Frías-Hiciano, M. Moreno-Jiménez, G. N. Cruz, Z. E. García-Batista, K. Guerra-Peña, and L. A. Medrano, "Focusing on cybersickness: pervasiveness, latent trajectories, susceptibility, and effects on the virtual reality experience," *Virtual Reality*, vol. 26, no. 4, pp. 1347–1371, 2022.

[104] L. Tychsen and P. Foeller, "Effects of immersive virtual reality headset viewing on young children: visuomotor function, postural stability, and motion sickness," *American journal of ophthalmology*, vol. 209, pp. 151–159, 2020.

[105] H. Oh and W. Son, "Cybersickness and its severity arising from virtual reality content: A comprehensive study," *Sensors*, vol. 22, no. 4, p. 1314, 2022.

[106] T. C. Peck, L. E. Sockol, and S. M. Hancock, "Mind the gap: The underrepresentation of female participants and authors in virtual reality research," *IEEE transactions on visualization and computer graphics*, vol. 26, no. 5, pp. 1945–1954, 2020.

[107] K. Stanney, C. Fidopiastis, and L. Foster, "Virtual reality is sexist: but it does not have to be," *Frontiers in Robotics and AI*, vol. 7, p. 4, 2020.

[108] A. Mazloumi Gavgani, F. R. Walker, D. M. Hodgson, and E. Nalivaiko, "A comparative study of cybersickness during exposure to virtual reality and "classic" motion sickness: are they different?" *Journal of Applied Physiology*, vol. 125, no. 6, pp. 1670–1680, 2018.

[109] J. Treleaven, J. Battershill, D. Cole, C. Fadelli, S. Freestone, K. Lang, and H. Sarig-Bahat, "Simulator sickness incidence and susceptibility during neck motion-controlled virtual reality tasks," *Virtual Reality*, vol. 19, pp. 267–275, 2015.

[110] J. M. Fulvio, M. Ji, and B. Rokers, "Variations in visual sensitivity predict motion sickness in virtual reality," *Entertainment Computing*, vol. 38, p. 100423, 2021.

[111] S. Katsigiannis, R. Willis, and N. Ramzan, "A qoe and simulator sickness evaluation of a smart-exercise-bike virtual reality system via user feedback and physiological signals," *IEEE Transactions on Consumer Electronics*, vol. 65, no. 1, pp. 119–127, 2018.

[112] S. Zhang, A. Kurogi, and Y. Ono, "Vr sickness in continuous exposure to live-action 180° video," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2019, pp. 1269–1270.

[113] D. Risi and S. Palmisano, "Effects of postural stability, active control, exposure duration and repeated exposures on hmd induced cybersickness," *Displays*, vol. 60, pp. 9–17, 2019.

[114] J. M. Mittelstädt, J. Wacker, and D. Stelling, "Emotional and cognitive modulation of cybersickness: The role of pain catastrophizing and body awareness," *Human Factors*, vol. 61, no. 2, pp. 322–336, 2019.

[115] E. Altena, Y. Daviaux, E. Sanz-Arigita, E. Bonhomme, É. de Sevin, J.-A. Micoulaud-Franchi, S. Bioulac, and P. Philip, "How sleep problems contribute to simulator sickness: Preliminary results from a realistic driving scenario," *Journal of sleep research*, vol. 28, no. 2, p. e12677, 2019.

[116] N.-G. Kim and B.-S. Kim, "The effect of retinal eccentricity on visually induced motion sickness and postural control," *Applied Sciences*, vol. 9, no. 9, p. 1919, 2019.

[117] I. M. Arafat, S. M. S. Ferdous, and J. Quarles, "The effects of cybersickness on persons with multiple sclerosis," in *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, 2016, pp. 51–59.

[118] W. Xu, H.-N. Liang, Q. He, X. Li, K. Yu, Y. Chen *et al.*, "Results and guidelines from a repeated-measures design experiment comparing standing and seated full-body gesture-based immersive virtual reality exergames: Within-subjects evaluation," *JMIR serious games*, vol. 8, no. 3, p. e17972, 2020.

[119] R. Pot-Kolder, W. Veling, J. Counotte, and M. Van Der Gaag, "Anxiety partially mediates cybersickness symptoms in immersive virtual reality environments," *Cyberpsychology, Behavior, and Social Networking*, vol. 21, no. 3, pp. 187–193, 2018.

[120] M. Kaufeld, J. Bourdeinik, L. M. Prinz, M. Mundt, and H. Hecht, "Emotions are associated with the genesis of visually induced motion sickness in virtual reality," *Experimental Brain Research*, vol. 240, no. 10, pp. 2757–2771, 2022.

[121] A. Y. Tychkov, N. S. Bofanova, A. K. Alimuradov, D. S. Chernyshov, and I. Miltykh, "Development of city of the future scene to assess the user experience in a virtual reality environment," in *2022 6th Scientific School Dynamics of Complex Networks and their Applications (DCNA)*. IEEE, 2022, pp. 282–285.

[122] A. Widyanti and H. N. Hafizhah, "The influence of personality, sound, and content difficulty on virtual reality sickness," *Virtual Reality*, vol. 26, no. 2, pp. 631–637, 2022.

[123] L. R. Rebenitsch, *Cybersickness prioritization and modeling.* Michigan State University, 2015.

[124] J. Wang, R. Shi, Z. Xiao, X. Qin, and H.-N. Liang, "Effect of render resolution on gameplay experience, performance, and simulator sickness in virtual reality games," *Proceedings of the ACM on Computer Graphics and Interactive Techniques*, vol. 5, no. 1, pp. 1–15, 2022.

[125] S. Palmisano, R. S. Allison, J. Teixeira, and J. Kim, "Differences in virtual and physical head orientation predict sickness during active head-mounted display-based virtual reality," *Virtual Reality*, pp. 1–21, 2022.

[126] I. Tošić, D. Hoffman, and N. Balram, "Effect of latency on simulator sickness in smartphone virtual reality," *Journal of the Society for Information Display*, vol. 29, no. 7, pp. 561–572, 2021.

[127] S. Kawamura and R. Kijima, "Effect of head mounted display latency on human stability during quiescent standing on one foot," in *2016 IEEE Virtual Reality (VR).* IEEE, 2016, pp. 199–200.

[128] K. Brunnström, E. Dima, T. Qureshi, M. Johanson, M. Andersson, and M. Sjöström, "Latency impact on quality of experience in a virtual reality simulator for remote control of machines," *Signal Processing: Image Communication*, vol. 89, p. 116005, 2020.

[129] J.-P. Stauffert, F. Niebling, and M. E. Latoschik, "Effects of latency jitter on simulator sickness in a search task," in *2018 IEEE conference on virtual reality and 3D user interfaces (VR).* IEEE, 2018, pp. 121–127.

[130] J. Mittelstaedt, J. Wacker, and D. Stelling, "Effects of display type and motion control on cybersickness in a virtual bike simulator," *Displays*, vol. 51, pp. 43–50, 2018.

[131] S.-M. Choy, E. Cheng, R. H. Wilkinson, I. Burnett, and M. W. Austin, "Quality of experience comparison of stereoscopic 3d videos in different projection devices:

flat screen, panoramic screen and virtual reality headset," *Ieee Access*, vol. 9, pp. 9584–9594, 2021.

[132] P. Sajjadi, J. Zhao, J. O. Wallgrün, P. La Femina, and A. Klippel, "Influence of hmd type and spatial ability on experiences and learning in place-based education," in *2021 7th International Conference of the Immersive Learning Research Network (iLRN)*. IEEE, 2021, pp. 1–8.

[133] D. Monteiro, H.-N. Liang, J. Wang, H. Chen, and N. Baghaei, "An in-depth exploration of the effect of 2d/3d views and controller types on first person shooter games in virtual reality," in *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 2020, pp. 713–724.

[134] B. Keshavarz, A. E. Philipp-Muller, W. Hemmerich, B. E. Riecke, and J. L. Campos, "The effect of visual motion stimulus characteristics on vection and visually induced motion sickness," *Displays*, vol. 58, pp. 71–81, 2019.

[135] K. K. Kwok, A. K. Ng, and H. Y. Lau, "Effect of navigation speed and vr devices on cybersickness," in *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 2018, pp. 91–92.

[136] T. Cui, Y. Yang, and Y. Guo, "Evaluation of height and speed effects on the comfort of vr motion picture display," in *2021 International Conference on Culture-oriented Science & Technology (ICCST)*. IEEE, 2021, pp. 426–430.

[137] H. TT Tran, N. P. Ngoc, C. T. Pham, Y. J. Jung, and T. C. Thang, "A subjective study on user perception aspects in virtual reality," *Applied sciences*, vol. 9, no. 16, p. 3384, 2019.

[138] F. Soyka, E. Kokkinara, M. Leyrer, H. Buelthoff, M. Slater, and B. Mohler, "Turbulent motions cannot shake vr," in *2015 IEEE Virtual Reality (VR)*. IEEE, 2015, pp. 33–40.

[139] Y. Tai, Y. Yang, and X. Wang, "Development of vr motion sickness test platform based on ue," in *2022 International Conference on Culture-Oriented Science and Technology (CoST)*. IEEE, 2022, pp. 169–173.

[140] R. Venkatakrishnan, R. Venkatakrishnan, R. G. Anaraky, M. Volonte, B. Knijnenburg, and S. V. Babu, "A structural equation modeling approach to understand the relationship between control, cybersickness and presence in virtual reality," in *2020 IEEE conference on virtual reality and 3D user interfaces (VR)*. IEEE, 2020, pp. 682–691.

[141] R. Venkatakrishnan, R. Venkatakrishnan, A. Bhargava, K. Lucaites, H. Solini, M. Volonte, A. Robb, S. V. Babu, W.-C. Lin, and Y.-X. Lin, "Comparative evaluation of the effects of motion control on cybersickness in immersive virtual environments," in *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2020, pp. 672–681.

[142] E. Langbehn, P. Lubos, and F. Steinicke, "Evaluation of locomotion techniques for room-scale vr: Joystick, teleportation, and redirected walking," in *Proceedings of the Virtual Reality International Conference-Laval Virtual*, 2018, pp. 1–9.

[143] D. Lim, S. Shirai, J. Orlosky, P. Ratsamee, Y. Uranishi, and H. Takemura, "Evaluation of user interfaces for three-dimensional locomotion in virtual reality," in *Proceedings of the 2022 ACM Symposium on Spatial User Interaction*, 2022, pp. 1–9.

[144] K. Rahimi, C. Banigan, and E. D. Ragan, "Scene transitions and teleportation in virtual reality and the implications for spatial awareness and sickness," *IEEE transactions on visualization and computer graphics*, vol. 26, no. 6, pp. 2273–2287, 2018.

[145] M. Al Zayer, P. MacNeilage, and E. Folmer, "Virtual locomotion: a survey," *IEEE transactions on visualization and computer graphics*, vol. 26, no. 6, pp. 2315–2334, 2018.

[146] J. Marengo, P. Lopes, and R. Boulic, "On the influence of the supine posture on simulation sickness in virtual reality," in *2019 IEEE Conference on Games (CoG)*. IEEE, 2019, pp. 1–8.

[147] A. Tiiro, "Effect of visual realism on cybersickness in virtual reality," *University of Oulu*, vol. 350, 2018.

[148] K. Lim, J. Lee, K. Won, N. Kala, and T. Lee, "A novel method for vr sickness reduction based on dynamic field of view processing," *Virtual Reality*, vol. 25, pp. 331–340, 2021.

[149] F. Wu and E. S. Rosenberg, "Asymmetric lateral field-of-view restriction to mitigate cybersickness during virtual turns," in *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2022, pp. 103–111.

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

[151] P. Bala, I. Oakley, V. Nisi, and N. J. Nunes, "Dynamic field of view restriction in 360 video: Aligning optical flow and visual slam to mitigate vims," in *Proceedings of the 2021 CHI conference on human factors in computing systems*, 2021, pp. 1–18.

[152] J. Teixeira and S. Palmisano, "Effects of dynamic field-of-view restriction on cybersickness and presence in hmd-based virtual reality," *Virtual Reality*, vol. 25, no. 2, pp. 433–445, 2021.

[153] M. Al Zayer, I. B. Adhanom, P. MacNeilage, and E. Folmer, "The effect of field-of-view restriction on sex bias in vr sickness and spatial navigation performance," in *Proceedings of the 2019 CHI conference on human factors in computing systems*, 2019, pp. 1–12.

[154] F. Buttussi and L. Chittaro, "Effects of different types of virtual reality display on presence and learning in a safety training scenario," *IEEE transactions on visualization and computer graphics*, vol. 24, no. 2, pp. 1063–1076, 2017.

[155] J. Guna, G. Geršak, I. Humar, M. Krebl, M. Orel, H. Lu, and M. Pogačnik, "Virtual reality sickness and challenges behind different technology and content settings," *Mobile Networks and Applications*, vol. 25, pp. 1436–1445, 2020.

[156] U. A. Chattha, U. I. Janjua, F. Anwar, T. M. Madni, M. F. Cheema, and S. I. Janjua, "Motion sickness in virtual reality: An empirical evaluation," *IEEE Access*, vol. 8, pp. 130 486–130 499, 2020.

[157] S. Ang and J. Quarles, "You're in for a bumpy ride! uneven terrain increases cybersickness while navigating with head mounted displays," in *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2022, pp. 428–435.

[158] J. Guna, G. Geršak, I. Humar, J. Song, J. Drnovšek, and M. Pogačnik, "Influence of video content type on users' virtual reality sickness perception and physiological response," *Future Generation Computer Systems*, vol. 91, pp. 263–276, 2019.

[159] D. Monteiro, H.-N. Liang, W. Xu, M. Brucker, V. Nanjappan, and Y. Yue, "Evaluating enjoyment, presence, and emulator sickness in vr games based on first-and third-person viewing perspectives," *Computer Animation and Virtual Worlds*, vol. 29, no. 3-4, p. e1830, 2018.

[160] W. Al-Ashwal, H. Asadi, S. Mohamed, S. Alsanwy, L. Kooijman, D. Nahavandi, A. A. Alqumsan, and S. Nahavandi, "Cybersickness measurement and evaluation during flying a helicopter in different weather conditions in virtual reality," in *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2021, pp. 2152–2157.

[161] C. N. Aldaba and Z. Moussavi, "Effects of virtual reality technology locomotive multisensory motion stimuli on a user simulator sickness and controller intuitiveness during a navigation task," *Medical & biological engineering & computing*, vol. 58, pp. 143–154, 2020.

[162] S. Cao, K. Nandakumar, R. Babu, and B. Thompson, "Game play in virtual reality driving simulation involving head-mounted display and comparison to desktop display," *Virtual Reality*, vol. 24, pp. 503–513, 2020.

[163] T. Porcino, E. O. Rodrigues, F. Bernardini, D. Trevisan, and E. Clua, "Identifying cybersickness causes in virtual reality games using symbolic machine learning algorithms," *Entertainment Computing*, vol. 41, p. 100473, 2022.

[164] T. Kuosmanen, "The effect of visual detail on cybersickness: Predicting symptom severity using spatial velocity," 2019.

[165] M. Pouke, A. Tiiro, S. M. LaValle, and T. Ojala, "Effects of visual realism and moving detail on cybersickness," in *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2018, pp. 665–666.

[166] K. M. Stanney, R. S. Kennedy, and J. M. Drexler, "Cybersickness is not simulator sickness," in *Proceedings of the Human Factors and Ergonomics Society annual meeting*, vol. 41, no. 2. SAGE Publications Sage CA: Los Angeles, CA, 1997, pp. 1138–1142.

[167] S. Bouchard, M. Berthiaume, G. Robillard, H. Forget, C. Daudelin-Peltier, P. Renaud, C. Blais, and D. Fiset, "Arguing in favor of revising the simulator sickness questionnaire factor structure when assessing side effects induced by immersions in virtual reality," *Frontiers in Psychiatry*, vol. 12, p. 739742, 2021.

[168] H. K. Kim, J. Park, Y. Choi, and M. Choe, "Virtual reality sickness questionnaire (vrsq): Motion sickness measurement index in a virtual reality environment," *Applied ergonomics*, vol. 69, pp. 66–73, 2018.

[169] W. B. Stone III, "Psychometric evaluation of the simulator sickness questionnaire as a measure of cybersickness," Ph.D. dissertation, Iowa State University, 2017.

[170] J. E. Bos, S. C. de Vries, M. L. van Emmerik, and E. L. Groen, "The effect of internal and external fields of view on visually induced motion sickness," *Applied ergonomics*, vol. 41, no. 4, pp. 516–521, 2010.

[171] M. L. van Emmerik, S. C. de Vries, and J. E. Bos, "Internal and external fields of view affect cybersickness," *Displays*, vol. 32, no. 4, pp. 169–174, 2011.

[172] B. Keshavarz and H. Hecht, "Axis rotation and visually induced motion sickness: the role of combined roll, pitch, and yaw motion," *Aviation, space, and environmental medicine*, vol. 82, no. 11, pp. 1023–1029, 2011.

[173] ——, "Validating an efficient method to quantify motion sickness," *Human factors*, vol. 53, no. 4, pp. 415–426, 2011.

[174] ——, "Pleasant music as a countermeasure against visually induced motion sickness," *Applied ergonomics*, vol. 45, no. 3, pp. 521–527, 2014.

[175] R. S. Kennedy, J. E. Fowlkes, K. S. Berbaum, and M. G. Lilienthal, "Use of a motion sickness history questionnaire for prediction of simulator sickness." *Aviation, Space, and Environmental Medicine*, vol. 63, no. 7, pp. 588–593, 1992.

[176] B. D. Lawson, "Motion sickness scaling." 2014.

[177] J. F. Golding, "Motion sickness susceptibility questionnaire revised and its relationship to other forms of sickness," *Brain research bulletin*, vol. 47, no. 5, pp. 507–516, 1998.

[178] ——, "Predicting individual differences in motion sickness susceptibility by questionnaire," *Personality and Individual differences*, vol. 41, no. 2, pp. 237–248, 2006.

[179] B. Keshavarz, B. Murovec, N. Mohanathas, and J. F. Golding, "The visually induced motion sickness susceptibility questionnaire (vimssq): estimating individual susceptibility to motion sickness-like symptoms when using visual devices," *Human factors*, vol. 65, no. 1, pp. 107–124, 2023.

[180] R. H. So and W. Lo, "Cybersickness: an experimental study to isolate the effects of rotational scene oscillations," in *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)*. IEEE, 1999, pp. 237–241.

[181] R. H. So, W. Lo, and A. T. Ho, "Effects of navigation speed on motion sickness caused by an immersive virtual environment," *Human factors*, vol. 43, no. 3, pp. 452–461, 2001.

[182] R. H. So, A. Ho, and W. Lo, "A metric to quantify virtual scene movement for the study of cybersickness: Definition, implementation, and verification," *Presence*, vol. 10, no. 2, pp. 193–215, 2001.

[183] P. J. Gianaros, E. R. Muth, J. T. Mordkoff, M. E. Levine, and R. M. Stern, "A questionnaire for the assessment of the multiple dimensions of motion sickness," *Aviation, space, and environmental medicine*, vol. 72, no. 2, p. 115, 2001.

[184] M. S. Dennison, A. Z. Wisti, and M. D'Zmura, "Use of physiological signals to predict cybersickness," *Displays*, vol. 44, pp. 42–52, 2016.

[185] R. Islam, Y. Lee, M. Jaloli, I. Muhammad, D. Zhu, P. Rad, Y. Huang, and J. Quarles, "Automatic detection and prediction of cybersickness severity using deep neural networks from user's physiological signals," in *2020 IEEE international symposium on mixed and augmented reality (ISMAR)*. IEEE, 2020, pp. 400–411.

[186] Y.-T. Lin, Y.-Y. Chien, H.-H. Wang, F.-C. Lin, and Y.-P. Huang, "65-3: The quantization of cybersickness level using eeg and ecg for virtual reality head-mounted display," in *SID Symposium Digest of Technical Papers*, vol. 49, no. 1. Wiley Online Library, 2018, pp. 862–865.

[187] M. Recenti, C. Ricciardi, R. Aubonnet, I. Picone, D. Jacob, H. Á. Svansson, S. Agnarsdóttir, G. H. Karlsson, V. Baeringsdóttir, H. Petersen *et al.*, "Toward predicting motion sickness using virtual reality and a moving platform assessing brain, muscles, and heart signals," *Frontiers in Bioengineering and Biotechnology*, vol. 9, p. 635661, 2021.

[188] M. Aksoy, C. E. Ufodiama, A. D. Bateson, S. Martin, and A. U. Asghar, "A comparative experimental study of visual brain event-related potentials to a working memory task: virtual reality head-mounted display versus a desktop computer screen," *Experimental Brain Research*, vol. 239, pp. 3007–3022, 2021.

[189] D. Jeong, S. Yoo, and J. Yun, "Cybersickness analysis with eeg using deep learning algorithms," in *2019 IEEE conference on virtual reality and 3D user interfaces (VR)*. IEEE, 2019, pp. 827–835.

[190] T. Magaki and M. Vallance, "Developing an accessible evaluation method of vr cybersickness," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2019, pp. 1072–1073.

[191] C.-Y. Liao, S.-K. Tai, R.-C. Chen, and H. Hendry, "Using eeg and deep learning to predict motion sickness under wearing a virtual reality device," *Ieee Access*, vol. 8, pp. 126 784–126 796, 2020.

[192] E. S. Pane, A. Z. Khoirunnisaa, A. D. Wibawa, and M. H. Purnomo, "Identifying severity level of cybersickness from eeg signals using cn2 rule induction algorithm," in *2018 international conference on intelligent informatics and biomedical sciences (ICIIBMS)*, vol. 3. IEEE, 2018, pp. 170–176.

[193] M. Dennison Jr, M. D'Zmura, A. Harrison, M. Lee, and A. Raglin, "Improving motion sickness severity classification through multi-modal data fusion," in *Artificial*

intelligence and machine learning for multi-domain operations applications, vol. 11006. SPIE, 2019, pp. 277–286.

[194] X. Li, C. Zhu, C. Xu, J. Zhu, Y. Li, and S. Wu, "Vr motion sickness recognition by using eeg rhythm energy ratio based on wavelet packet transform," *Computer methods and programs in biomedicine*, vol. 188, p. 105266, 2020.

[195] M. A. Mawalid, A. Z. Khoirunnisa, M. H. Purnomo, and A. D. Wibawa, "Classification of eeg signal for detecting cybersickness through time domain feature extraction using naïve bayes," in *2018 international conference on computer engineering, network and intelligent multimedia (CENIM)*. IEEE, 2018, pp. 29–34.

[196] A. Z. Khoirunnisaa, E. S. Pane, A. D. Wibawa, and M. H. Purnomo, "Channel selection of eeg-based cybersickness recognition during playing video game using correlation feature selection (cfs)," in *2018 2nd international conference on biomedical engineering (IBIOMED)*. IEEE, 2018, pp. 48–53.

[197] C.-T. Lin, S.-F. Tsai, and L.-W. Ko, "Eeg-based learning system for online motion sickness level estimation in a dynamic vehicle environment," *IEEE transactions on neural networks and learning systems*, vol. 24, no. 10, pp. 1689–1700, 2013.

[198] A. Garcia-Agundez, C. Reuter, H. Becker, R. Konrad, P. Caserman, A. Miede, and S. Göbel, "Development of a classifier to determine factors causing cybersickness in virtual reality environments," *Games for health journal*, vol. 8, no. 6, pp. 439–444, 2019.

[199] R. Islam, Y. Lee, M. Jaloli, I. Muhammad, D. Zhu, and J. Quarles, "Automatic detection of cybersickness from physiological signal in a virtual roller coaster simulation," in *2020 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW)*. IEEE, 2020, pp. 648–649.

[200] N. Martin, N. Mathieu, N. Pallamin, M. Ragot, and J.-M. Diverrez, "Virtual reality sickness detection: an approach based on physiological signals and machine learning," in *2020 IEEE international symposium on mixed and augmented reality (ISMAR)*. IEEE, 2020, pp. 387–399.

[201] S. Palmisano, B. Arcioni, and P. J. Stapley, "Predicting vection and visually induced motion sickness based on spontaneous postural activity," *Experimental brain research*, vol. 236, pp. 315–329, 2018.

[202] T. A. Stoffregen, C.-H. Chang, F.-C. Chen, and W.-J. Zeng, "Effects of decades of physical driving on body movement and motion sickness during virtual driving," *PLoS One*, vol. 12, no. 11, p. e0187120, 2017.

[203] K. M. Stanney, R. S. Kennedy, J. M. Drexler, and D. L. Harm, "Motion sickness and proprioceptive aftereffects following virtual environment exposure," *Applied ergonomics*, vol. 30, no. 1, pp. 27–38, 1999.

[204] M. H. Draper, E. S. Viirre, T. A. Furness, and V. J. Gawron, "Effects of image scale and system time delay on simulator sickness within head-coupled virtual environments," *Human factors*, vol. 43, no. 1, pp. 129–146, 2001.

[205] J.-R. Chardonnet, M. A. Mirzaei, and F. Mérienne, "Visually induced motion sickness estimation and prediction in virtual reality using frequency components analysis of postural sway signal," in *International conference on artificial reality and telexistence eurographics symposium on virtual environments*, 2015, pp. 9–16.

[206] X. Dong and T. A. Stoffregen, "Postural activity and motion sickness among drivers and passengers in a console video game," in *Proceedings of the human factors and ergonomics society annual meeting*, vol. 54, no. 18. SAGE Publications Sage CA: Los Angeles, CA, 2010, pp. 1340–1344.

[207] S. J. Villard, M. B. Flanagan, G. M. Albanese, and T. A. Stoffregen, "Postural instability and motion sickness in a virtual moving room," *Human factors*, vol. 50, no. 2, pp. 332–345, 2008.

[208] H. B.-L. Duh, H. Abi-Rached, D. E. Parker, and T. A. Furness, "Effects on balance disturbance of manipulating depth of an independent visual background in a stereographic display," in *Proceedings of the human factors and ergonomics society annual meeting*, vol. 45, no. 27. SAGE Publications Sage CA: Los Angeles, CA, 2001, pp. 1882–1885.

[209] M. Dennison and M. D'Zmura, "Effects of unexpected visual motion on postural sway and motion sickness," *Applied ergonomics*, vol. 71, pp. 9–16, 2018.

[210] L. Rebenitsch and B. Quinby, "Cybersickness and postural sway using hmd orientation," in *Virtual, Augmented and Mixed Reality. Multimodal Interaction: 11th International Conference, VAMR 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26–31, 2019, Proceedings, Part I 21.* Springer, 2019, pp. 500–509.

[211] C. A. Tirado Cortes, H.-T. Chen, and C.-T. Lin, "Analysis of vr sickness and gait parameters during non-isometric virtual walking with large translational gain," in *The 17th international conference on virtual-reality continuum and its applications in industry*, 2019, pp. 1–10.

[212] S. Litleskare, "The relationship between postural stability and cybersickness: it's complicated–an experimental trial assessing practical implications of cybersickness etiology," *Physiology & Behavior*, vol. 236, p. 113422, 2021.

[213] Y. Y. Kim, H. J. Kim, E. N. Kim, H. D. Ko, and H. T. Kim, "Characteristic changes in the physiological components of cybersickness," *Psychophysiology*, vol. 42, no. 5, pp. 616–625, 2005.

[214] P. A. Howarth, "Oculomotor changes within virtual environments," *Applied Ergonomics*, vol. 30, no. 1, pp. 59–67, 1999.

[215] C. Diels, K. Ukai, and P. A. Howarth, "Visually induced motion sickness with radial displays: effects of gaze angle and fixation," *Aviation, space, and environmental medicine*, vol. 78, no. 7, pp. 659–665, 2007.

[216] S. Yang and J. E. Sheedy, "Effects of vergence and accommodative responses on viewer's comfort in viewing 3d stimuli," in *Stereoscopic Displays and Applications XXII*, vol. 7863. SPIE, 2011, pp. 231–243.

[217] D. Jeong, M. Jeong, U. Yang, and K. Han, "Eyes on me: Investigating the role and influence of eye-tracking data on user modeling in virtual reality," *Plos one*, vol. 17, no. 12, p. e0278970, 2022.

[218] E. Chang, H. T. Kim, and B. Yoo, "Predicting cybersickness based on user's gaze behaviors in hmd-based virtual reality," *Journal of Computational Design and Engineering*, vol. 8, no. 2, pp. 728–739, 2021.

[219] J. Wang, H.-N. Liang, D. Monteiro, W. Xu, and J. Xiao, "Real-time prediction of simulator sickness in virtual reality games," *IEEE Transactions on Games*, 2022.

[220] P. Lopes, N. Tian, and R. Boulic, "Eye thought you were sick! exploring eye behaviors for cybersickness detection in vr," in *Proceedings of the 13th ACM SIGGRAPH Conference on Motion, Interaction and Games*, 2020, pp. 1–10.

[221] Y. Wang, J.-R. Chardonnet, and F. Merienne, "Vr sickness prediction for navigation in immersive virtual environments using a deep long short term memory model," in *2019 IEEE conference on virtual reality and 3D user interfaces (VR)*. IEEE, 2019, pp. 1874–1881.

[222] Y. Li, A. Liu, and L. Ding, "Machine learning assessment of visually induced motion sickness levels based on multiple biosignals," *Biomedical signal processing and control*, vol. 49, pp. 202–211, 2019.

[223] N. Padmanaban, T. Ruban, V. Sitzmann, A. M. Norcia, and G. Wetzstein, "Towards a machine-learning approach for sickness prediction in 360 stereoscopic videos," *IEEE transactions on visualization and computer graphics*, vol. 24, no. 4, pp. 1594–1603, 2018.

[224] A. Dosovitskiy, P. Fischer, E. Ilg, P. Hausser, C. Hazirbas, V. Golkov, P. Van Der Smagt, D. Cremers, and T. Brox, "Flownet: Learning optical flow with convolutional networks," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 2758–2766.

[225] H. G. Kim, W. J. Baddar, H.-t. Lim, H. Jeong, and Y. M. Ro, "Measurement of exceptional motion in vr video contents for vr sickness assessment using deep convolutional autoencoder," in *Proceedings of the 23rd ACM symposium on virtual reality software and technology*, 2017, pp. 1–7.

[226] T. Hirzle, M. Cordts, E. Rukzio, J. Gugenheimer, and A. Bulling, "A critical assessment of the use of ssq as a measure of general discomfort in vr head-mounted displays," in

Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, 2021, pp. 1–14.

[227] M. Yu, R. Zhou, H. Wang, and W. Zhao, "An evaluation for vr glasses system user experience: The influence factors of interactive operation and motion sickness," *Applied ergonomics*, vol. 74, pp. 206–213, 2019.

[228] A. Singla, S. Göring, D. Keller, R. R. R. Rao, S. Fremerey, and A. Raake, "Assessment of the simulator sickness questionnaire for omnidirectional videos," in *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2021, pp. 198–206.

[229] V. Sevinc and M. I. Berkman, "Psychometric evaluation of simulator sickness questionnaire and its variants as a measure of cybersickness in consumer virtual environments," *Applied ergonomics*, vol. 82, p. 102958, 2020.

[230] R. Li *et al.*, "A pilot study on electroencephalogram-based evaluation of visually induced motion sickness, j," *Imaging Technol*, vol. 64, no. 2, 2020.

[231] S. Kye, C.-I. Moon, J. Lee, and O. Lee, "A study on the appropriateness of visual-related eeg electrodes for cybersickness measurement," in *2022 IEEE Sensors*. IEEE, 2022, pp. 1–4.

[232] J. Heo and G. Yoon, "Eeg studies on physical discomforts induced by virtual reality gaming," *Journal of Electrical Engineering & Technology*, vol. 15, pp. 1323–1329, 2020.

[233] A. D. Wibawa, S. Mardi, S. Nugroho, A. Z. Khoirunnisaa *et al.*, "Eeg visualization for cybersickness detection during playing 3d video games," in *2019 International seminar on intelligent technology and its applications (ISITIA)*. IEEE, 2019, pp. 325–330.

[234] S. S. Yeo, J. W. Kwon, and S. Y. Park, "Eeg-based analysis of various sensory stimulation effects to reduce visually induced motion sickness in virtual reality," *Scientific Reports*, vol. 12, no. 1, p. 18043, 2022.

[235] T. Gruden, N. B. Popović, K. Stojmenova, G. Jakus, N. Miljković, S. Tomažič, and J. Sodnik, "Electrogastrography in autonomous vehicles—an objective method for assessment of motion sickness in simulated driving environments," *Sensors*, vol. 21, no. 2, p. 550, 2021.

[236] K.-M. Jang, M. Kwon, S. G. Nam, D. Kim, and H. K. Lim, "Estimating objective (eeg) and subjective (ssq) cybersickness in people with susceptibility to motion sickness," *Applied ergonomics*, vol. 102, p. 103731, 2022.

[237] H. K. Lim, K. Ji, Y. S. Woo, D.-u. Han, D.-H. Lee, S. G. Nam, and K.-M. Jang, "Test-retest reliability of the virtual reality sickness evaluation using electroencephalography (eeg)," *Neuroscience Letters*, vol. 743, p. 135589, 2021.

[238] U. Celikcan, "Detection and mitigation of cybersickness via eeg-based visual comfort improvement," in *2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*. IEEE, 2019, pp. 1–4.

[239] B. Keshavarz and S. Berti, "Integration of sensory information precedes the sensation of vection: a combined behavioral and event-related brain potential (erp) study," *Behavioural brain research*, vol. 259, pp. 131–136, 2014.

[240] R. K. Kundu, R. Islam, P. Calyam, and K. A. Hoque, "Truvr: Trustworthy cybersickness detection using explainable machine learning," in *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 2022, pp. 777–786.

[241] J. Kim, W. Kim, H. Oh, S. Lee, and S. Lee, "A deep cybersickness predictor based on brain signal analysis for virtual reality contents," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 10 580–10 589.

[242] T. M. Lee, J.-C. Yoon, and I.-K. Lee, "Motion sickness prediction in stereoscopic videos using 3d convolutional neural networks," *IEEE transactions on visualization and computer graphics*, vol. 25, no. 5, pp. 1919–1927, 2019.

[243] R. K. Kundu, A. Rahman, and S. Paul, "A study on sensor system latency in vr motion sickness," *Journal of Sensor and Actuator Networks*, vol. 10, no. 3, p. 53, 2021.

[244] Y. Lee and M. Alamaniotis, "Unsupervised eeg cybersickness prediction with deep embedded self organizing map," in *2020 IEEE 20th international conference on bioinformatics and bioengineering (BIBE)*. IEEE, 2020, pp. 538–542.

[245] H. G. Kim, H.-T. Lim, S. Lee, and Y. M. Ro, "Vrsa net: Vr sickness assessment considering exceptional motion for 360 vr video," *IEEE transactions on image processing*, vol. 28, no. 4, pp. 1646–1660, 2018.

[246] S. Lee, S. Kim, H. G. Kim, and Y. M. Ro, "Assessing individual vr sickness through deep feature fusion of vr video and physiological response," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 5, pp. 2895–2907, 2021.

[247] K. Kim, S. Lee, H. G. Kim, M. Park, and Y. M. Ro, "Deep objective assessment model based on spatio-temporal perception of 360-degree video for vr sickness prediction," in *2019 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2019, pp. 3192–3196.

[248] S. Kim, S. Lee, and Y. M. Ro, "Estimating vr sickness caused by camera shake in vr videography," in *2020 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2020, pp. 3433–3437.

[249] M. Du, H. Cui, Y. Wang, and H. B.-L. Duh, "Learning from deep stereoscopic attention for simulator sickness prediction," *IEEE Transactions on Visualization and Computer Graphics*, vol. 29, no. 2, pp. 1415–1423, 2021.

[250] B. Arcioni, S. Palmisano, D. Apthorp, and J. Kim, "Postural stability predicts the likelihood of cybersickness in active hmd-based virtual reality," *Displays*, vol. 58, pp. 3–11, 2019.

[251] R. Islam, K. Desai, and J. Quarles, "Towards forecasting the onset of cybersickness by fusing physiological, head-tracking and eye-tracking with multimodal deep fusion network," in *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 2022, pp. 121–130.

[252] S. P. Smith, "Comparing virtual environments for cybersickness using a cumulative optical flow entropy metric," *IEEE Access*, vol. 9, pp. 68 898–68 904, 2021.

[253] R. Islam, K. Desai, and J. Quarles, "Cybersickness prediction from integrated hmd's sensors: A multimodal deep fusion approach using eye-tracking and head-tracking data," in *2021 IEEE international symposium on mixed and augmented reality (ISMAR)*. IEEE, 2021, pp. 31–40.

[254] H.-L. Duh, J. Lin, R. V. Kenyon, D. E. Parker, and T. A. Furness, "Effects of field of view on balance in an immersive environment," in *Proceedings IEEE Virtual Reality 2001.* IEEE, 2001, pp. 235–240.

[255] G. Li, O. Onuoha, M. McGill, S. Brewster, C. P. Chen, and F. Pollick, "Comparing autonomic physiological and electroencephalography features for vr sickness detection using predictive models," in *2021 IEEE Symposium Series on Computational Intelligence (SSCI).* IEEE, 2021, pp. 01–08.

[256] N. Kobayashi, R. Iinuma, Y. Suzuki, T. Shimada, and M. Ishikawa, "Using bio-signals to evaluate multi discomfort in image viewing-balancing visually induced motion sickness and field of view," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).* IEEE, 2015, pp. 6198–6201.

[257] M. Venturino and M. J. Wells, "Head movements as a function of field-of-view size on a helmet-mounted display," in *Proceedings of the Human Factors Society Annual Meeting*, vol. 34, no. 19. SAGE Publications Sage CA: Los Angeles, CA, 1990, pp. 1572–1576.

[258] A. F. Seay, D. M. Krum, L. Hodges, and W. Ribarsky, "Simulator sickness and presence in a high fov virtual environment," in *Proceedings IEEE Virtual Reality 2001.* IEEE, 2001, pp. 299–300.

[259] M. Bolas, J. A. Jones, I. McDowall, and E. Suma, "Dynamic field of view throttling as a means of improving user experience in head mounted virtual environments," May 9 2017, uS Patent 9,645,395.

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

[261] F. Wu and E. Suma Rosenberg, "Adaptive field-of-view restriction: Limiting optical flow to mitigate cybersickness in virtual reality," in *Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology*, 2022, pp. 1–11.

[262] F. Wu, G. S. Bailey, T. Stoffregen, and E. Suma Rosenberg, "Don't block the ground: Reducing discomfort in virtual reality with an asymmetric field-of-view restrictor," in *Proceedings of the 2021 ACM Symposium on Spatial User Interaction*, 2021, pp. 1–10.

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

[264] G.-Y. Nie, H. B.-L. Duh, Y. Liu, and Y. Wang, "Analysis on mitigation of visually induced motion sickness by applying dynamical blurring on a user's retina," *IEEE transactions on visualization and computer graphics*, vol. 26, no. 8, pp. 2535–2545, 2019.

[265] Y.-X. Lin, R. Venkatakrishnan, R. Venkatakrishnan, E. Ebrahimi, W.-C. Lin, and S. V. Babu, "How the presence and size of static peripheral blur affects cybersickness in virtual reality," *ACM Transactions on Applied Perception (TAP)*, vol. 17, no. 4, pp. 1–18, 2020.

[266] K. Carnegie and T. Rhee, "Reducing visual discomfort with hmds using dynamic depth of field," *IEEE computer graphics and applications*, vol. 35, no. 5, pp. 34–41, 2015.

[267] R. Hussain, M. Chessa, and F. Solari, "Mitigating cybersickness in virtual reality systems through foveated depth-of-field blur," *Sensors*, vol. 21, no. 12, p. 4006, 2021.

[268] J. J.-W. Lin, H. Abi-Rached, D.-H. Kim, D. E. Parker, and T. A. Furness, "A "natural" independent visual background reduced simulator sickness," in *Proceedings of the human factors and ergonomics society annual meeting*, vol. 46, no. 26. SAGE Publications Sage CA: Los Angeles, CA, 2002, pp. 2124–2128.

[269] H. B.-L. Duh, D. E. Parker, and T. A. Furness, "An "independent visual background" reduced balance disturbance envoked by visual scene motion: implication for alleviating simulator sickness," in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2001, pp. 85–89.

[270] ——, "An independent visual background reduced simulator sickness in a driving simulator," *Presence: Teleoperators & Virtual Environments*, vol. 13, no. 5, pp. 578–588, 2004.

[271] T. Nguyen-Vo, B. E. Riecke, and W. Stuerzlinger, "Simulated reference frame: A cost-effective solution to improve spatial orientation in vr," in *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2018, pp. 415–422.

[272] D. M. Whittinghill, B. Ziegler, T. Case, and B. Moore, "Nasum virtualis: A simple technique for reducing simulator sickness," in *Games developers conference (GDC)*, vol. 74, 2015.

[273] C. Wienrich, C. K. Weidner, C. Schatto, D. Obremski, and J. H. Israel, "A virtual nose as a rest-frame-the impact on simulator sickness and game experience," in *2018 10th international conference on virtual worlds and games for serious applications (VS-Games)*. IEEE, 2018, pp. 1–8.

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

[275] F. Bonato, A. Bubka, and W. W. Krueger, "A wearable device providing a visual fixation point for the alleviation of motion sickness symptoms," *Military medicine*, vol. 180, no. 12, pp. 1268–1272, 2015.

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

[277] Y. Fujii and T. Seno, "The effect of optical flow motion direction on vection strength," *i-Perception*, vol. 11, no. 1, p. 2041669519899108, 2020.

[278] S. H. Park, B. Han, and G. J. Kim, "Mixing in reverse optical flow to mitigate vection and simulation sickness in virtual reality," in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 2022, pp. 1–11.

[279] H. Buhler, S. Misztal, and J. Schild, "Reducing vr sickness through peripheral visual effects," in *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2018, pp. 517–9.

[280] P. Bala, I. Oakley, V. Nisi, and N. Nunes, "Staying on track: a comparative study on the use of optical flow in 360 video to mitigate vims," in *ACM International Conference on Interactive Media Experiences*, 2020, pp. 82–93.

[281] R. Lou, F. Mérienne, R. H. So, T.-T. Chan, and D. Bechmann, "Geometric simplification for reducing optic flow in vr," in *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 2022, pp. 682–685.

[282] R. Xiao and H. Benko, "Augmenting the field-of-view of head-mounted displays with sparse peripheral displays," in *Proceedings of the 2016 CHI conference on human factors in computing systems*, 2016, pp. 1221–1232.

[283] L. Andre and R. Coutellier, "Cybersickness and evaluation of a remediation system: a pilot study," in *2019 International Conference on 3D Immersion (IC3D)*. IEEE, 2019, pp. 1–6.

[284] Y. Ikei, V. Yem, K. Tashiro, T. Fujie, T. Amemiya, and M. Kitazaki, "Live stereoscopic 3d image with constant capture direction of 360 cameras for high-quality visual telepresence," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2019, pp. 431–439.

[285] K. K. Ra and J. J. Clark, "Decoupled hybrid 360° panoramic stereo video," in *2019 International conference on 3D vision (3DV)*. IEEE, 2019, pp. 386–394.

[286] C. Groth, J.-P. Tauscher, N. Heesen, M. Hattenbach, S. Castillo, and M. Magnor, "Omnidirectional galvanic vestibular stimulation in virtual reality," *IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 5, pp. 2234–2244, 2022.

[287] M. Sra, A. Jain, and P. Maes, "Adding proprioceptive feedback to virtual reality experiences using galvanic vestibular stimulation," in *Proceedings of the 2019 CHI conference on human factors in computing systems*, 2019, pp. 1–14.

[288] S. Weech, T. Wall, and M. Barnett-Cowan, "Reduction of cybersickness during and immediately following noisy galvanic vestibular stimulation," *Experimental Brain Research*, vol. 238, no. 2, pp. 427–437, 2020.

[289] S. Weech, J. Moon, and N. F. Troje, "Influence of bone-conducted vibration on simulator sickness in virtual reality," *PloS one*, vol. 13, no. 3, p. e0194137, 2018.

[290] S. Jung, R. Li, R. McKee, M. C. Whitton, and R. W. Lindeman, "Floor-vibration vr: mitigating cybersickness using whole-body tactile stimuli in highly realistic vehicle driving experiences," *IEEE Transactions on Visualization & Computer Graphics*, vol. 27, no. 05, pp. 2669–2680, 2021.

[291] Y. Sawada, Y. Itaguchi, M. Hayashi, K. Aigo, T. Miyagi, M. Miki, T. Kimura, and M. Miyazaki, "Effects of synchronised engine sound and vibration presentation on visually induced motion sickness," *Scientific reports*, vol. 10, no. 1, pp. 1–10, 2020.

[292] H. Yoon, S. Lee, J. Park, Y. Choi, and S. Cho, "Development of racing game using motion seat," in *2017 International Symposium on Ubiquitous Virtual Reality (ISUVR)*. IEEE, 2017, pp. 4–7.

[293] M. Rietzler, T. Hirzle, J. Gugenheimer, J. Frommel, T. Dreja, and E. Rukzio, "Vrspinning: Exploring the design space of a 1d rotation platform to increase the perception of self-motion in vr," in *Proceedings of the 2018 Designing Interactive Systems Conference*, 2018, pp. 99–108.

[294] S. Cmentowski, F. Kievelitz, and J. H. Krueger, "Outpace reality: A novel augmented-walking technique for virtual reality games," *Proceedings of the ACM on Human-Computer Interaction*, vol. 6, no. CHI PLAY, pp. 1–24, 2022.

[295] Y. Farmani and R. J. Teather, "Viewpoint snapping to reduce cybersickness in virtual reality," in *Proceedings of the 44th graphics interface conference*, 2018, pp. 168–175.

[296] S. Freitag, B. Weyers, and T. W. Kuhlen, "Efficient approximate computation of scene visibility based on navigation meshes and applications for navigation and scene analysis," in *2017 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, 2017, pp. 134–143.

[297] J. L. Dorado and P. A. Figueroa, "Ramps are better than stairs to reduce cybersickness in applications based on a hmd and a gamepad," in *2014 IEEE symposium on 3D user interfaces (3DUI)*. IEEE, 2014, pp. 47–50.

[298] P. Hu, Q. Sun, P. Didyk, L.-Y. Wei, and A. E. Kaufman, "Reducing simulator sickness with perceptual camera control," *ACM Transactions on Graphics (TOG)*, vol. 38, no. 6, pp. 1–12, 2019.

[299] M. Isaza, J. Zhang, K. Kim, C. Mei, and R. Guo, "Mono-stereoscopic camera in a virtual reality environment: Case study in cybersickness," in *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*. IEEE, 2019, pp. 1–4.

[300] P. Kourtesis, S. Collina, L. A. Doumas, and S. E. MacPherson, "Validation of the virtual reality neuroscience questionnaire: maximum duration of immersive virtual reality sessions without the presence of pertinent adverse symptomatology," *Frontiers in human neuroscience*, vol. 13, p. 417, 2019.

[301] S. Freitag, B. Weyers, and T. W. Kuhlen, "Examining rotation gain in cave-like virtual environments," *IEEE transactions on visualization and computer graphics*, vol. 22, no. 4, pp. 1462–1471, 2016.

[302] S. E. Stromberg, M. E. Russell, and C. R. Carlson, "Diaphragmatic breathing and its effectiveness for the management of motion sickness," *Aerospace medicine and human performance*, vol. 86, no. 5, pp. 452–457, 2015.

[303] R. Liu, C. Zhuang, R. Yang, and L. Ma, "Effect of economically friendly acustimulation approach against cybersickness in video-watching tasks using consumer virtual reality devices," *Applied Ergonomics*, vol. 82, p. 102946, 2020.

[304] A. Iskenderova, F. Weidner, and W. Broll, "Drunk virtual reality gaming: exploring the influence of alcohol on cybersickness," in *Proceedings of the annual symposium on computer-human interaction in play*, 2017, pp. 561–572.

[305] M. Kaufeld, K. De Coninck, J. Schmidt, and H. Hecht, "Chewing gum reduces visually induced motion sickness," *Experimental brain research*, vol. 240, no. 2, pp. 651–663, 2022.

[306] H.-C. Lien, W. M. Sun, Y.-H. Chen, H. Kim, W. Hasler, and C. Owyang, "Effects of ginger on motion sickness and gastric slow-wave dysrhythmias induced by circular vection," *American Journal of Physiology-Gastrointestinal and Liver Physiology*, vol. 284, no. 3, pp. G481–G489, 2003.

[307] C. Schartmüller and A. Riener, "Sick of scents: Investigating non-invasive olfactory motion sickness mitigation in automated driving," in *12th International conference on automotive user interfaces and interactive vehicular applications*, 2020, pp. 30–39.

[308] N. Ranasinghe, P. Jain, D. Tolley, S. Karwita Tailan, C. C. Yen, and E. Y.-L. Do, "Exploring the use of olfactory stimuli towards reducing visually induced motion sickness in virtual reality," in *Proceedings of the 2020 ACM Symposium on Spatial User Interaction*, 2020, pp. 1–9.

[309] A. Paroz and L. E. Potter, "Investigating external airflow and reduced room temperature to reduce virtual reality sickness," in *Proceedings of the 33rd Australian Conference on Human-Computer Interaction*, 2021, pp. 198–207.

[310] A. Mao, K. Barnes, L. Sharpe, A. L. Geers, S. G. Helfer, K. Faasse, and B. Colagiuri, "Using positive attribute framing to attenuate nocebo side effects: A cybersickness study," *Annals of Behavioral Medicine*, vol. 55, no. 8, pp. 769–778, 2021.

[311] J. Smyth, P. Jennings, P. Bennett, and S. Birrell, "A novel method for reducing motion sickness susceptibility through training visuospatial ability–a two-part study," *Applied Ergonomics*, vol. 90, p. 103264, 2021.

[312] D. Narciso, M. Bessa, M. Melo, A. Coelho, and J. Vasconcelos-Raposo, "Immersive 360 video user experience: impact of different variables in the sense of presence and cybersickness," *Universal Access in the Information Society*, vol. 18, pp. 77–87, 2019.

[313] J. Munafo, M. Diedrick, and T. A. Stoffregen, "The virtual reality head-mounted display oculus rift induces motion sickness and is sexist in its effects," *Experimental brain research*, vol. 235, pp. 889–901, 2017.

[314] E. C. Regan and K. R. Price, "The frequency of occurrence and severity of side-effects of immersion virtual reality." *Aviation, space, and environmental medicine*, 1994.

[315] S. Burigat and L. Chittaro, "Navigation in 3d virtual environments: Effects of user experience and location-pointing navigation aids," *International Journal of Human-Computer Studies*, vol. 65, no. 11, pp. 945–958, 2007. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1071581907000985

[316] J. Dodiya and V. N. Alexandrov, "Use of auditory cues for wayfinding assistance in virtual environment: music aids route decision," in *Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology*, ser. VRST '08. New York, NY, USA: Association for Computing Machinery, 2008, p. 171–174. [Online]. Available: https://doi.org/10.1145/1450579.1450615

[317] A. Parush and D. Berman, "Navigation and orientation in 3d user interfaces: the impact of navigation aids and landmarks," *International Journal of Human-Computer Studies*, vol. 61, no. 3, pp. 375–395, 2004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1071581904000448

[318] A. D. Angélique Montuwy and B. Cahour, "Helping older pedestrians navigate in the city: comparisons of visual, auditory and haptic guidance instructions in a virtual environment," *Behaviour & Information Technology*, vol. 38, no. 2, pp. 150–171, 2019. [Online]. Available: https://doi.org/10.1080/0144929X.2018.1519035

[319] R. Alghofaili, Y. Sawahata, H. Huang, H.-C. Wang, T. Shiratori, and L.-F. Yu, "Lost in style: Gaze-driven adaptive aid for vr navigation," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ser. CHI '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 1–12. [Online]. Available: https://doi.org/10.1145/3290605.3300578

[320] C. R. Bruns and B. C. Chamberlain, "The influence of landmarks and urban form on cognitive maps using virtual reality," *Landscape and Urban Planning*, vol. 189, pp. 296–306, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169204618303104

[321] B. Kruysman, A. P. Maneri, Y. Yang, K. Vaidyanathan, D. Webb, and M. Schoen, "Project hubble: Multi-user xr collaboration tool," in *ACM SIGGRAPH 2022 Immersive Pavilion*, 2022, pp. 1–2.

[322] A. Wang, M. Thompson, C. Uz-Bilgin, and E. Klopfer, "Authenticity, interactivity, and collaboration in virtual reality games: Best practices and lessons learned," *Frontiers in Virtual Reality*, vol. 2, p. 734083, 2021.

[323] C.-H. Chen and X. Li, "Visualizing off-screen targets: Effects of response time and visual cue design on users' wayfinding performance using a dynamic peephole interface," *International Journal of Human–Computer Interaction*, vol. 37, no. 13, pp. 1220–1230, 2021. [Online]. Available: https://doi.org/10.1080/10447318.2021.1876356

[324] R. Li, "Spatial learning in smart applications: enhancing spatial awareness through visualized off-screen landmarks on mobile devices," in *Smart Spaces and Places*. Routledge, 2021, pp. 89–101.

[325] T. Gonçalves, A. P. Afonso, M. B. Carmo, and P. Paulo, "Halodot: Visualization of the relevance of off-screen objects," *Proceedings of SIACG*, pp. 117–120, 2011.

[326] A. Cockburn, A. Karlson, and B. B. Bederson, "A review of overview+detail, zooming, and focus+context interfaces," *ACM Computing Surveys (CSUR)*, vol. 41, no. 1, 2009.

[327] K. Danyluk, B. Ens, B. Jenny, and W. Willett, "A design space exploration of worlds in miniature," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–15.

[328] S. S. Sarkar M, "Stretching the rubber sheet: A metaphor for viewing large layouts on small screens," in *Proceedings of the 6th annual ACM symposium on User interface software and technology*. Association for Computing Machinery, 1993, pp. 81–91.

[329] D. Miau and S. Feiner, "Personalized compass: A compact visualization for direction and location," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2016, pp. 5114–5125.

[330] F. Bork, C. Schnelzer, U. Eck, and N. Navab, "Towards efficient visual guidance in limited field-of-view head-mounted displays," *IEEE transactions on visualization and computer graphics*, vol. 24, no. 11, pp. 2983–2992, 2018.

[331] P. Perea, D. Morand, and L. Nigay, "Halo3d: A technique for visualizing off-screen points of interest in mobile augmented reality," in *Proceedings of the 29th Conference on l'Interaction Homme-Machine*, 2017, pp. 43–51.

[332] L. Chittaro and S. Burigat, "3d location-pointing as a navigation aid in virtual environments," in *InProceedings of the working conference on Advanced visual interfaces*, 2004, pp. 267–274.

[333] S. Feiner, B. MacIntyre, and D. Seligmann, "Knowledge-based augmented reality," *Communications of the ACM*, vol. 36, no. 7, pp. 53–62, 1993.

[334] F. Biocca, A. Tang, C. Owen, and F. Xiao, "Attention funnel: omnidirectional 3d cursor for mobile augmented reality platforms," in *Proceedings of the SIGCHI conference on Human Factors in computing systems*, 2006, pp. 1115–1122.

[335] M. Tonnis and G. Klinker, "Effective control of a car driver's attention for visual and acoustic guidance towards the direction of imminent dangers," in *2006 IEEE/ACM International Symposium on Mixed and Augmented Reality*, 2006, pp. 13–22.

[336] M. Tonnis, C. Sandor, G. Klinker, C. Lange, and H. Bubb, "Experimental evaluation of an augmented reality visualization for directing a car driver's attention," in *Fourth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'05)*. IEEE, 2005, pp. 56–59.

[337] U. Gruenefeld, D. Lange, L. Hammer, S. Boll, and W. Heuten, "Flyingarrow: pointing towards out-of-view objects on augmented reality devices," in *Proceedings of the 7th ACM international symposium on pervasive displays*, 2018, pp. 1–6.

[338] S. Hu, J. Malloch, and D. Reilly, "A comparative evaluation of techniques for locating out of view targets in virtual reality," in *Graphics Interface 2021*, 2021, pp. 1–11.

[339] U. Gruenefeld, L. Prädel, and W. Heuten, "Improving search time performance for locating out-of-view objects in augmented reality," in *Proceedings of Mensch Und Computer 2019*, 2019, pp. 481–485.

[340] R. E. Ziemer, W. H. Tranter, and D. R. Fannin, *Signals and systems: continous and discrete*. 978-0134964560: Prentice-Hall International Inc., 1984.

[341] F. H. Durgin, K. Gigone, and R. Scott, "Perception of visual speed while moving." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 31, no. 2, p. 339, 2005.

[342] M. van Gisbergen, M. Kovacs, F. Campos, M. van der Heeft, and V. Vugts, "What we don't know. the effect of realism in virtual reality on experience and behaviour," in *Augmented Reality and Virtual Reality*. Cham: Springer, 2019, pp. 45–57.

[343] A. Murata, "Effects of duration of immersion in a virtual reality environment on postural stability," *International Journal of Human-Computer Interaction*, vol. 17, no. 4, pp. 463–477, 2004.

[344] A. N. R. Chandra, F. El Jamiy, and H. Reza, "A survey on simulation sickness in virtual environments," 2021.

[345] Z. D. Siegel and J. W. Kelly, "Walking through a virtual environment improves perceived size within and beyond the walked space," *Attention, Perception, & Psychophysics*, vol. 79, no. 1, pp. 39–44, 2017.

[346] C. Archetti, M. G. Speranza, and D. Vigo, "Chapter 10: Vehicle routing problems with profits," in *Vehicle Routing: Problems, Methods, and Applications, Second Edition*. SIAM, 2014, pp. 273–297.

[347] I.-M. Chao, B. L. Golden, and E. A. Wasil, "The team orienteering problem," *European journal of operational research*, vol. 88, no. 3, pp. 464–474, 1996.

[348] L. Perron and V. Furnon, "Or-tools," Google, 2019. [Online]. Available: https://developers.google.com/optimization/

[349] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, "Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness," *The international journal of aviation psychology*, vol. 3, no. 3, pp. 203–220, 1993.

[350] B. Sarupuri, S. Jung, S. Hoermann, M. C. Whitton, and R. W. Lindeman, "Testbed evaluation of multi-travel mode in virtual reality," 2020.

[351] J. Marin-Morales, C. Llinares, J. Guixeres, and M. Alcañiz, "Emotion recognition in immersive virtual reality: From statistics to affective computing," *Sensors*, vol. 20, no. 18, 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/18/5163

[352] J. Heyse, M. Torres Vega, T. De Jonge, F. De Backere, and F. De Turck, "A personalised emotion-based model for relaxation in virtual reality," *Applied Sciences*, vol. 10, no. 17, 2020. [Online]. Available: https://www.mdpi.com/2076-3417/10/17/6124

[353] Y. Y. Kim, E. N. Kim, M. J. Park, K. S. Park, H. D. Ko, and H. T. Kim, "The application of biosignal feedback for reducing cybersickness from exposure to a virtual environment," *Presence: Teleoperators and Virtual Environments*, vol. 17, no. 1, pp. 1–16, 2008.

[354] T. Xue, A. E. Ali, G. Ding, and P. Cesar, "Investigating the relationship between momentary emotion self-reports and head and eye movements in hmd-based 360° vr video watching," in *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, ser. CHI EA '21. New York, NY, USA: Association for Computing Machinery, 2021. [Online]. Available: https://doi.org/10.1145/3411763.3451627

[355] L. Tabbaa, R. Searle, S. M. Bafti, M. M. Hossain, J. Intarasisrisawat, M. Glancy, and C. S. Ang, "Vreed: Virtual reality emotion recognition dataset using eye tracking & physiological measures," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 5, no. 4, dec 2022. [Online]. Available: https://doi.org/10.1145/3495002

[356] J. Z. Lim, J. Mountstephens, and J. Teo, "Emotion recognition using eye-tracking: taxonomy, review and current challenges," *Sensors*, vol. 20, no. 8, p. 2384, 2020.

[357] B. Sarupuri, S. Hoermann, F. Steinicke, and R. W. Lindeman, "Triggerwalking: a biomechanically-inspired locomotion user interface for efficient realistic virtual walking," in *Proceedings of the 5th Symposium on Spatial User Interaction*, ser. SUI '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 138–147. [Online]. Available: https://doi.org/10.1145/3131277.3132177

[358] J. Sykes and S. Brown, "Affective gaming: measuring emotion through the gamepad," in *CHI '03 Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA

'03. New York, NY, USA: Association for Computing Machinery, 2003, p. 732–733. [Online]. Available: https://doi.org/10.1145/765891.765957

[359] S. Brambilla, G. Boccignone, N. A. Borghese, L. A. Ripamonti *et al.*, "Between the buttons: Stress assessment in video games using players' behavioural data," in *Scitepress Proceedings*, 2022, pp. 59–69.

[360] J. A. Russell, "A circumplex model of affect." *Journal of personality and social psychology*, vol. 39, no. 6, p. 1161, 1980.

[361] T. Eerola and J. K. Vuoskoski, "A comparison of the discrete and dimensional models of emotion in music," *Psychology of Music*, vol. 39, no. 1, pp. 18–49, 2011.

[362] D. Västfjäll, "Emotion induction through music: A review of the musical mood induction procedure," *Musicae Scientiae*, vol. 5, no. 1_suppl, pp. 173–211, 2001.

[363] B. J. Li, J. N. Bailenson, A. Pines, W. J. Greenleaf, and L. M. Williams, "A public database of immersive vr videos with corresponding ratings of arousal, valence, and correlations between head movements and self report measures," *Frontiers in psychology*, vol. 8, p. 2116, 2017.

[364] S. C. Steinhaeusser, S. Oberdörfer, S. von Mammen, M. E. Latoschik, and B. Lugrin, "Joyful adventures and frightening places–designing emotion-inducing virtual environments," *Frontiers in Virtual Reality*, vol. 3, p. 919163, 2022.

[365] M. M. Bradley and P. J. Lang, "Measuring emotion: the self-assessment manikin and the semantic differential," *Journal of behavior therapy and experimental psychiatry*, vol. 25, no. 1, pp. 49–59, 1994.

[366] S. Tikadar and S. Bhattacharya, "A novel method to build and validate an affective state prediction model from touch-typing," in *Human-Computer Interaction – INTERACT 2019*, D. Lamas, F. Loizides, L. Nacke, H. Petrie, M. Winckler, and P. Zaphiris, Eds. Cham: Springer International Publishing, 2019, pp. 99–119.

[367] D. Maheshwari, S. Ghosh, R. Tripathy, M. Sharma, and U. R. Acharya, "Automated accurate emotion recognition system using rhythm-specific deep convolutional neural network technique with multi-channel eeg signals," *Computers*

in Biology and Medicine, vol. 134, p. 104428, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0010482521002225

[368] J. N. Latta and D. J. Oberg, "A conceptual virtual reality model," *IEEE Computer Graphics and Applications*, vol. 14, no. 1, pp. 23–29, 1994.

[369] M. Slater and M. V. Sanchez-Vives, "Enhancing our lives with immersive virtual reality," *Frontiers in Robotics and AI*, vol. 3, p. 74, 2016.

[370] J. Kim, J. Cha, H. Lee, and S. Kim, "Hand-free natural user interface for vr hmd with ir based facial gesture tracking sensor," in *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, 2017, pp. 1–2.

[371] D. Bersak, G. McDarby, N. Augenblick, P. McDarby, D. McDonnell, B. McDonald, and R. Karkun, "Intelligent biofeedback using an immersive competitive environment," in *Paper at the designing ubiquitous computing games workshop at UbiComp*, 2001, pp. 1–6.

[372] L. E. Nacke, M. Kalyn, C. Lough, and R. L. Mandryk, "Biofeedback game design: using direct and indirect physiological control to enhance game interaction," in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2011, pp. 103–112.

[373] A. Parnandi, B. Ahmed, E. Shipp, and R. Gutierrez-Osuna, "Chill-out: Relaxation training through respiratory biofeedback in a mobile casual game," in *International Conference on Mobile Computing, Applications, and Services*. Springer, 2013, pp. 252–260.

[374] T. M. Porcino, E. Clua, D. Trevisan, C. N. Vasconcelos, and L. Valente, "Minimizing cyber sickness in head mounted display systems: Design guidelines and applications," in *2017 IEEE 5th International Conference on Serious Games and Applications for Health (SeGAH)*, 2017, pp. 1–6.

[375] J. Desnoyers-Stewart, E. R. Stepanova, P. Pasquier, and B. E. Riecke, "Jel: Connecting through breath in virtual reality," in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–6.

[376] C. Rockstroh, J. Blum, and A. S. Göritz, "A mobile vr-based respiratory biofeedback game to foster diaphragmatic breathing," *Virtual Reality*, vol. 25, pp. 539–552, 2021.

[377] C. Davies and J. Harrison, "Osmose: towards broadening the aesthetics of virtual reality," *ACM SIGGRAPH Computer Graphics*, vol. 30, no. 4, pp. 25–28, 1996.

[378] R. Bossenbroek, A. Wols, J. Weerdmeester, A. Lichtwarck-Aschoff, I. Granic, and M. M. J. W. van Rooij, "Efficacy of a virtual reality biofeedback game (deep) to reduce anxiety and disruptive classroom behavior: Single-case study," *JMIR Ment Health*, vol. 7, no. 3, p. e16066, Mar 2020. [Online]. Available: http://mental.jmir.org/2020/3/e16066/

[379] F. Soyka, M. Leyrer, J. Smallwood, C. Ferguson, B. E. Riecke, and B. J. Mohler, "Enhancing stress management techniques using virtual reality," in *Proceedings of the ACM Symposium on Applied Perception*, ser. SAP '16. New York, NY, USA: Association for Computing Machinery, 2016, p. 85–88. [Online]. Available: https://doi.org/10.1145/2931002.2931017

[380] M. Prpa, E. R. Stepanova, T. Schiphorst, B. E. Riecke, and P. Pasquier, "Inhaling and exhaling: How technologies can perceptually extend our breath awareness," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 1–15. [Online]. Available: https://doi.org/10.1145/3313831.3376183

[381] E. M. Murtagh, J. L. Mair, E. Aguiar, C. Tudor-Locke, and M. H. Murphy, "Outdoor walking speeds of apparently healthy adults: A systematic review and meta-analysis," *Sports Medicine*, vol. 51, pp. 125–141, 2021.

[382] D. Zielasko, Y. C. Law, and B. Weyers, "Take a look around–the impact of decoupling gaze and travel-direction in seated and ground-based virtual reality utilizing torso-directed steering," in *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2020, pp. 398–406.

[383] Z. Zhao, Y. Li, and H.-N. Liang, "Leanon: Simulating balance vehicle locomotion in virtual reality," in *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 2023, pp. 415–424.

[384] A. Kitson, A. M. Hashemian, E. R. Stepanova, E. Kruijff, and B. E. Riecke, "Comparing leaning-based motion cueing interfaces for virtual reality locomotion," in *2017 IEEE Symposium on 3d user interfaces (3DUI)*. IEEE, 2017, pp. 73–82.

[385] C. Chourpiliadis and A. Bhardwaj, "Physiology, respiratory rate," 2019. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK537306/

[386] C.-H. Shih, N. Tomita, Y. X. Lukic, Á. H. Reguera, E. Fleisch, and T. Kowatsch, "Breeze: Smartphone-based acoustic real-time detection of breathing phases for a gamified biofeedback breathing training," *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies*, vol. 3, no. 4, pp. 1–30, 2019.

[387] A. Cannavò, D. Calandra, F. G. Pratticò, V. Gatteschi, and F. Lamberti, "An evaluation testbed for locomotion in virtual reality," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 3, pp. 1871–1889, 2021.

[388] A. Likas, N. Vlassis, and J. J. Verbeek, "The global k-means clustering algorithm," *Pattern recognition*, vol. 36, no. 2, pp. 451–461, 2003.

[389] M. Ester, H.-P. Kriegel, J. Sander, X. Xu *et al.*, "A density-based algorithm for discovering clusters in large spatial databases with noise." in *kdd*, vol. 96, 1996, pp. 226–231.

[390] M. B. Carmo, A. P. Afonso, A. Ferreira, A. P. Cláudio, and G. Silva, "Poi awareness, relevance and aggregation for augmented reality," in *2016 20th International Conference Information Visualisation (IV)*. IEEE, 2016, pp. 300–305.

[391] D. Chatzopoulos and P. Hui, "Readme: A real-time recommendation system for mobile augmented reality ecosystems," in *Proceedings of the 24th ACM international conference on Multimedia*, 2016, pp. 312–316.

[392] K. Y. Lam, L. H. Lee, and P. Hui, "A2w: Context-aware recommendation system for mobile augmented reality web browser," in *Proceedings of the 29th ACM International Conference on Multimedia*, 2021, pp. 2447–2455.

List of Publications

Publications from Thesis Work:

Refereed Journals

1. **Nilotpal Biswas**, Arpit Singh, and Samit Bhattacharya, "Augmented 3D arrows for visualizing off-screen Points of Interest without clutter", *Displays*, Volume 79, Issue 1, pp. 102502, Sept 2023.
   DOI: `https://doi.org/10.1016/j.displa.2023.102502` [Chapter 7]

2. **Nilotpal Biswas**, Debangshu Banerjee, and Samit Bhattacharya, "Realistic walking experience for system-automated virtual reality tour.", *Multimedia Tools and Applications*, Volume 82, Issue 11, pp. 17453-17470, May 2023.
   DOI: `https://doi.org/10.1007/s11042-022-14035-z` [Chapter 3]

3. **Nilotpal Biswas**, Anamitra Mukherjee and Samit Bhattacharya, ""Are you feeling sick?" – A systematic literature review of cybersickness in virtual reality.", *ACM Computing surveys*, Volume 56, Issue 11, pp. 1-38, June 2024.
   DOI: `https://doi.org/10.1145/3670008` [Chapter 2]

Refereed Conferences

4. **Nilotpal Biswas**, Debangshu Banerjee, and Samit Bhattacharya, "Minimising the duration of a system-controlled virtual reality tour", *In 2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Singapore, October 2022, pp. 600-604, IEEE*
   DOI: https://doi.org/10.1109/ISMAR-Adjunct57072.2022.00124 [Chapter 4]

5. **Nilotpal Biswas**, Arpit Singh, and Samit Bhattacharya, "AroundArrow: Off-Screen POI visualization for handheld Augmented Reality in vertically dense regions ", *In 2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Singapore, October 2022, pp. 570-575, IEEE*
   DOI: https://doi.org/10.1109/ISMAR-Adjunct57072.2022.00119 [Chapter 7]

6. **Nilotpal Biswas**, Debangshu Banerjee, and Samit Bhattacharya, "Natural walking speed prediction in Virtual Reality while using target selection-based locomotion", *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology, New York, NY, USA, October 2021, pp. 1-3, ACM.*
   DOI: https://doi.org/10.1145/3489849.3489944 [Chapter 3]

7. **Nilotpal Biswas**, and Samit Bhattacharya, "Finding a range of perceived natural visual walking speed for stationary travelling techniques in VR", *In 2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Bari, Italy, October 2021, pp. 209-211, IEEE*
   DOI: https://doi.org/10.1109/ISMAR-Adjunct54149.2021.00050 [Chapter 4]

Under Review Publications

8. **Nilotpal Biswas**, and Samit Bhattacharya, "Recognising Emotional State using consumer-based HMD and VR controllers while navigating in a VR environment.", *Multimedia Tools and Applications (MTAP).* [Chapter 5]

9. **Nilotpal Biswas**, and Samit Bhattacharya, "BreathWalk: Enhancing VR Navigation through Breathing-Based Locomotion.", *India HCI 2024* . [Chapter 6]

Brief Bio

Nilotpal Biswas began his Ph.D. program in the Department of Computer Science and Engineering (CSE) at the Indian Institute of Technology (IIT) Guwahati, India, in December 2017. At IIT Guwahati, he was part of the User-Centric Computing and Networking (UCNET) Lab within the CSE Department. Before embarking on his doctoral studies, he served as an assistant project engineer in the CSE Department at IIT Guwahati. Throughout his Ph.D., he also worked as a teaching assistant for Coursera and NPTEL MOOCs, working under the guidance of Dr. Samit Bhattacharya. He earned his Bachelor of Technology (B.Tech) degree in CSE from the Central Institute of Technology Kokrajhar, India, in 2016. His doctoral research is centered on Navigation in Extended Reality. Outside of his academic pursuits, Nilotpal enjoys singing, playing musical instruments, reading, and painting.

Contact Information

E-mail: nilot176101101@iitg.ac.in
nilotpal.biswas126@gmail.com

Address: S/o: Subhash Chandra Biswas, Lachit road, Mazbat,
PO/PS: Mazbat, Dist: Udalguri, Assam - 784507, India