Development of trustworthy intelligent avatars in virtual immersion

Azadeh Hadadi

ÉCOLE DOCTORALE SCIENCES ET MÉTIERS DE L'INGÉNIEUR

Laboratoire d'Ingénierie des Systèmes Physiques et Numériques (**LISPEN**)

Campus de Cluny

THÈSE

présentée par : **Azadeh HADADI**
soutenue le : **25 Novembre 2024**

pour obtenir le **Doctorat** *délivré par*

L'Ecole Nationale Supérieure d'Arts et Métiers

Spécialité : **Informatique**

Development of trustworthy intelligent avatars in virtual immersion

THESE dirigée par :
M. CHARDONNET Jean-Rémy et Mme. OVTCHAROVA Jivka

et co-encadrée par :
M. GUILLET Christophe

Jury

Prof. Antonio CAPOBIANCO	Professeur des Universités ICube, Université de Strasbourg	Président
Prof. Eike LANGBEHN	Professeur, HAW Hamburg	Rapporteur
Prof. Franck GECHTER	Professeur des Universités CIAD, UTBM	Rapporteur
Prof. Andres BUSTILLO	Professeur, Université de Burgos, Espagne	Rapporteur
Prof. Jivka OVTCHAROVA	Professeure, IMI, Karlsruhe Institute of Technology	Co-directrice
Dr. Christophe GUILLET	Maître de Conférences LISPEN, Arts et Métiers	Co-encadrant
Prof. Jean-Rémy CHARDONNET	Professeure des Universités LISPEN, Arts et Métiers	Directeur
Prof. Barbara DEML	Professeure, IFAB, Karlsruhe Institute of Technology	Invitée

THÈSE

Acknowledgments

Throughout the writing of this dissertation, I have received immense support and encouragement from many wonderful individuals. This thesis would not have been possible without their inspiration and assistance.

I owe my deepest gratitude to my supervisors, Prof. Jean-Rémy Chardonnet, Dr. Christophe Guillet, and Prof. Dr. Jivka Ovtcharova. Their constructive feedback, valuable suggestions, enthusiastic discussions, and steadfast support were instrumental in the completion of this thesis. I appreciate the constructive advice and comments from the thesis committee.

I would also like to thank Dr. Mikhail Langovoy, who supervised me during my first year. I sincerely appreciate his time and attention throughout that pivotal period.

My warmest appreciation goes to M. Ali Mirzaei, who served as my co-supervisor. His expertise and deep knowledge were invaluable to my research, and his insightful feedback greatly sharpened my thinking, elevating the quality of my work. His guidance in the realms of virtual reality and research has been a significant asset to this thesis.

I am profoundly grateful to all the members of the Institut Arts et Métiers. I especially thank Prof. Frédéric Merienne, the former director of the institute, as well as Géraldine Roux, Thomas Lamy, and Julien Ryard for their unwavering support during my time in France and beyond—merci beaucoup! I would also like to acknowledge my Ph.D. colleagues, particularly Yuyang Wang and Théo Combe, who graduated in recent years. Their continued support after graduation has meant a great deal to me and has significantly enriched my experience.

I extend my gratitude to the members of the Karlsruhe Institute of Technology (KIT) for their support, particularly Anjela Mayer, Tengyu Li, Victor Häfner, Felix Michels, Michael Grethler, Sonja Ohler, and Thomas Mayer. My sincere thanks go to Polina Häfner for her invaluable advice throughout

my experimental work. Her expertise and unwavering support provided me with the confidence to overcome challenges.

I also wish to express my gratitude to the French-German Institute program for funding my stay at the Karlsruhe Institute of Technology.

Finally, my deepest and sincerest gratitude goes to my family for their continuous and unparalleled love and support. I am thankful to my sisters for always being there for me as friends. I am forever indebted to my parents for providing me with the opportunities and experiences that have shaped who I am today. They selflessly encouraged me to explore new directions in life and pursue my own destiny. This journey would not have been possible without them, and I dedicate this milestone to them.

Résumé

La réalité virtuelle (VR) est de plus en plus adoptée dans des domaines tels que l'industrie, l'éducation et la santé, mais des défis freinent son acceptation complète. Les avatars VR, destinés à représenter les utilisateurs dans des environnements virtuels, ne tiennent souvent pas compte des états physiologiques, des réponses émotionnelles et des nuances comportementales des utilisateurs, entraînant des problèmes tels que le cybermalaise, et l'inconfort.

Cette thèse introduit un avatar intelligent et digne de confiance visant à améliorer l'immersion virtuelle. La "fiabilité" englobe la perception de la fiabilité et de la crédibilité par les utilisateurs, l'adaptabilité étant cruciale. En utilisant l'intelligence artificielle (AI) et la fusion de capteurs, nous proposóns "SmartSimVR", un cadre flexible pour développer des avatars intelligents auto-adaptatifs.

Pour valider ce concept, nous avons implémente SmartSimVR dans un cas d'utilisation de simulateur de conduite, où la voiture agit comme un avatar non humanoïde représentant l'utilisateur. Le modèle vise à minimiser le cybermalaise, en apprenant continuellement à partir de données objectives et subjectives. Si le cybermalaise, est détecté, l'avatar ajuste son accélération, créant une expérience personnalisée.

Une expérience a comparé un avatar intelligent centré sur l'humain avec un homologue non intelligent pour évaluer l'impact sur la réduction du cybermalaise,. Notre analyse complète, couvrant divers domaines—y compris le temps, la fréquence, le temps-fréquence et les aspects topologiques—a produit des résultats éclairants, fournissant une base solide pour nos conclusions.

Notamment, en se concentrant sur les métriques d'énergie par ondelettes, nous avons observé un effet positif dans une gamme de mesures physiologiques et comportementales dans la condition de l'avatar intelligent par rapport à l'avatar non intelligent.

De plus, notre analyse des données a identifié plusieurs mesures objectives qui pourraient servir

de signatures du cybermalaise,. Nous avons noté des réductions significatives et substantielles de la vitesse oculaire, de la vitesse angulaire des yeux et de l'activité électrodermale (EDA) dans la condition de l'avatar intelligent à travers toutes les analyses. Ces résultats suggèrent fortement que ces signaux physiologiques sont en effet des indicateurs sensibles de l'expérience de cybermalaise„ renforçant ainsi l'efficacité de l'avatar intelligent.

L'analyse de corrélation a encore renforcé nos résultats, révélant de fortes corrélations positives entre l'énergie par ondelettes de la vitesse oculaire, de la vitesse angulaire des yeux et des signaux EDA avec les scores globaux du Questionnaire de Mal de Simulateur (SSQ). Les résultats subjectifs et objectifs fournissent des preuves convaincantes de l'efficacité de l'avatar intelligent auto-adaptatif dans l'atténuation du cybermalaise, et l'amélioration de l'expérience utilisateur en VR, soutenant ainsi notre hypothèse.

Mots-clés : Avatar, Intelligence Artificielle, Immersion Virtuelle, Fiabilité, Fusion de capteurs, Simulation de conduite en temps Réel, cybermalaise, Étude utilisateur.

Abstract

Virtual reality (VR) is becoming widely adopted in fields like industry, education, and healthcare, yet challenges hinder its full acceptance. VR avatars, intended to represent users in virtual environments, often fail to consider users' physiological states, emotional responses, and behavioral nuances, leading to issues like cybersickness and discomfort.

This thesis introduces a trustworthy, intelligent avatar aimed at enhancing virtual immersion. "Trustworthiness" encompasses users' perceived reliability and credibility, with adaptability crucial. Using artificial intelligence (AI) and sensor fusion, we propose "SmartSimVR," a flexible framework for developing auto-adaptive intelligent avatars.

To validate this concept, we implemented SmartSimVR in a driving simulator use case, where the car acts as a non-humanoid avatar representing the user. The model aims to minimize cybersickness by continuously learning from objective and subjective data. If cybersickness is detected, the avatar adjusts its acceleration, creating a customized experience.

An experiment compared a human-centered intelligent avatar with a non-intelligent counterpart to assess the impact on reducing cybersickness. Our comprehensive analysis, spanning various domains—including time, frequency, time-frequency, and topological aspects—yielded insightful results, providing a robust foundation for our conclusions.

Notably, focusing on wavelet energy metrics, we observed positive effect in a range of physiological and behavioral measures in the intelligent avatar condition compared to the non-intelligent avatar.

Additionally, our data analysis identified several objective measures that may serve as signatures of cybersickness. We noted significant and substantial reductions in eye velocity, eye angular velocity, and electrodermal activity (EDA) for the intelligent avatar condition across all analyses. These findings strongly suggest that these physiological signals are indeed sensitive indicators of the cybersickness

experience, further reinforcing the effectiveness of the intelligent avatar.

Correlation analysis further reinforced our findings, revealing strong positive correlations between the wavelet energy of eye velocity, eye angular velocity, and EDA signals with overall Simulator Sickness Questionnaire (SSQ) scores. Both subjective and objective results provide compelling evidence of the auto-adaptive intelligent avatar's effectiveness in mitigating cybersickness and enhancing the user experience in VR, thereby supporting our hypothesis.

Keywords : Avatar, Artificial Intelligence, Virtual immersion, Trustworthiness, Sensor fusion, Real-time driving simulation, Cybersickness, User study.

Contents

List of Tables

List of Figures

Introduction

Context

The present thesis was carried out under a cotutelle between France and Germany. The research centers participating in this joint thesis were the LISPEN ("Laboratoire d'Ingénierie des Systémes Physiques et Numériques (EA 7515)") and IMI ("Institut für Informationsmanagement im Ingenieurwesen"), respectively belonging to ENSAM ("Ecole Nationale Supérieure d'Arts et Métiers") in France and KIT ("Karlsruhe Institute of Technology"), under the "French-German Institute for Industry of the Future". It was partly supported by a French-German University (UFA-DFH) grant, No. CDFA 03-19. Adhering to the predefined plan, the timetable was structured into two phases. The first phase encompassed an extensive exploration of theoretical foundations and pertinent literature on user experience in virtual immersion, virtual avatars, behavioral factors, and topological data analysis. This phase was meticulously conducted at the French partner institute. The subsequent phase centered on developing the final demonstrator, entailing experimenting with a panel of users. This pivotal phase was carried out at the German partner institute.

Background and motivation

Virtual Reality (VR) has emerged as a transformative technology, offering immersive experiences that blur the boundaries between the physical and digital realms. Central to the VR experience are avatars, digital representations of users within the virtual environment. Avatars facilitate user engagement and interaction, bridging the real and virtual worlds. However, the effectiveness of avatars in accurately representing users is a topic of ongoing exploration and debate.

Avatars in VR possess two primary aspects: appearance and functionality. Appearance focuses on visual resemblance, striving to create digital representations that closely mirror users' physical

attributes. Advancements in graphics and animation technologies have made significant strides in achieving lifelike avatars, enabling users to recognize themselves visually within the virtual space. This aspect of avatars aims to foster familiarity and self-identification, contributing to a more immersive user experience.

On the other hand, functionality pertains to the avatar's capabilities and reliability in replicating the user's actions and behaviors. Avatars can mimic the user's movements, gestures, and expressions in real-time by leveraging motion capture systems, skeletal tracking, and natural language processing. This aspect aims to enhance the sense of embodiment and immersion, enabling users to engage with the virtual environment more intuitively and naturally.

However, despite significant advancements in appearance and functionality, it is essential to assess the safety and consideration of user states in avatar representation. Immersive environments profoundly impact users' psychological and physiological well-being, making safety in VR paramount. Currently, avatars often fail to provide a truly safe representation of users, neglecting their emotional and physical conditions. By ensuring that avatars are responsive to users' states and reactions, we can foster a more supportive and enriching VR experience.

The central motivation of this thesis was introducing the concept of a trustworthy, intelligent avatar for virtual immersion. In the context of VR avatars, "trustworthiness" refers to the perceived reliability, credibility, and confidence that users experience when interacting with the virtual environment. Trustworthiness in avatars arises from a range of factors, with one of the most crucial being adaptability enabling them to dynamically adjust their behavior, responses, and environmental interactions. By applying artificial intelligence (AI) techniques, avatars can be given this crucial adaptable nature. The result is a trustworthy intelligent avatar - a customized, self-adapting avatar equipped with robust AI capabilities.

Contribution

Motivated by the need to bridge the gap between avatars and user's state and safety, this thesis aims to critically examine the current state of avatars in VR, emphasizing their functionality aspects and explicitly focusing on the development of first-person avatars. By exploring emerging technologies and research directions, this research seeks to contribute to developing more secure, adaptive, and

user-centered avatar systems that align with the needs and preferences of human users in the virtual realm.

These avatars would be equipped with human-centered AI to align with user state. To achieve this goal, a crucial consideration was given to the physiological and behavioral state of the users. This virtual model facilitates auto-adaptation and enhances the user's sense of engagement during virtual experiments. In other words, it served as an innovative and dynamic representation of the user, effectively capturing their current state and responsively addressing their specific needs within the virtual environment. The thesis makes several significant contributions in terms of scientific and technical developments, including:

- Conducted a comprehensive literature review on intelligent avatars and provided a clear, well-defined definition for trustworthy intelligent avatar concept.
- Developed a novel, modular, and scalable framework for creating a real-time, customizable human-centric intelligent avatar that dynamically adapts to user state. The framework utilizes sensor fusion techniques to integrate data from multiple sensors seamlessly.
- Conducted comprehensive data analysis and studies across diverse domains, including statistical and topological analyses, to investigate the impact of a trustworthy intelligent avatar on user experience.

Manuscript Organization

Figure 1 illustrates the thesis structure and the interrelationships between the various chapters. The research content focuses primarily on avatars. The following provides a summary of the content covered in each thesis chapter.

Chapter 1's literature review covers several vital areas. It examines the concepts of VR and immersion. The chapter also delves into avatars in VR, including discussions on their visual appearance, interactivity, the Proteus effect, functional capabilities, and embodiment. Additionally, the literature review defines "trustworthiness" in the context of VR. Finally, the chapter explores the concept of intelligent avatars and their characteristics.

Chapter 2 delves into the research approach, elaborates on the thesis's specific objectives in greater detail, and presents and discusses the associated research questions.

Chapter 3 focuses on the core concept of a "Trustworthy Intelligent Avatar," discussing aspects like adaptive avatar behavior, the hybrid avatar/agent model, human-centered intelligent avatars, and the challenges of real-time learning.

Chapter 4 investigates the integration of AI and machine learning (ML) in VR platforms. It highlights vital AI/ML applications, the mathematical foundations of machine learning, and the benefits of stream learning. The chapter presents "SmartSimVR's" software architecture, implemented in our use case as a driving simulator, where users control a car as a non-humanoid avatar. This approach enhances user representation and immersion while focusing on safety by monitoring user states to reduce cybersickness through AI-driven prediction and detection. It also explores cybersickness causes and alleviation strategies and discusses multi-modal measurement techniques, including speech-driven self-reporting for enhanced user experiences in VR.

The experimental design, data collection procedure, and participant information are covered in Chapter 5. Then this chapter comprehensively explores the results and analysis, encompassing assessing model performance during training, statistical analysis, the time domain, the frequency domain, and the time-frequency domain, as well as the advantages of the wavelet transform. It also delves into the convergence of objective and subjective measures and persistent homology to extract insights from the data.

The concluding chapter (Chapter 6) summarizes the study's key findings and highlights the research's overall contributions and implications. Additionally, it outlines potential future research directions.

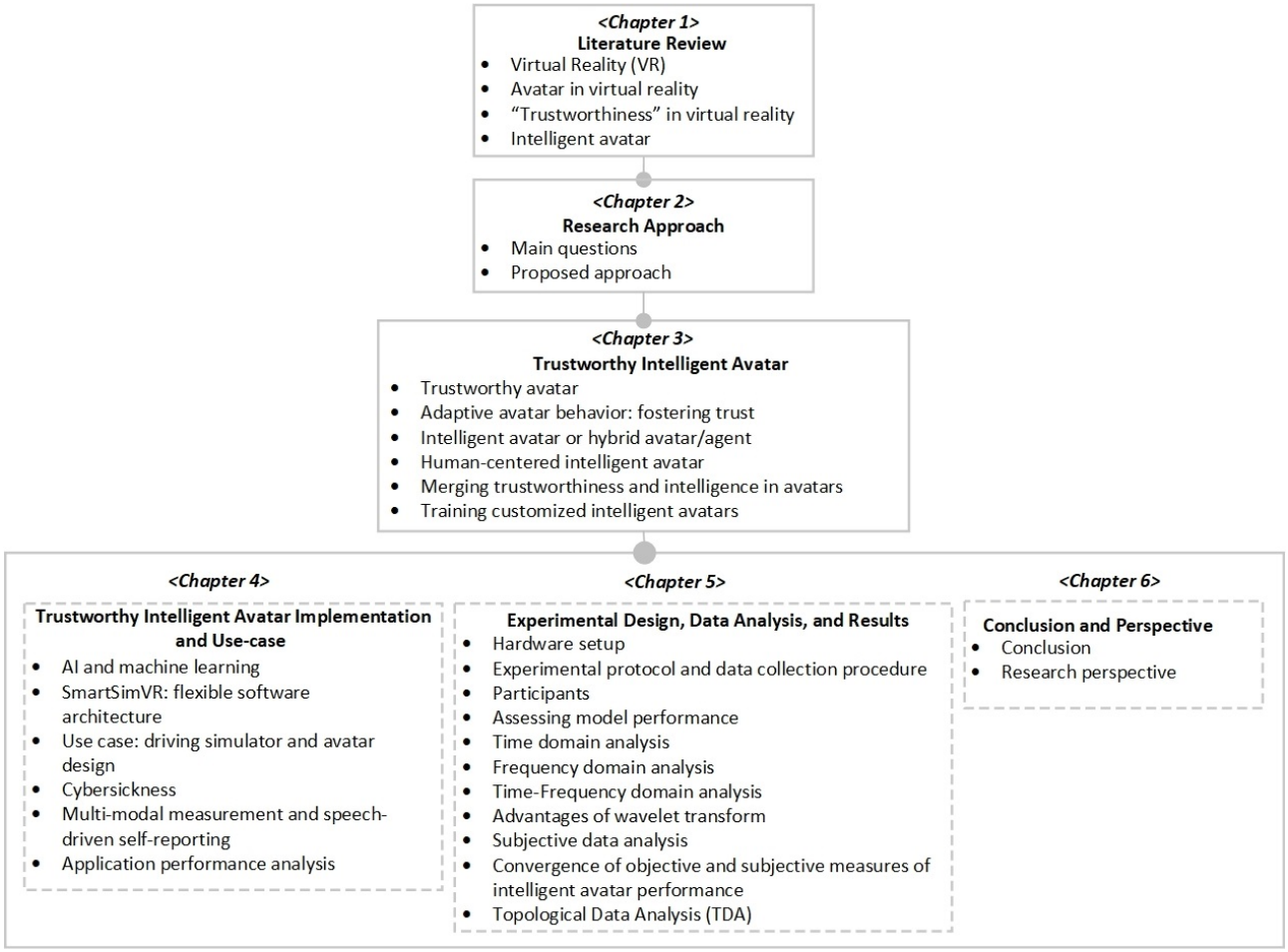


Figure 1: Overview of the thesis structure

Chapter 1

Literature Review

Content

This chapter introduces Virtual Reality (VR) as a transformative technology that immerses users in simulated environments, enhancing their interactions and experiences. A central focus is on avatars representing users in these virtual spaces. The chapter examines various aspects of avatars, including their interactivity, design (humanoid vs. non-humanoid), and the Proteus Effect, which explores the influence of avatars on user behavior.

Additionally, avatars' functional capabilities and embodiment are discussed, emphasizing their role in facilitating communication and user engagement. The emergence of intelligent avatars introduces a new dimension, as they can adapt and respond to user interactions, further enhancing the VR experience.

1.1 Virtual Reality (VR)

> Virtual reality (VR) is a computer-generated simulation of a three-dimensional environment that allows users to interact with and navigate through that environment using specialized hardware.

Simulated environments have a rich history that predates the term "Virtual Reality (VR)," with notable examples such as the Sensorama and the Headsight. The Sensorama (see Figure 1.1), developed by Morton Heilig in 1961 [91], provided users with a multi-sensory experience by combining visuals, sound, vibration, and scents to simulate a motorcycle ride. Similarly, the Headsight, created by Philco Corporation in the 1960s [143], employed a head-mounted display system that allowed users to view remote locations through a camera mounted on a moving platform, enabling exploration and remote operation. Figure 1.2 demonstrates the Headsight.

These early endeavors in simulating environments laid the groundwork for the development of VR. The term "Virtual Reality (VR)" gained popularity in the late eighties, representing a digital realm created through computer technology that appears to exist but does not physically exist in the real world. Since then, VR hardware has significantly advanced, enabling increasingly immersive and interactive experiences.

VR is a concept situated on a spectrum known as the Reality-Virtuality Continuum. This framework, initially introduced by Milgram and Kishino in 1994 [148], distinguishes between the "real world" at one end and "VR" at the other as depicted in Figure 1.3. The real world refers to the physical reality

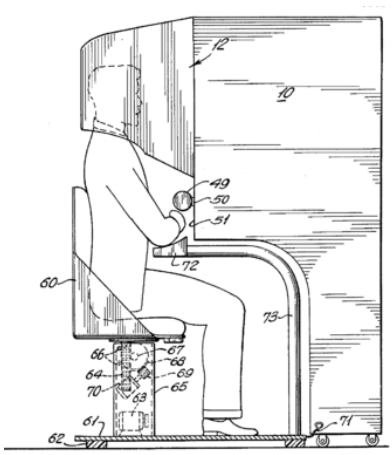


Figure 1.1: The Sensorama [91]

Figure 1.2: Headsight by Philco's Engineer [143]

we perceive and interact with. In contrast, VR represents a fully simulated environment that immerses users in an artificial realm. Between these extremes lies the realm of "mixed reality," which blends the real world and virtual/simulated elements to create a hybrid experience. The Reality-Virtuality Continuum has been revised recently, acknowledging that VR experiences can encompass more than just visual displays [210]. The updated perspective recognizes the importance of stimulating all senses within virtual environments (VE). It highlights that many current VR applications fall within mixed reality (see Figure 1.4).

Various definitions of VR have been proposed over the years. For instance, after analyzing existing definitions, Steuer [222] stated, "A VR is defined as a real or simulated environment in which a perceiver experiences telepresence". Another definition by Arnaldi et al. [13] describes it as a scientific and technical field that utilizes computer science and behavioral interfaces to simulate the behavior of

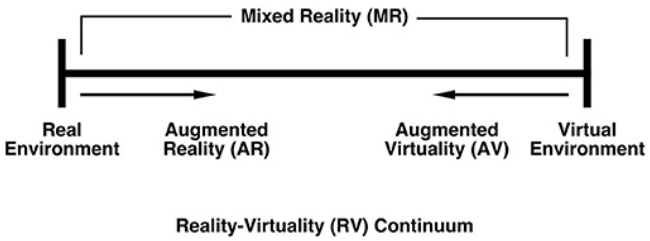


Figure 1.3: Milgram and Kishino's reality-virtuality continuum [148]

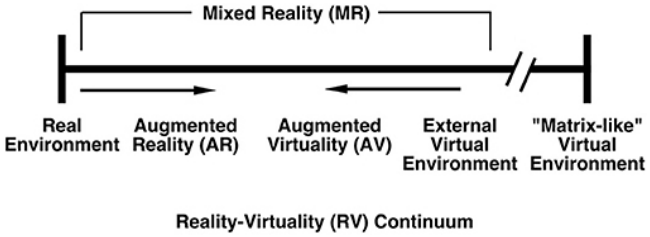


Figure 1.4: Revised Reality-Virtuality Continuum: Highlighting the inclusion of the External Virtual Environment (traditionally called "Virtual Reality") within the Mixed Reality (MR) spectrum [210]

3D entities in a virtual world. This simulated world interacts in real-time with users through sensory-motor channels, creating a pseudo-natural immersive experience. Han et al. [86] highlight that VR has emerged as a powerful technology that enables users to immerse themselves in computer-generated environments, offering a unique and interactive experience beyond the boundaries of the physical world. Škola et al. [211] define VR as "a set of images and sounds, produced by a computer, that seem to represent a place or a situation that a person can take part in." This definition encapsulates the essence of VR as a simulated environment that immerses individuals in a digital world, blurring the lines between physical and virtual experiences.

VR goes beyond the mere creation of a computer-generated world; it embodies a dynamic interaction between users and the virtual environment. This immersive experience involves two-way communication, allowing users to engage in sensori-motor actions that influence their perception of the space. Through this interaction, users can manipulate objects, navigate landscapes, and respond to stimuli, explore and interact with fantastical and imaginative realms [31]. Ultimately, VR blends sensory input and motor output, providing a rich and interactive experience that enhances user involvement.

Head-Mounted Displays (HMDs) have played a significant role in the various apparatus used to achieve immersion, which was used in this thesis. HMDs are devices worn on the head that

deliver visual and auditory stimuli directly to the user's senses. These devices encompass the user's field of view and contribute to the sense of presence (the sense of being there, see Section 1.1.1) within the virtual environment. Notably, Ivan Sutherland et al. [229] are credited with creating one of the earliest HMD prototypes in 1968, laying the foundation for subsequent advancements in this field. This immersive medium holds tremendous potential for applications in diverse fields, including entertainment, education, training, healthcare, and beyond, inspiring new possibilities and applications.

However, alternative hardware options, such as the Cave Automatic Virtual Environment (CAVE), have also emerged [48]. CAVE is a specially designed immersive room with multiple screens onto which the virtual environment is projected, enveloping users in a visually immersive experience [47]. Yin et al. [255] highlight the use of cutting-edge technologies in creating and delivering VR training experiences. This ongoing development signifies the commitment to pushing the boundaries of VR and providing users with increasingly immersive and realistic virtual environments.

These advancements in hardware technology have expanded the range of possibilities for experiencing VR. Whether through HMDs or alternative setups like CAVE, users now have the flexibility to choose the hardware configuration that best suits their needs and preferences, empowering them to tailor their VR experience.

1.1.1 Immersion

Immersion is a fundamental aspect of VR systems. It describes the extent to which VR technology immerses users, enabling them to experience a heightened sense of presence.

This immersion facilitates a heightened sense of presence, allowing users to experience a strong physical connection to the digital space. Achieving a high level of immersion is crucial for creating compelling and realistic VR experiences. Immersive virtual environments have the potential to captivate users, transport them to new worlds, and provide a sense of agency (perception of control and authorship over one's actions, see Section 1.2.5.2) and engagement. Immersive experiences can evoke a "sense of presence," which means being detached from one's physical surroundings and genuinely inhabiting the virtual environment [132].

Sense of Presence is a psychological construct that reflects the feeling of truly inhabiting a virtual environment, leading to realistic reactions and experiences within that space.

Presence and immersion are distinct concepts in VR. Presence is a psychological construct that varies based on individuals' mental representations. It emerges as a response to immersion [212] and manifests in users' realistic reactions to the virtual world.

By continually advancing the technologies and techniques contributing to immersion, researchers and developers are pushing the boundaries of what is possible in VR, creating increasingly realistic, captivating, and immersive user experiences. Servotte et al. [201] emphasize the significance of creating highly immersive and engaging environments that induce a sense of presence in users. For example, advanced graphics rendering techniques such as real-time ray tracing have been employed to achieve more realistic and visually immersive environments [144]. Spatial audio techniques significantly enhance emotional response and immersion in virtual environments, as validated by novel biometric sensor measures in the study [245]. Moreover, Haptic feedback systems like haptic gloves enhance user immersion by providing tactile sensations, which are crucial for teleoperation tasks and realistic haptic perception in telepresence scenarios [170].

Today, the potential applications of VR and immersion are remarkably diverse and have a profound impact across numerous fields like training [251] [154], healthcare [42], communication rehabilitation [33] [57], and even tourism marketing [211] [100], and so on. These examples illustrate the broad spectrum of applications of VR and immersion, showcasing their potential to transform industries and enhance user experiences. Continued research and development in these areas are essential to harnessing the full capabilities of VR and expanding its applications further.

1.1.2 Engaging with Virtual Environments: Interactions and Experiences

Interacting with virtual environments involves a dynamic action-perception loop, where users perform actions and receive feedback from the environment. As highlighted by Argelaguet and Andujar [11], the range of actions that can be performed in the virtual world encompasses various real-world activities such as walking, grabbing objects, and more.

In virtual environment interaction, users provide input through different modalities, e.g., motion or voice. This input is transmitted to the VR application via input devices. The environment, a key

player in this process, processes the input and generates feedback using output devices to provide a coherent response to the user's actions, thereby enhancing the user experience.

One approach to feedback generation is through one-to-one mapping or isomorphic interaction, where the user's motion is directly mapped to a corresponding motion performed by their virtual representation [87]. However, another form of interaction in virtual environments involves non-isomorphic or anisomorphic interaction [58]. Anisomorphic interaction applies transformations to the user's input, resulting in distorted feedback.

Feedback is critical in virtual environment interaction, providing user information through various sensory channels, primarily visual, auditory, and haptic feedback [250]. Visual feedback confirms the consequences of users' actions or informs them about the current state of the application. Auditory feedback, a key contributor to the overall user experience, can provide additional cues or alerts, enhancing the overall user experience. Haptic feedback, which simulates tactile sensations, enables users to feel virtual objects' texture, weight, or resistance, improving the sense of presence and realism.

VR offers a diverse range of interaction possibilities within virtual environments. These interactions can be classified into different categories [30]. The taxonomy includes navigation, selection, manipulation, and system input. Each category represents a distinct aspect of user engagement with the virtual world.

- Navigation in VR involves users' movements translating into camera translations and rotations within the virtual environment. Users can explore the virtual space by walking, moving their heads, or using handheld controllers to navigate the environment. Various techniques, such as teleportation or redirected walking, have been explored to enhance user locomotion and minimize discomfort.

- Selection and manipulation, fundamental interactions in VR, enable users to choose objects and perform transformations on them. Selection allows users to identify and interact with specific virtual entities, such as picking up an object or activating a control. Manipulation involves applying transformations to objects, such as translating, rotating, or scaling them. These interactions are typically facilitated through handheld controllers or hand-tracking technologies, which give users a sense of agency and control over the virtual elements. Furthermore, system input encompasses interactions that modify parameters within the VR application, access menus,

or input text. These interactions enable users to customize their experience, adjust settings, or provide input for text-based interactions, such as typing or entering commands. While system input is integral to VR applications, it often shares similarities with selection and manipulation techniques.

Researchers and developers can create more immersive and user-friendly VR experiences by understanding and advancing these interaction paradigms. Enhancements in tracking technologies, input devices, and interaction design continue to push the boundaries of what is possible within virtual environments, offering users increasingly intuitive and natural ways to interact with the virtual world.

1.2 Avatars in Virtual Reality

Avatar is a virtual representation of a user in a virtual environment, serving as an intermediary between the user's physical body and the digital space. It enables interaction and manipulation of the virtual world and facilitates feedback through various sensory modalities.

With the advent of digital platforms and the internet, people began seeking ways to express themselves to others in online environments. As text-based communication proved to be limited, there arose a need for visual representations to convey one's identity to fellow users. This led to the emergence of profile pictures, commonly referred to as avatars, initially popularized in video games and novels. Over time, the term 'avatar' has been expanded to encompass all forms of user representation in virtual environments. From video games and chat platforms to VR, avatars serve as users' virtual identities [198]. They can be customized and act as users' virtual proxies in the digital world. The definition of 'avatar' may vary across communities and researchers. Nowak and Fox [163] provide a comprehensive definition, stating that an avatar is a digital representation of a human user that facilitates interaction with other users, entities, or the environment. Additionally, Miao et al. [146] define avatars as entities capable of embodying artificial agents within immersive VR settings.

Avatars have rapidly developed in the VR domain, finding diverse applications that span from facial avatar reconstruction [256] to interventions addressing body image [174]. One crucial area where avatars are extensively used is in collaborating systems. The role of avatars in facilitating communication and interaction within virtual environments is vital for virtual collaboration [20]. Avatars enable users to convey themselves through gestures and body language, which are fundamental for social in-

teractions [189] [191]. Additionally, they visually represent other users' activities, enhancing the sense of presence and awareness within the virtual environment. Avatars also impact users' capabilities and offer information about their position in the virtual world. The appearance of an avatar reflects how users want to present themselves to others and their self-perception [240]. Users' interactions and emotional responses with avatars in virtual environments have been studied extensively. The studies and research efforts by Majaranta and Bulling [137], Zhang et al. [261], and Waltemate et al. [242] collectively contribute to the expanding body of knowledge concerning the relationship between users and their avatars in virtual environments. These studies explore visual attention, emotional expression, customization, and immersion, all aimed at deepening our understanding of this significant aspect of virtual interaction.

The study of the user-avatar relationship is an interdisciplinary exploration that sheds light on the intricate dynamics and profound influence that avatars have on shaping users' experiences and interactions within virtual environments. This exploration encompasses various dimensions, including embodied interaction and body ownership, thereby contributing to avatars' multifaceted definition and functionality in VR. The complexity of this interdisciplinary study is genuinely intriguing.

1.2.1 Avatar Interactivity: Immersing in Virtual Environments

In real life, our bodies serve as the means for performing actions through our motor system and receiving feedback through our sensory system. However, the avatar is crucial in VR as the intermediary between our physical bodies and the virtual environment. Avatar interactivity becomes a fundamental aspect of engaging with virtual environments, enabling users to interact and manipulate the virtual world actively through their avatars. The avatar serves as a representation of the actions performed by users. It facilitates feedback, encompassing visual cues and other sensory modalities, mirroring the feedback experienced in the real world.

Depending on the capabilities of the VR system, users can receive auditory, haptic, and proprioceptive feedback through their avatars. Auditory feedback can enhance the sense of presence and spatial awareness, providing users with audio cues that match their actions or the virtual environment's characteristics [241]. Haptic feedback, such as vibrations or force feedback, can simulate tactile sensations, allowing users to feel virtual objects or surfaces [96]. Proprioceptive feedback, which encompasses the sense of body position, movement, and muscle force, is crucial in shaping our perception of the body

and its interactions with the environment [177]. It can be conveyed through the avatar's movements and animations. The integration of proprioceptive feedback into VR experiences has the potential to provide users with a heightened sense of presence and embodiment within virtual environments.

The level of interactivity between users and avatars has been shown to significantly impact the overall immersion, presence, and user experience in VR settings. Research has demonstrated that increased avatar interactivity increases feelings of agency and ownership, enhancing the sense of embodiment and presence within the virtual environment [213].

One key aspect of avatar interactivity is the ability for users to manipulate virtual objects and interact with the environment naturally and intuitively. This can be achieved through various input modalities, including hand gestures, motion controllers, or haptic feedback devices.

Furthermore, avatar interactivity can extend beyond basic object manipulation, including social interactions with other avatars or virtual characters. The ability to communicate, gesture, or collaborate with virtual entities can significantly enhance the sense of social presence and create a more realistic and engaging experience. Studies have shown that when users interact with responsive and lifelike virtual characters, their social presence is heightened, increasing engagement and enjoyment of the virtual environment [253].

In addition to the immediate benefits of engagement and presence, avatar interactivity also holds promise for applications in various fields such as education, training, and therapy. Avatar interactivity can facilitate experiential learning, skill acquisition, and rehabilitation by allowing users to participate and manipulate virtual scenarios actively. For example, performing virtual procedures through interactive avatars in medical training simulations can provide a safe and realistic learning environment, promoting skill development and knowledge transfer [106].

In conclusion, avatar interactivity is crucial in engaging with virtual environments. By actively enabling users to interact with the virtual world and other virtual entities, avatar interactivity enhances the sense of agency, ownership, embodiment, and presence. This, in turn, leads to more immersive and captivating virtual experiences with numerous potential applications in fields such as education, training, and therapy.

1.2.2 Appearance: Humanoid or Non-Humanoid Avatar

As previously discussed, one aspect of avatars in VR is their appearance. However, the question arises: Is it necessary for an avatar to resemble a human?

Traditionally, humanoid avatars have been the prevailing choice in VR experiences, aiming to mimic the appearance and movements of real humans. This approach has been intended to create a sense of familiarity and empathy, as humans naturally tend to relate more easily to other human-like entities. Moreover, such an appearance enhances the sense of presence and social interaction in specific applications. Research has explored various aspects of humanoid avatars, including their use in digital twins, telepresence platforms, and VR communication systems. For instance, an inflatable cybernetic avatar (CA) with a humanoid upper body has been proposed as a lightweight, safe, and cost-effective option for physical human-robot interaction, allowing real-time collection of human behavior [160]. Additionally, studies have upgraded miniature biped robots into humanoid telepresence platforms to enhance social interactions, demonstrating successful operator control and high telepresence experiences [49]. Furthermore, advancements in VR communication systems aim to realistically render users' movements, facial expressions, and speech onto virtual avatars in real time for immersive experiences in virtual environments.

However, with the advancement of VR technology and the growing understanding of users' diverse needs and preferences, the concept of non-humanoid avatars has gained traction. These avatars can provide distinct advantages and unique opportunities for self-expression and immersion in virtual environments.

By adopting non-humanoid avatars, users can surpass the limitations of human form and explore a vast range of creative possibilities. They can assume animal, fantastical, or abstract forms, enabling them to embody different perspectives, engage in imaginative experiences, and enhance control precision [99] [168]. Such avatars encourage self-expression and significantly enhance immersion by offering a fresh and captivating way to interact with the virtual world, keeping users engaged and captivated.

Moreover, non-humanoid avatars can align with the context and theme of specific VR applications. For example, in a game set in a futuristic sci-fi universe, players might choose robotic or alien avatars that seamlessly blend with the environment, enhancing the overall sense of believability and enjoyment. Non-humanoid avatars, such as a scorpion, rhino, or bird, offer distinct gameplay experiences with

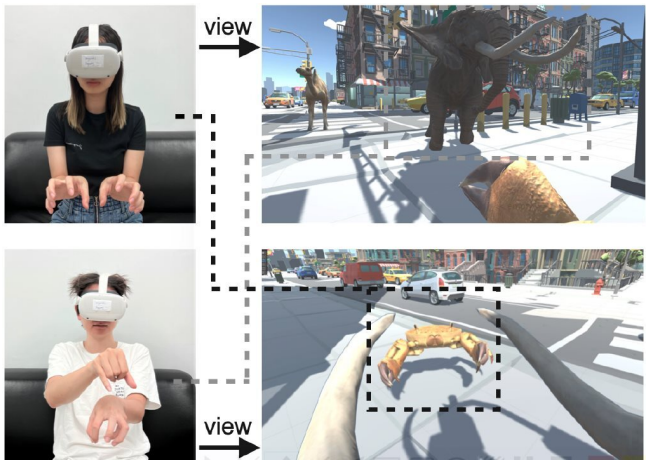


Figure 1.5: Users immerse themselves in a virtual world as a non-humanoid avatar, assuming the roles of a crab and an elephant, engaging in social interactions from a first-person perspective [99]

extraordinary abilities, creating a sense of virtual body ownership that enhances player enjoyment [119]. A car in a driving simulator can serve as a non-human avatar, expanding beyond humanoid forms to align with the specific context and purpose of the VR application [181]. Additionally, non-humanoid avatars can contribute to accessibility and inclusivity in VR. They provide alternative options for individuals with physical or cognitive limitations better suited to non-human representations. By embracing diverse avatar designs, VR experiences can become more welcoming and accommodating for a broader range of users.

It is important to note that the choice between humanoid and non-humanoid avatars ultimately depends on the goals, context, and preferences of the VR experience. While humanoid avatars can facilitate relatability and social interaction, non-humanoid avatars offer unique opportunities for creative expression, immersion, and pushing the boundaries of virtual experiences.

1.2.3 The Bidirectional Influence of Users and Avatars: Unraveling the Proteus Effect

The search results provide insights into the bidirectional relationship between users and their avatars in virtual environments, meaning they have an ongoing interaction and influence. This bidirectional relationship plays a significant role in the manifestation of the Proteus effect, which refers to the phenomenon where a person's behavior in a virtual environment is influenced by the characteristics of their avatar [176].

On one hand, the user controls the avatar's actions, movements, and decisions within the virtual

environment. The user's inputs and intentions are translated into the behaviors and appearances of the avatar. The user's characteristics and preferences can influence the design and customization of their avatars. The diary study found that people with disabilities preferred using avatars with disability signifiers, allowing them to better represent their identity and experiences in the virtual space [260]. Users wanted to enhance the aesthetics and interactivity of these avatars further to reflect their real-world selves better.

On the other hand, research has shown that the appearance and capabilities of an avatar can shape the user's self-perception and subsequent behavior. As the user embodies and identifies with the avatar, they can experience a psychological shift in their behavior, attitudes, and self-perception. This is where the Proteus effect comes into play. The avatar's characteristics, traits, and behaviors can positively impact how users perceive themselves and how they behave in the virtual environment, offering a hopeful and optimistic outlook. For example, suppose users embody an avatar that is tall, attractive or possesses powerful traits. In that case, they might feel more confident, assertive, and inclined to take on leadership roles within the virtual world. Conversely, suppose users embody an avatar that is small, unattractive or exhibits traits associated with negative stereotypes. In that case, they might experience lowered self-esteem and submissiveness or conform to social expectations associated with those traits.

Understanding this bidirectional relationship and the potential for the Proteus effect is crucial for designing immersive and transformative VR experiences. It enlightens developers and researchers about how avatars can enhance users' engagement, self-expression, and psychological well-being within virtual environments, keeping them informed and up-to-date.

1.2.4 Functional Capabilities of Avatars in Virtual Environments

Avatars in virtual environments possess various functional capabilities that enhance the user experience and facilitate interaction within the virtual world. These capabilities include visual representation, animations, locomotion, object manipulation, and communication abilities. Animations are crucial in bringing avatars to life, enabling them to move, gesture, and express emotions, enhancing realism, and facilitating non-verbal communication [55]. Akbas et al. [2] designed and prototyped a VR dance experience called Virtual Dance Mirror. A dancer's bodily movements are reflected on a 3D avatar model using a motion-capture suit in this system. This approach explored the functionality

of avatars in VR by embodying dancers' movements on 3D models, enabling enhanced collaboration, and facilitating non-verbal communication in immersive VR experiences. Such innovations not only showcase the potential of avatars to extend physical capabilities but also inspire the creation of new forms of expression and interaction within virtual environments.

The locomotion capabilities of avatars allow users to navigate and explore the virtual environment. Avatars can be controlled through various means, such as joystick-based movement, gesture-based locomotion, or brain-computer interfaces [1]. Object manipulation is another crucial functional capability of avatars, enabling users to interact with virtual objects. Avatars can be equipped with tools or virtual hands that allow users to grasp, manipulate, and interact with virtual objects naturally and intuitively [12]. Knierim et al. [115] analyzed the effects of avatar hands and physical keyboards on typing performance in VR. The findings highlighted that the design and functionality of avatars have a direct impact on user interaction. This suggests further research into optimizing user input methods for virtual environments. Exploring different ways to enhance user input and interaction with avatars can lead to more intuitive and efficient experiences in VR.

Furthermore, avatars possess communication abilities that enable users to interact with other avatars or virtual characters within the virtual environment. This includes voice, text, and non-verbal communication through gestures or facial expressions [78]. These communication capabilities enhance social interaction and collaboration among users in virtual environments, fostering a sense of presence and co-presence [69]. A study by Lo [131] examined the impact of chatbots and conversational agents on educational settings. Although not explicitly focused on avatars, the study provided insights into the potential role of interactive virtual entities in enhancing educational experiences. This suggests that future research could explore integrating avatars with chatbots to create more engaging and immersive educational environments in VR. Roth et al. [192] examined the relationship between avatar realism and social interaction quality in VR. The study revealed that realistic avatars positively influenced social interaction quality, emphasizing the potential of avatars to facilitate meaningful social experiences in virtual environments. This finding suggests that future research should explore the impact of avatar realism on social presence and interpersonal communication in VR.

Moreover, in social communication rehabilitation for autistic children, avatars offer a valuable means of interaction by leveraging visual, speaking, and aural cues. Through these cues, avatars facilitate engagement and help reduce repetitive behaviors [200]. This application underscores the

effectiveness of avatars as tools in therapeutic interventions, supporting the development of social skills and communication in a controlled and nurturing virtual environment.

In conclusion, avatars in virtual environments exhibit a range of functional capabilities that contribute to an immersive and interactive user experience. Their abilities in visual representation, animations, locomotion, object manipulation, and communication facilitate engagement, presence, and social interaction within virtual environments.

1.2.5 Embodiment in Avatars

As discussed in section 1.2, avatars can represent users in VR. Through this representation, users can experience a profound sense of embodiment, perceiving the avatar as their physical body within the virtual environment [109]. This definition decomposed the sense of embodiment into three subcomponents:

- Sense of Body-Ownership (SoBO)
- Sense of Agency (SoA)
- Sense of Self-Location (SoSL)

In the subsequent sections, these components will be explored individually to understand their role in embodiment better.

1.2.5.1 Sense of Body Ownership (SoBO)

One of the essential aspects of avatars is body ownership. Body ownership in avatars refers to the psychological phenomenon where users perceive and experience the virtual body of an avatar as their own [28]. It entails a subjective feeling of ownership and identification with the avatar's body as if it were an extension of the user's physical self. Body ownership is primarily associated with the appearance aspect of avatars, as the visual representation of the avatar's body plays a crucial role in creating the illusion of embodiment and ownership. The appearance aspect of avatars focuses on the visual representation of the avatar, including attributes such as body shape, facial features, clothing, and accessories. The design of the avatar's appearance significantly influences users' perception, identification, and emotional connection with the virtual character.

When users feel a strong sense of body ownership in an avatar, they feel like they are inhabiting and controlling the avatar's body [37]. This aspect of avatars is crucial, but it is also the very essence of creating immersive and engaging VR experiences. It enhances the feeling of being "present" in the virtual world and can contribute to a more intuitive and natural interaction with the environment and other virtual entities.

Researchers and developers employ various techniques to promote body ownership in avatars, such as realistic body representations, real-time motion tracking, sensory feedback such as seeing the avatar's movements synchronized with their actions, haptic feedback that corresponds to the interactions in the virtual environment, and alignment between the user's actions and the avatar's movements. These techniques aim to bridge the gap between the user's physical body and the virtual representation, fostering a strong sense of ownership and embodiment in the avatar.

The findings of Ogawa et al. [166] indicated that enhancing body ownership in avatars can be achieved by utilizing realistic body representations, as realism influences object size perception in body-based scaling, fostering a stronger sense of embodiment. Lugrinet al. [134] conducted a novel experiment, a first of its kind, to explore the impact of avatar realism on the illusion of virtual body ownership (IVBO) in immersive virtual environments, with full-body avatar embodiment and freedom of movement, revealed that each avatar elicited a relatively high level of illusion. Kim et al. [112] suggested the beneficial effects of synchronized body motion and matched body size between a user and avatar on the user's subjective experience in IVR, which can boost the effects of VR applications in entertainment, psychotherapy, and education. Van der Waal et al. [239] found that embodying an overweight avatar in VR led to negative affective responses, indicating a link between avatar appearance and body ownership.

Body ownership in avatars is closely related to the distinction between first-person and third-person perspectives.

1.2.5.1.1 First-Person Avatar First-person avatars embody the user's point of view within the virtual environment, providing a perspective that matches the user's visual and auditory senses, as is shown in Figure 1.6. When using a first-person avatar, users experience the virtual world as if directly interacting with it. They see what the avatar sees and hear what the avatar hears, creating a sense of immersion and presence. One of the critical goals of first-person avatars is to establish a strong

sense of body ownership. Users perceive the virtual body as their own, feeling a sense of embodiment and control over the avatar's movements and actions. This embodiment enhances the feeling of being "present" in the VR experience as users interact with the environment through the perspective of their avatar.

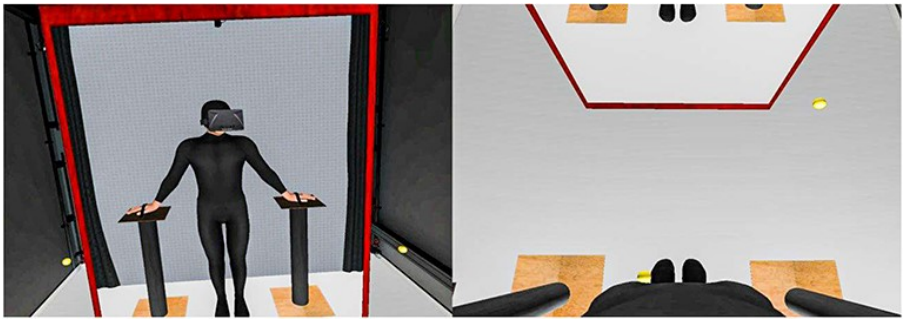


Figure 1.6: On the left, a first-person perspective of an avatar with a front view, as seen through a mirror reflection. On the right, a view of the participant's virtual body as they look down [159]

.

Waltemate et al. [242] emphasized that users experience the virtual world as if they are directly interacting with it when using first-person avatars. This highlights the immersive nature of first-person avatars, providing a perspective that matches the user's visual and auditory senses. Furthermore, Gorisse et al. [80] found that first-person perspectives in immersive virtual environments significantly contribute to embodied users' presence and performance analysis, indicating the importance of first-person avatars in creating a solid sense of presence. Argelaguet et al. [12] highlighted the impact of interaction and the role of virtual hand representation. Their study demonstrated the significance of interaction in virtual embodiment and its effects on virtual hand representation, shedding light on the importance of interaction design in enhancing the user experience with first-person avatars.

Pavone et al. [171] further explored Embodiment in VR and examined the electrocortical signatures of monitoring errors in an avatar's actions seen from a first-person perspective. This study provided insights into the neural correlates of embodying others in immersive VR, contributing to understanding the cognitive processes underlying embodiment in VR.

However, there are knowledge gaps regarding the impact of specific learning styles on the user's experience with first-person avatars in VR. Gilakjani [77] highlighted the importance of visual, auditory, and kinaesthetic learning styles in education, suggesting potential future research directions in investigating the influence of learning styles on the perception and interaction with first-person avatars

in VR.

1.2.5.1.2 Third-Person Avatar Third-person avatars are external representations of the user within the virtual environment. Users view the third-person avatar from an external viewpoint (see Figure 1.7), often observing the avatar and its actions within the virtual world. While users have a degree of control over the avatar's behavior, the sense of body ownership may be less pronounced than in first-person avatars [54]. Users perceive the third-person avatar as a separate entity rather than an extension of themselves. This external perspective gives users a broader view of the virtual world, enabling them to observe the avatar's movements, actions, and interactions with the environment.

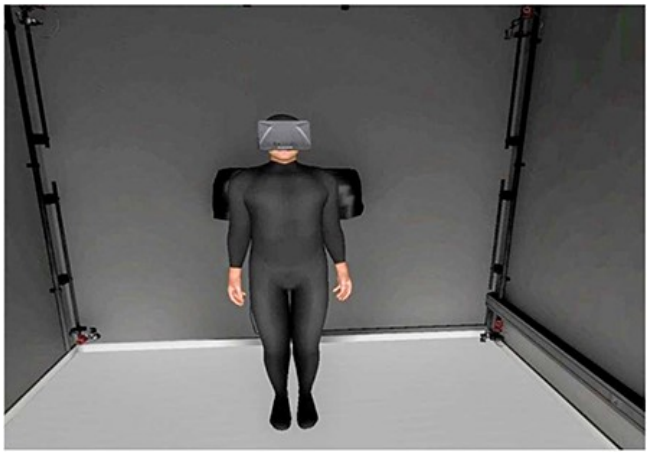


Figure 1.7: A third-person perspective of the avatar, as seen from the participant's point of view [159]
.

Slater et al. [214] explored the first-person experience of body transfer in VR. Their study focused on the sense of body ownership and found that users perceive first-person avatars as an extension of themselves. This finding suggests that the sense of body ownership may be more pronounced in first-person avatars than in third-person avatars. However, the study did not directly investigate third-person avatars, indicating a knowledge gap in understanding the differences in body ownership between first-person and third-person perspectives.

Maselli and Slater [142] delved into the building blocks of the full-body ownership illusion in VR. While their study primarily focused on the mechanisms underlying body ownership illusions, the findings can be extrapolated to understand how users perceive and relate to third-person avatars. The study highlighted multi-sensory integration and agency as key components of the full-body ownership

illusion. These findings provide valuable insights into the cognitive and perceptual processes that may influence the user's perception of third-person avatars.

1.2.5.2 Sense of Agency (SoA)

The sense of agency, which refers to the perception of control and authorship over one's actions, is a crucial aspect of embodiment. While extensively studied in domains outside of VR [169], it plays a significant role in understanding the user-avatar relationship within virtual environments. A strong sense of agency involves recognizing oneself as the cause of action and experiencing coherent feedback in response to those actions. This dual aspect of agency, namely the feeling and judgment of agency, has been distinguished by Jeunet et al. [98], drawing upon previous psychological studies.

By delving into the sense of agency within virtual embodiment, researchers aim to uncover how users perceive and interpret their actions and the subsequent feedback they receive within virtual environments [116] [70]. This exploration considers factors such as the user's ability to control and manipulate their avatar, the alignment between their intentions and the avatar's actions, and the degree to which the feedback from the virtual environment corresponds to their actions. Understanding the sense of agency in VR is pertinent to enhancing user experiences and has implications for designing interactive and immersive virtual environments.

Furthermore, investigating the sense of agency in VR can shed light on the mechanisms underlying the presence and identification of avatars [124]. When users experience a strong sense of agency, they are likelier to feel a sense of ownership and embodiment of their avatar, leading to a heightened sense of presence and engagement within the virtual environment. This interplay between the sense of agency, embodiment, and presence contributes to a more immersive and captivating virtual experience.

1.2.5.3 Sense of Self-Location (SoSL)

The sense of self-location, self-identification, or spatial presence relates to the user's perception of their location within the virtual environment. Research has shown that the level of body representation, such as full-body avatars, significantly impacts spatial awareness, self-presence, and spatial presence during virtual locomotion [219] [264]. Additionally, studies have demonstrated that the brain can perceive an avatar's body as its own in VR, highlighting the importance of motor predictions in establishing self-location and self-identification within virtual environments [85]. Furthermore, the

perspective from which the avatar is observed plays a significant role in influencing this sense. First-person avatars give users a stronger sense of self-location as they perceive themselves within the virtual world through the avatar's perspective. In contrast, when using third-person avatars, this sense of self-location tends to be weaker [27]. This sense of self-location contributes to a more immersive and embodied experience.

The concept of self-location finds its roots in real-life experiments like the rubber hand illusion (RHI), where the perceived location of a fake hand differs from the actual location of the real hand [61]. In VR, achieving a co-located avatar becomes easier when tracking is accurate. As a result, this component has garnered less attention in the VR community in recent times.

While the sense of self-location has received less emphasis in VR research due to advancements in tracking technology, its significance is recognized in scenarios where deliberate manipulation or disruption of the user's perception is desired. These applications provide opportunities to explore the boundaries of self-location perception in virtual environments and offer unique experiences that challenge the traditional sense of embodiment.

1.3 "Trustworthiness" in Virtual Reality

"Trustworthiness" in VR refers to the degree to which users perceive VR systems, applications, or agents as reliable, credible, and secure. Adopting and effectively using VR technologies is crucial, especially in applications involving sensitive information or critical tasks. Trustworthiness encompasses several dimensions: transparency, privacy, ethical use, user training, and overall user experience.

Research has shown that in virtual medical interactions, patients' trust in their virtual doctor can impact pain perception [9]. Studies on animated virtual characters have demonstrated that trustworthiness can be experimentally manipulated and graded, affecting social traits and user perceptions [204]. In vehicular clouds (VCs), trustworthiness is essential for ensuring data privacy, integrity, and resource availability, highlighting the significance of security features in virtual environments [3]. Trustworthiness is critical in virtual communities like Second Life, influencing behaviors and interactions within these digital spaces [71]. Furthermore, a study utilized a multi-sensory approach involving electroencephalography (EEG), galvanic skin response, and Heart Rate Variability to measure trust in a virtual agent and investigate the relationship between trust and cognitive load [84]. In this study,

trustworthiness in VR was defined as the perceived reliability of a virtual agent. The findings revealed that EEG alpha band power was an indicator of trustworthiness, suggesting a correlation between biosignals and trust in VR interactions.

Generally, several factors contribute to trustworthiness in various contexts, such as online platforms or human interactions.

- **Authenticity:** In avatars, users trust avatars that exhibit authentic and genuine behaviors. This includes avatars accurately representing the intended character or entity and behaving consistently and coherently. Avatars perceived as authentic are more likely to be trusted by users. Mikeska et al. [147] delved into the perceptions of avatar authenticity in virtual environments, particularly among pre-service teachers. The study examined how variations in authenticity perceptions among pre-service teachers engaging in simulated teaching experiences can impact their trust in the avatars. This research provides valuable insights into the role of authenticity in establishing trustworthiness in avatars. It underscores that users are more likely to trust avatars perceived as authentic and genuine, thereby highlighting the significant influence of authenticity on user trust. Huang and Jung [93] proposed a multi-layered perceived authenticity model, which delves into the perception of authenticity in virtual characters. The model demonstrates that virtual characters do not necessarily have to be perceived as humans to be considered authentic by their human interactants. This implies that avatars can possess their unique characteristics and behaviors while still being perceived as authentic, reinforcing the role of authenticity in building trustworthiness. Moradinezhad and Solovey [153] investigated how trust is influenced by the cooperativeness of an Embodied Virtual Agent (EVA) and an individual's prior experience with other agents. They found that participants may still choose the agent's suggested answer (which can be incorrect) over theirs, even if they are sure their answer is correct.

- **Reliability:** Trustworthy avatars are reliable and consistent in their performance. They consistently respond to user input, provide accurate information, and fulfill their intended functions without unexpected errors or inconsistencies. Reliability builds trust in the avatar's abilities and enhances the user experience. Visser et al. [53] investigated the impact of anthropomorphism on trust resilience in cognitive agents and found that trustworthy avatars are reliable and consistent in their performance. [46] explored the impact of self-representation (full body Self

Avatar vs. Just Controllers) in a collaborative virtual environment (CVE) and the consistency of self-representation among the users. They concluded that consistent self-representation in CVE leads to increased trust and productivity, highlighting the importance of reliability in avatar performance.

- **Transparency:** Transparency involves providing users with clear information about how the VR system operates, what data it collects, and how it is used. This openness helps build user trust by ensuring no hidden practices exist. Trustworthy avatars demonstrate a transparent approach, offering clear and understandable explanations of their behaviors, actions, and decision-making processes. This transparency is not just a feature but a necessity in helping users comprehend and anticipate the avatar's responses, fostering trust and minimizing uncertainty. Zaman et al. [258] discussed the importance of transparency in enhancing trust in XR technology in education and training. Marino et al. [140] conducted interviews with users of Second Life® to delve into avatar transparency in virtual worlds and its connection to establishing trustworthiness. Their preliminary analysis provides valuable insights into the importance of transparency in avatars and its impact on building trust. Understanding the relationship between transparency and trustworthiness aids in developing avatars that users perceive as reliable, credible, and trustworthy. By emphasizing transparency, avatars can enhance the quality of their interactions and promote a sense of openness, clarity, and predictability. Users appreciate avatars that are forthcoming about their actions and intentions, as it creates an environment of trust and reduces uncertainty. Incorporating transparency as a fundamental characteristic of avatars contributes to building trust and fostering positive user-avatar relationships. Narag et al. [158] suggested using the Bayesian Theory of Mind approach to infer user avatar intentions can contribute to transparency in shared virtual environments. This transparency, in turn, can enhance users' trust in the avatar, as they have a clearer understanding of its motivations and actions.

- **Privacy:** Privacy concerns protecting user data from unauthorized access and misuse. Ensuring robust privacy measures is essential for building trust, especially in applications that handle personal or sensitive information. Zaman et al. [258] highlight privacy as a critical consideration for establishing trust in XR technology.

- **Ethical Use:** Ethical use involves adhering to moral principles and standards in designing and

deploying VR systems. This includes avoiding manipulative practices and ensuring that the technology is used for beneficial purposes. [258] also includes ethical use as a component of its comprehensive framework for enhancing trust in XR technology.

- **User Training:** Providing adequate training and support to users helps them understand and effectively use VR systems, increasing their confidence and trust in the technology. [258] mentions user training as a crucial aspect of building trust in XR technology.

- **User Experience:** A positive user experience characterized by intuitive interfaces, responsive interactions, and realistic simulations enhances the perceived trustworthiness of VR systems. Yildirim [254] discusses how the design of conversational agents in VR, including their embodiment, expressiveness, and responsiveness, influences their perceived trustworthiness.

1.4 Intelligent Avatar

The term' Intelligent Avatar (IA)' refers to an avatar that is equipped with advanced artificial intelligence (AI) capabilities [72]. These capabilities enable the avatar to exhibit intelligent behavior, make decisions, and interact with users sophisticatedly [128]. To enhance its intelligence and responsiveness, an intelligent avatar may incorporate various AI techniques, such as natural language processing, machine learning, computer vision, and other cognitive abilities. Intelligent avatars play a crucial role in the intersection of AI and VR, enhancing user experiences and interaction in virtual environments.

Rule-based systems (production systems or expert systems) are the simplest forms of AI that rely on predefined rules and conditions to facilitate adaptation in VR [82]. These techniques allow for static adjustments based on specific triggers or user inputs. Using rule-based adaptation frameworks like the 3QS-adaptation framework, virtual environments can dynamically adjust to users' needs and mitigate issues like cybersickness [237]. These systems monitor performance and security anomalies in real-time, triggering adaptations to enhance user experience while maintaining application functionality. Additionally, VR technology generates virtual gaming environments that adapt based on real-world features identified by head-mounted display devices, enabling customized and immersive gaming experiences [155].

In a more advanced mode, other AI techniques bring higher adaptability to VR avatars. These

techniques encompass a range of methods, including optimization algorithms, genetic algorithms, machine learning, deep learning, and reinforcement learning. Optimization algorithms present a powerful tool for enhancing adaptive avatars and user experience in VR. One effective approach involves iteratively adjusting avatar characteristics based on user performance [262]. Moreover, optimization-based reset algorithms can significantly improve user interaction in VR applications like Redirected Walking (RDW) by reducing the resets needed [130]. These advancements showcase the potential of optimization algorithms in tailoring avatars and VR experiences to users' needs and preferences, ultimately enhancing immersion and interaction in virtual environments. Moreover, genetic algorithms [4] have demonstrated their efficacy in enhancing the adaptation of non-player characters in games, yielding promising results [205].

Research suggests that machine learning-based predictions can improve input performance by compensating for system latency and predicting future avatars' movements.

Some of the research studies employed reinforcement learning (RL) techniques. Kastanis and Slater [102] explored using RL techniques to train virtual characters to interact with participants and guide them toward specific goals. The study emphasized the lifelike responses of individuals in immersive virtual environments and utilized these responses to accomplish desired tasks. Porssut et al. [175] conducted a study that focused on a particular method utilizing RL to determine the maximum distortion level that can go unnoticed by individuals. The objective was to prevent subjects from detecting distortions in the avatar's movements as long as the distortion remained below a certain threshold.

Machine learning-based predictions of full-body avatar movements in VR can compensate for system latency but only provide subjective benefits under specific conditions, as shown in the study by Schwind et al. [199]. Machine learning algorithms can play a pivotal role in VR therapy for phobias by customizing virtual environments according to users' emotional states. This customized approach allows for the creation of tailored exposure scenarios to facilitate effective treatment. The work of Van et al. [238] is a testament to the potential of machine learning in this field, offering hope for the future of mental health treatment.

Using machine learning techniques, VR sensorimotor rehabilitation can develop customized avatars capable of predicting motion within latency constraints. Bălan et al. [19] studied this approach, which has shown promising results in addressing sensorimotor deficits and enhancing patients' immersive

experiences. The work of Axenie et al. [16] further demonstrates the potential of machine learning in this field, offering optimism about the future of healthcare technology.

Guerrero Vásquez et al. [83] proposed animating 3D models to transform them into virtual agents or avatars, which could be used in Human-Computer interaction processes. They also propose developing algorithms combining artificial intelligence and human intelligence to improve interaction processes.

Chen et al. [41] proposed transferring facial motion from a regular video to a 3D photo-realistic avatar by learning a deep-learning lighting model. This method, in combination with a high-quality 3D face tracking algorithm, provides a method for subtle and robust facial motion transfer.

The study by Aquino et al. [51] underscores the use of avatars (or intelligent agents) in virtual environments to enhance the analysis of the environment, monitor user actions, and identify user profiles. By interpreting user intentions, agents contribute to constructing more precise and dynamic adaptive virtual environments.

Gongora et al. [79] described a novel methodology for creating efficient AI-based agents capable of adapting. They suggested a novel differentiated combination of these techniques to enable the self-adapting process for intelligent agents used to represent characters in virtual environments.

Zhang and Tsai [263] discussed the application of adaptive VR with AI techniques in sports training, showcasing the potential for intelligent adaptive avatars using machine learning in virtual environments. They designed a motion capture data algorithm based on behavior string, which successfully improves the advantages of VR technology.

Hassan et al. [88] described an interview training program that utilizes an artificially intelligent Child Avatar in VR for training Child Protective Services (CPS) and police personnel in interviewing abused children. The program uses the Unity game engine and incorporates various artificial intelligence-based technologies, including dialogue models, talking visual avatars, text-to-speech, and speech-to-text components. It showcases the application of intelligent avatars in VR for a specific training scenario, where the avatars simulate interactions with abused children during interviews. Integrating artificial intelligence technologies enhances the realism and effectiveness of the training program.

Zhang et al. [259] conducted research on human motion recognition in virtual environments dur-

ing multimedia interactions. They proposed an algorithm that utilizes linear decision and support vector machine (SVM) techniques for motion classification and recognition. The algorithm employs linear discriminant analysis in combination with a kernel function to project training samples into a high-dimensional space, enabling the extraction of an optimal classification feature vector. This approach effectively addresses nonlinearity and enhances sample differentiation. The paper also utilizes a genetic algorithm for parameter optimization of the SVM, taking advantage of its strengths in multi-dimensional space optimization. The accurate motion recognition provided by this algorithm can greatly benefit intelligent avatars, enabling them to respond realistically and appropriately to user movements.

Kruszewski and Mahamad. [122] explored the development of AI-powered avatars for immersive AR/VR experiences utilizing webcams and deep learning techniques. Their research enables users to engage in cyberspace interactions through diverse digital representations. The study presents a tangible implementation of avatars within a novel system, emphasizing the exciting possibilities of avatar-based experiences in virtual environments.

Rincon et al. [186] designed a framework called JACALIVE, especially for the execution and adaptation of Intelligent Virtual Environments. This framework facilitates the development of this environment, efficiently and realistically managing the evolution of parameters to adapt to the physical world. While the focus is more on the overall environment than avatars specifically, the framework's ability to manage and adjust parameters can have implications for the behavior and interaction of avatars within the virtual environment.

Wang et al. [244] proposed a method for real-time facial retargeting. The method utilizes deep learning and feature transfer to generate blend shapes for target avatars based on given sources. It employs a Variational Autoencoder (VAE) and a Multilayer Perceptron (MLP) model to achieve the translation between source and target avatars' latent spaces. The results show that the proposed method outperforms existing methods in generating realistic blend shapes.

1.5 Summary

In this chapter, we discussed the role of avatars in virtual environments. Avatars are visual representations people use to express their identities online. They serve as users' virtual identities

and can be customized to act as their virtual proxies. Avatars facilitate interaction with other users, entities, or virtual environments.

Avatar interactivity is crucial for engaging with virtual environments. The avatar intermediates the user's physical body and the virtual world. It allows users to interact and manipulate the virtual environment actively, providing feedback through visual cues and other sensory modalities such as auditory, haptic, and proprioceptive feedback. Increased avatar interactivity enhances immersion, presence, and user experience in VR.

Trustworthiness in VR refers to virtual entities or characters' perceived reliability, credibility, and integrity. Different factors contribute to trustworthiness in various contexts, such as avatars. However, it's important to note that current intelligent avatars in VR have limitations in terms of autonomy and adaptability. They often rely on predefined scripts or responses, making interactions feel scripted and less immersive. This underscores the need for avatars with greater autonomy and adaptability, capable of learning and dynamically adjusting their behavior based on user interactions and changing contexts.

Building on this foundation, the next chapter will outline the primary approach used to investigate the development of these adaptable intelligent avatars. Following that, we will delve deeper into how trustworthiness can be integrated into the design of intelligent avatars, as well as the associated limitations and challenges.

It is important to note that while the concept of embodiment in avatars is significant within the broader context of VR, this thesis primarily emphasizes the functionality of intelligent avatars and aims to develop a general framework that can potentially accommodate various avatar use cases. This focus enables a more targeted investigation into the effectiveness of intelligent avatars in real-world applications. Additionally, the complexities and nuances associated with embodiment, while noteworthy, introduce layers of variability that may divert attention from the core objectives of this research. Consequently, although embodiment is acknowledged in the literature review as a relevant concept, it will not be a primary focus in the subsequent chapters.

Chapter 2

Research Approach

Content

A substantial body of research exists on embodiment, as exemplified by the work of Kilteni et al. [109]. However, as a pivotal element of the VR experience, Avatar design falls short in several critical aspects. Some of the neglected points in avatar design include: Inadequate Representation of Users and Neglecting Their State:

1. **Inadequate Representation of Users and Neglecting Their State:** Avatars in VR frequently overlook the user's physiological state, such as heart rate, respiration, or emotional responses. This omission limits the potential for adaptive and customized experiences within virtual environments. Nacke et al. [156] highlight the significance of integrating physiological feedback into avatars, enabling real-time adaptation and enhancing user engagement. Moreover, avatars in VR often fail to depict behavioral nuances accurately. These limitations undermine the potential benefits of social presence and user immersion within virtual environments. Bailenson et al. [17] emphasize the importance of embodying avatars with realistic features, as it significantly influences user behavior, attitudes, and social interactions in virtual environments. Safety Concerns

2. **Safety Concerns:** Another critical aspect overlooked in avatar design for VR is user safety. Avatars that do not prioritize safety considerations can lead to adverse effects, such as cybersickness, discomfort, or increased risk of physical injury. Bowman et al. [31] emphasize incorporating safety factors into avatar design, including ergonomic considerations, collision avoidance mechanisms, and minimizing sensory conflicts. By addressing these safety concerns, developers can create a safer virtual environment for users.

By examining these deficiencies, we can gain deeper insights into the challenges and opportunities associated with avatar design, ultimately paving the way for more engaging, personalized, and safe virtual experiences.

2.1 Research Questions

Research in avatars has traditionally focused on appearance rather than considering the user state. In this thesis, we want to answer these questions:

Q_1: *Which features and behaviors should be integrated into virtual avatars?*

Q_2: How can a customized avatar be developed to help improve the experience in virtual immersion?

2.2 Proposed Approach

Our approach encompasses a comprehensive understanding of users, considering their psychological and behavioral states. We aim to build a highly customized model of the user's virtual companion by integrating objective parameters such as electrodermal activity and eye movement. One prevalent challenge addressed using our model is cybersickness, which can lead to discomfort and dissatisfaction for some users. By addressing cybersickness as a critical indicator, we aimed to enhance overall safety in virtual environment. In this way, our model will focus on mitigating safety concerns induced by cybersickness, such as user disorientation and anxiety, ensuring a more supportive and enjoyable virtual experience.

We selected and developed a driving simulator as a VR application use case because these applications can be excellent for developing intelligent avatars due to their immersive and realistic environments. These simulators closely mimic real-world driving experiences, allowing users to engage in scenarios that elicit genuine responses. This realism is crucial for understanding how users interact with the virtual environment and how various factors can contribute to or mitigate feelings of discomfort. By immersing users in a familiar context, we enabled gather valuable data on physiological and psychological responses to driving simulations, providing a solid foundation for the intelligent agent's development.

Another reason for this selection was this kind of application's dynamic and controlled environments. These simulators allow real-time interaction, enabling the avatar to monitor users' physiological and behavioral indicators continuously. This capability facilitates prompt, personalized interventions, such as adjusting driving conditions to enhance user comfort. This process enables the intelligent avatar to learn from interactions and refine its algorithms, ultimately improving its effectiveness in mitigating cybersickness.

To achieve this, we relied on artificial intelligence tools that enable predicting and detecting the user's state. By harnessing AI algorithms and machine learning techniques, we can analyze the objective and subjective data obtained from the user during navigation in a virtual environment in real time. This data-driven approach ensures that the model continuously adapts and evolves, maintaining

its accuracy and relevance.

Our efforts have resulted in a next-generation VR experience characterized by a heightened level of safety. Equipped with human-centered intelligence, users can confidently navigate virtual environments, knowing that their virtual companion is attuned to their needs and working to optimize their experience.

This user-centric and adaptive approach aims to create a virtual environment that is not only immersive but also customized and tailored to each user, ensuring their safety at all times. The ultimate goal of our approach is to generate virtual twins of users. A virtual twin is a digital replica that reflects an individual's unique characteristics and responses in the virtual environment. By creating these personalized avatars, we can enhance the user experience by tailoring interactions and interventions based on their specific needs and preferences.

Chapter 3

Trustworthy Intelligent Avatar

Content

Avatars play a significant role in virtual environments, where trustworthiness is a crucial quality that instills reliability and confidence in users. In VR, factors such as visual realism, natural behavior, social cues, effective communication, and adaptability contribute to avatars' perceived trustworthiness.

Adaptive avatar behavior enhances trust by adjusting their responses and interactions to personalize the user experience. Moreover, adaptive avatars have diverse applications, including emotional induction, trait prediction, addressing sensorimotor deficits, and representing remote VR users in mixed-reality collaboration. Establishing trust between users and intelligent adaptive agents (IAA) in VR involves interdisciplinary research exploring psychological, philosophical, sociological, computational, and economic aspects. Efforts have been made to develop adaptive avatars through iterative approaches and closed-loop systems, enabling auto-adaptation and enhancing trustworthiness and user engagement.

This chapter explores the mechanisms, technologies, and approaches that enable avatars to exhibit adaptive behavior and foster trust in VR by delving into the concept of intelligent avatars (IA). While often used interchangeably, it is crucial to differentiate between avatars and agents to establish a precise definition of intelligent avatars. Avatars are virtual representations controlled by humans, facilitating self-expression and interaction with the virtual environment and other users. Conversely, agents are virtual representations controlled by computer algorithms, exhibiting autonomous behavior and decision-making capabilities. Intelligent avatars combine elements of avatars and agents, possessing advanced artificial intelligence capabilities such as natural language processing, machine learning, and computer vision. They engage in real-time interactions, exhibit adaptive behavior, make autonomous decisions, demonstrate emotional intelligence, and engage in multi-modal interactions.

However, developing intelligent avatars presents challenges related to natural language understanding, ethical and privacy concerns, the uncanny valley effect, and scalability and performance. Overcoming these challenges can lead to more immersive and interactive virtual experiences. By harnessing the power of adaptability, intelligent avatars can become more reliable, responsive, and customized, ultimately enhancing the user experience in virtual environments.

3.1 Trustworthy Avatar

In the realm of avatars, such as in the broad definition within VR (as outlined in section 1.3), being "trustworthy" refers to the perception or quality of an avatar that instills a sense of reliability, credibility, and confidence in users. In other words, a trustworthy avatar is one that users perceive as dependable, truthful, and capable of fulfilling its intended role or function [136]. Several factors can influence trustworthiness in avatars.

In the first chapter (section 1.3), some general factors were introduced including authenticity, reliability, and transparency, contributing to trustworthiness in various contexts. However, trustworthiness in avatars within virtual environments can involve a broader range of factors. These factors may include visual realism, natural and responsive behavior, consistent and predictable actions, appropriate social cues, effective user communication, and adaptability. However, the specific factors and their relative importance may vary depending on the particular application, research objective, or domain of study.

3.2 Adaptive Avatar Behavior: Fostering Trust

The adaptability of avatars is a crucial element in enhancing trustworthiness, as it allows them to adjust their behavior, responses, and environment. Adaptive avatars significantly improve user experiences in VR through customized interactions and tailored environments [178][18]. These avatars have various applications, such as inducing emotional states, predicting individual traits, and addressing sensorimotor deficits, particularly in scenarios like sensorimotor rehabilitation [35][16]. Furthermore, the role of adaptive avatars in representing remote VR users in mixed reality (MR) collaboration is paramount, as it enhances social presence and overall MR experience [172].

In addition, adaptive interfaces in VR can potentially enhance physical exercise motivation and intensity. They achieve this by responding to users' exertion levels and physiological data, ultimately improving training and gaming performance [34]. By offering a versatile tool, adaptive avatars in VR enable the creation of immersive, customized, and engaging virtual experiences that cater to individual users' needs and preferences. It's important to note that users highly value avatars that can dynamically adjust and cater to their requirements [243], highlighting the user-centric nature of adaptive avatars.

Sun and Botev [227] provided an overview of the numerous factors in establishing trust between users and Intelligent adaptive agents (IAA), spanning scientific disciplines as diverse as psychology, philosophy, sociology, computer science, and economics. Kritikos, Alevizopoulos, and D Koutsouris [121] presented a VR system that can recognize individual differences and readjust the VR scenarios during the simulation according to the treatment aims, which can increase the efficiency of VR stimulations for the treatment of central nervous system dysfunctions, as it provides numerically more controlled sessions without unexpected variations. Piumsomboon et al. [172]. introduced Mini-Me, an adaptive avatar for enhancing Mixed Reality (MR) remote collaboration between a local Augmented Reality (AR) user and a remote VR user. McIntosh et al. [145] proposed a systematic approach for iteratively adapting the avatar to perform better in a given task based on users' performance. The approach was evaluated in a target selection task, where the avatar's forearms were scaled to improve performance.

The intriguing question is how to grant an avatar an adaptable nature.

3.3 Intelligent Avatar or Hybrid Avatar/Agent

While reviewing studies, "avatar" and "agent" are occasionally used interchangeably. However, it is crucial to differentiate between these two concepts to establish a precise definition of an intelligent avatar.

Avatars and agents are concepts commonly used in virtual environments and artificial intelligence. While they can sometimes overlap in functionality, they serve different purposes and have different characteristics.

As a virtual representation, an avatar is perceived to be controlled by humans in a virtual environment (see Figure 3.1). It's closely tied to the user's identity and serves as a means of self-expression. More importantly, avatars enable users to interact with the virtual environment and other users. They serve as a medium through which users navigate, communicate, and engage in virtual experiences. Avatars can be customized to reflect the user's preferences or identity. Still, their primary role is to provide a user-centric interface and facilitate the user's presence and interaction within the virtual environment. On the other hand, an agent refers to a virtual representation perceived to be controlled by computer algorithms, focusing on autonomous behavior and task performance, as is de-

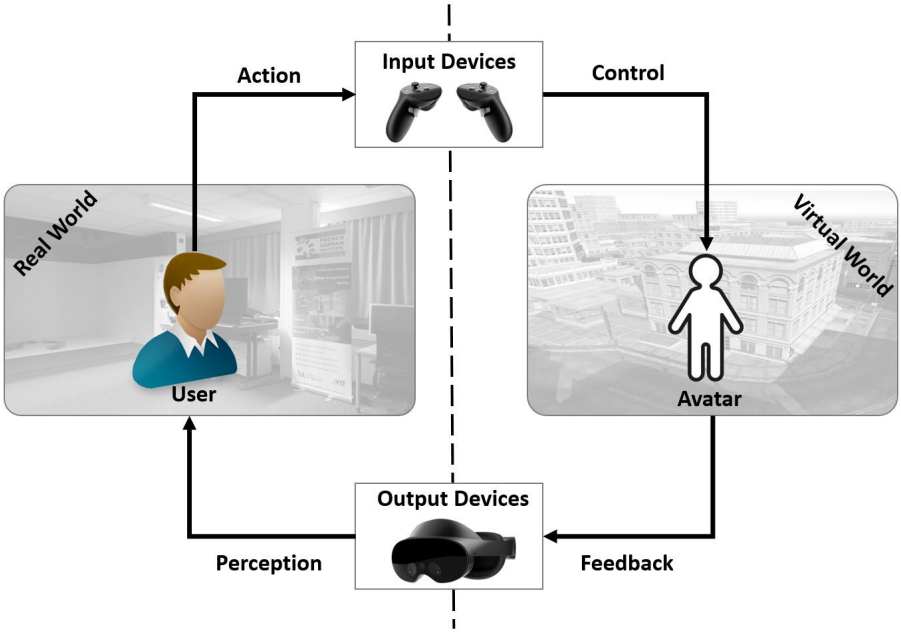


Figure 3.1: The user utilizes input devices to control the avatar, enabling interaction within the virtual environment. The avatar engages with the virtual environment through the user's actions, while the virtual environment responds by providing feedback. The user perceives this feedback through output devices, completing the continuous cycle of interaction and perception within the virtual environment.

picted in Figure 3.2. Agents are often autonomous entities driven by artificial intelligence algorithms, exhibiting intelligent behavior and decision-making capabilities. Unlike avatars, agents can operate without directly receiving input data from users or devices. In specific scenarios, agents can function autonomously or follow predefined rules without requiring real-time input from the user. They can analyze data, process information, and decide based on their programming or learned behavior. This ability allows agents to perform tasks and interact with the virtual environment or other entities self-directedly.

As explored in the section 1.4 on 'Intelligent Avatar' (IA), this concept represents a virtual user representation endowed with advanced artificial intelligence (AI) capabilities. These exceptional abilities empower the IA to demonstrate intelligent behaviors, employ reasoning and decision-making processes, and engage in sophisticated user interactions. By incorporating machine learning algorithms, the IA continuously refines its understanding of the user, becoming more effective in delivering tailored support and enhancing the overall immersive experience over time. Ultimately, the Intelligent Avatar aims to create a more intuitive and supportive virtual environment, significantly improving user satisfaction

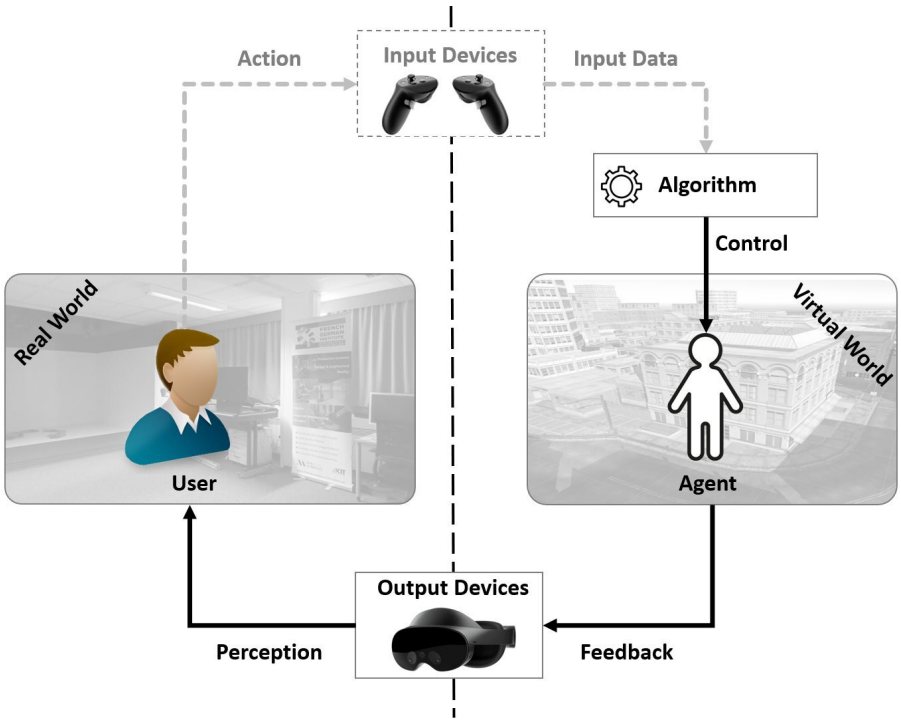


Figure 3.2: The agent's back-end algorithm can receive input data directly from the user through the input device. The agent autonomously performs actions based on its programming or learned behavior. It interacts with the virtual environment or other entities, gathering information and making decisions. The virtual environment or entities provide feedback or output in response to the agent's actions. The user perceives this feedback through user interface elements, such as output devices, facilitating their understanding of the agent's behavior and its impact on the virtual environment.

and outcomes.

However, when we consider the traditional definitions of avatars and agents, the question arises: Can we slot the intelligent avatar into one of these categories? The answer lies in the intelligent avatar's adaptability, which can exhibit traits of avatars and agents, depending on the context and application.

Incorporating AI into an avatar involves two distinct phases: learning and deploying a trained model for prediction or detection. The learning phase can take place either online or offline. Online learning refers to the AI gathering and analyzing data during user interactions with the avatar. On the other hand, offline learning involves collecting data solely during user interactions with the avatar in a virtual environment. The AI algorithm is applied to the collected data once the user's experience within the virtual environment is complete. In both cases, the learning phase aligns with the traditional

definition of an avatar, as depicted in Figure 3.3. This figure illustrates online learning within the action-perception loop. During this phase, the avatar typically represents a digital or virtual entity that a user can control.

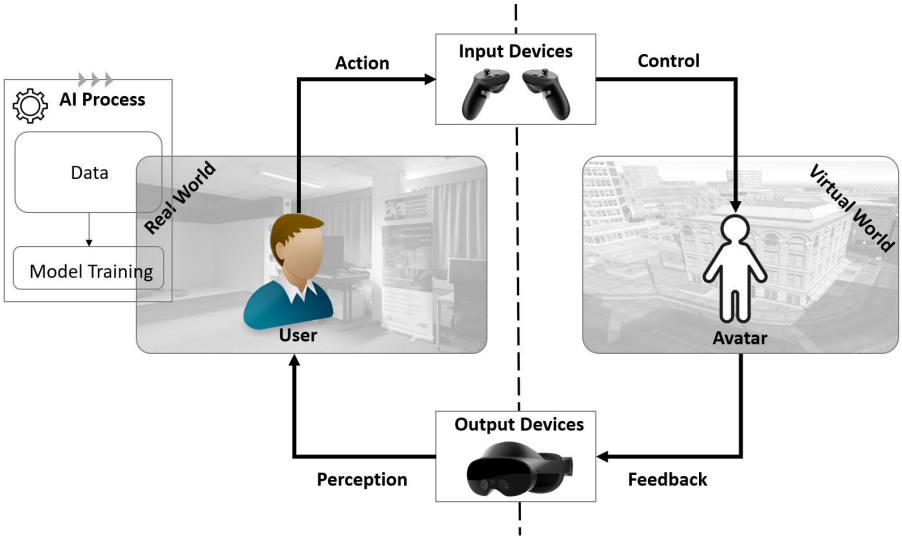


Figure 3.3: Avatar as a digital or virtual entity controlled by a user during the learning phase.

However, in the deployment phase, which primarily focuses on the intelligent avatar, a hybrid nature emerges, combining traits of avatars and agents. While a user typically controls an avatar, an agent can make autonomous decisions and interact with its environment. An intelligent avatar blends elements from both categories, as depicted in Figure 3.4, showcasing the hybrid characteristics during the deployment phase.

An intelligent avatar goes beyond the realm of a traditional avatar by incorporating techniques such as natural language processing, machine learning, computer vision, and other cognitive abilities to enhance its overall intelligence and responsiveness. This hybrid entity functions as an avatar by representing a digital or virtual presence that users can interact with. At the same time, it acts as an agent by leveraging its AI capabilities to autonomously analyze data, make decisions, and adapt its behavior based on the context and user interactions. This fusion of avatar and agent characteristics is expected to enable the intelligent avatar to provide a more immersive and interactive experience, potentially bridging the gap between human users and AI technology.

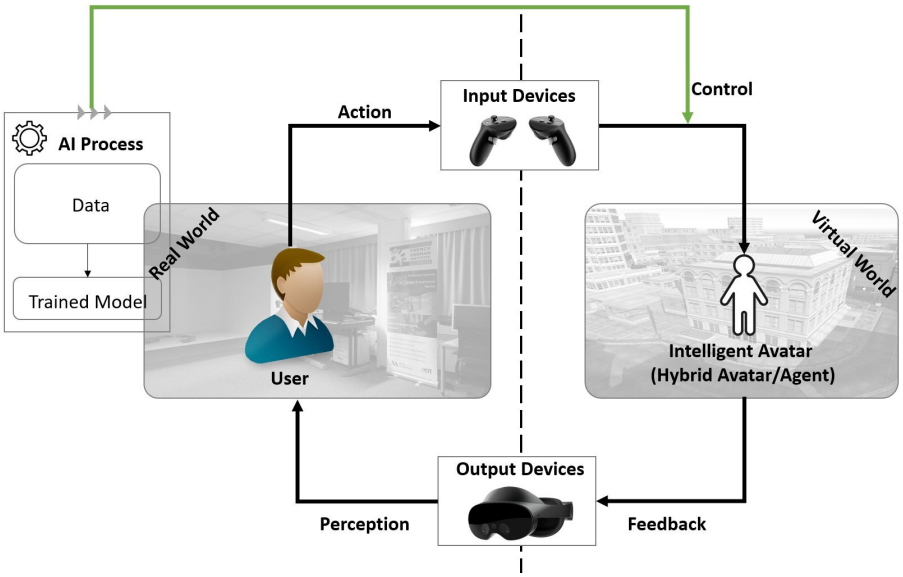


Figure 3.4: Hybrid characteristics of the avatar and agent during the deployment phase.

3.4 Human-Centered Intelligent Avatar

In addition to the "Intelligent Avatar", another concept is the avatar equipped with "Human-centered AI." This concept refers to an approach where AI systems, algorithms, and technologies used in the avatar prioritize human needs, preferences, and well-being [203]. The emphasis is on creating a customized and user-centric experience that aligns with the user's requirements. Indeed, the Human-Centered Intelligent Avatar is a subgroup or specific category within the broader concept of intelligent avatars.

While both concepts involve using AI in avatars, "Human-Centered AI in Avatar" emphasizes the user-centric approach. By considering and respecting user requirements through adapting the avatar to the user, human-centered AI aims to build trust between the user and the avatar [197]. This trust makes users feel comfortable, confident, and secure when interacting with the avatar. In this regard, Shneiderman [202] discussed the importance of human-centered AI, emphasizing the need for reliable, safe, and trustworthy AI systems. This work underscores the significance of trustworthiness in AI applications and highlights the broader implications of reliable AI on user experience and interaction.

3.5 Merging Trustworthiness and Intelligence in Avatars

Customization and adaptability are critical factors in establishing trust in VR avatars. Trustworthy avatars recognize and adapt to individual user characteristics, preferences, and needs, enhancing connection and trust. Through advanced AI techniques, these avatars can gather and analyze user data, such as past interactions, behaviors, and stated preferences, to tailor their responses and behaviors accordingly [18]. By dynamically adjusting their communication style, gestures, interaction with virtual environment, and even appearance, these avatars create a more customized experience, which leads to a stronger connection and builds trust over time [92].

Generally, avatars' ability to recognize and adapt to individual characteristics, preferences, emotional states, and contextual cues contributes to a more immersive and customized user experience, fostering a more profound sense of trust and connection within the virtual environment.

3.6 Training Customized Intelligent Avatars

In this section, we examine the training of customized intelligent avatars, beginning to showcase details of how these avatars adapt by gathering and analyzing user data during interactions. This real-time learning approach contrasts sharply with the limitations of offline learning, where static datasets hinder the ability to capture individual user nuances, leading to less personalized experiences. However, technical and practical obstacles are faced when implementing effective real-time training. Together, these subsections illustrate the importance of real-time learning for developing truly adaptive intelligent avatars while acknowledging the limitations and challenges of traditional methods.

3.6.1 Customized Intelligent Avatars Learning in Real-Time

To create a truly customized virtual twin of the user within a virtual environment, it is crucial for the avatar to "learn" from the user during their interactions. By leveraging real-time learning techniques, the avatar can continually gather and analyze the user's physiological and behavioral characteristics and responses within the virtual environment. This ongoing learning process allows the avatar to adapt and personalize its behavior. It creates a virtual representation that closely aligns with the user's characteristics and desires, resulting in a more immersive and tailored user experience.

With AI and machine learning algorithms, the avatar can learn to interpret these signals and customize its interactions accordingly. For example, if the user displays signs of anxiety or stress, the avatar can adjust its behavior to provide calming and supportive responses. By leveraging real-time training, the avatar can continually refine its understanding of the user's preferences, emotional states, and communication patterns, leading to a more accurate and personalized representation.

3.6.2 Limitations of Offline Avatar Learning

Offline avatar learning, where the intelligent avatar is trained using pre-recorded data or a static dataset, poses significant drawbacks compared to real-time, adaptive learning approaches. One of the primary limitations is the difficulty in capturing individual users' nuanced and distinctive characteristics. In offline learning, the avatar is typically trained using data gathered from a panel of users, each with unique behavior patterns, communication styles, and emotional responses. However, this generalized training approach may fail to model individual users' personalized nuances adequately. As a result, the avatar's behavior may need to align better with the specific preferences and contextual needs of the end-user, leading to a less personalized and engaging user experience.

Furthermore, offline training needs more responsiveness and adaptability to keep pace with the dynamic nature of human behavior and preferences. Even if the avatar was trained exclusively using the user's own pre-recorded data of the user, it could not capture the evolving nature of the user's needs and contextual factors in real-time. Consequently, the avatar's responses may become outdated and disconnected from the user's current state.

Another significant drawback of offline avatar learning is its limited adaptability to changing circumstances. Some virtual applications, such as driving simulators, involve dynamic environments where user preferences and states fluctuate over time. Once trained, offline learning models generally cannot self-adapt and update to reflect these changes. This can result in a mismatch between the avatar's learned behavior and the user's current preferences.

3.6.3 Challenges of Real-Time Learning

Real-time training for human-centered intelligent avatars presents several challenges that need to be addressed for effective implementation. Firstly, acquiring and processing real-time physiological and behavioral data requires robust and reliable sensing technologies. Accurate and timely capture of

signals such as heart rate, skin conductance, facial expressions, and body movements is crucial for the avatar to make informed decisions and adapt its behavior accordingly. However, if multiple sensors are used to acquire data with different frequencies, synchronizing them in real-time becomes challenging. Ensuring the availability of high-quality sensor data in real-time and the accurate synchronization of heterogeneous sensor data streams is essential for maintaining the integrity and reliability of the training process.

Another significant challenge in real-time training is the computational complexity involved. Personalized avatars often employ machine learning algorithms to analyze and interpret the user's data, allowing them to understand and respond appropriately. However, training these algorithms in real-time demands significant computational resources to quickly process and learn from the data stream. The system hosting the avatar must have sufficient computational power to handle the training process without causing delays or interruptions in the user's experience.

Additionally, maintaining privacy and security during real-time training is a critical concern. Personalized avatars rely on sensitive user data, including physiological signals and behavioral patterns. Safeguarding this information and ensuring user privacy is of utmost importance. Robust data encryption, secure storage, and adherence to privacy regulations are essential to mitigate potential risks and build user trust. Moreover, determining the stopping criteria for avatar learning and ensuring the achievement of a satisfactory level of reliability is a crucial task. Real-time training involves continuously updating the avatar's knowledge and behavior based on the user's cues and interactions.

However, it is important to define when the learning process should stop, and the avatar should transition into a reliable state. Without a clear stop point, the avatar may continue learning indefinitely, potentially leading to overfitting or an excessive adaptation that impairs its overall performance. Establishing appropriate criteria and thresholds to determine the point at which the avatar has learned enough and achieved a reliable state is crucial.

Addressing these challenges requires interdisciplinary efforts, including advancements in sensor technologies, computational capabilities, data privacy measures, and algorithmic development. Overcoming these issues will pave the way for more sophisticated and adaptable real-time training techniques, enabling personalized intelligent avatars to provide more prosperous and more immersive user experiences.

3.7 Summary

This chapter focuses on trustworthy intelligent avatars, which aim to create user reliability and confidence. Trustworthiness is achieved through authenticity, transparency, visual realism, natural behavior, effective communication, and adaptation. Auto-adaptation enhances trustworthiness and user engagement in virtual environments. Avatars can incorporate machine learning or artificial intelligence techniques to adapt their behavior, responses, or recommendations based on user feedback or previous experiences, improving their performance over time. Figure 3.5 is a snapshot of the relation between these four concepts.

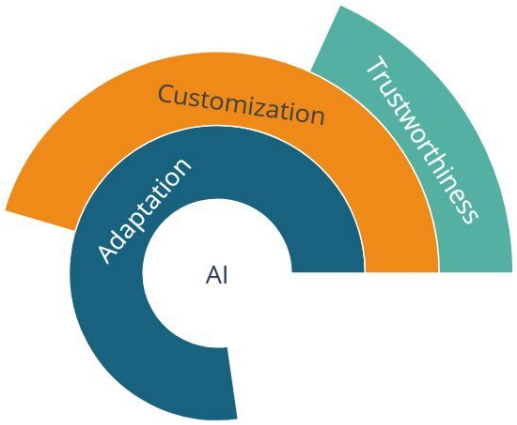


Figure 3.5: The conceptual hierarchical relationship between trustworthiness, customization, adaptation, and artificial intelligence. By leveraging artificial intelligence techniques, we can achieve a high degree of adaptive capabilities and customization in intelligent avatars, which in turn can foster a sense of trustworthiness and reliability in the user's interactions within the virtual environment.

Avatars and agents are distinct concepts in virtual environments and AI. Avatars serve as user representations and enable interaction, while agents exhibit autonomous behavior. Intelligent avatars bridge the gap between avatars and agents by combining AI capabilities with user control during learning and autonomous behavior during deployment. This integration prioritizes user control and aims to create a seamless and cohesive experience for users, where they can interact with avatars that possess intelligence and adaptability.

Offline avatar learning has limitations, including a lack of responsiveness, adaptability, and inability to capture individual user characteristics and preferences. On the other hand, real-time learning for avatars presents challenges such as robust sensing technologies, computational complexity, privacy and

security concerns, and determining stopping criteria. Interdisciplinary efforts are needed to overcome these challenges and enhance the user experience with personalized intelligent avatars.

The next chapter will outline the methodology employed to develop the trustworthy intelligent avatar and present the selected use case.

Chapter 4

Trustworthy Intelligent Avatar Implementation and Use-case

Content

As briefly mentioned in the Introduction, we have developed a driving simulator to demonstrate the practical implementation of our approach, serving as a compelling case study. Within our driving simulator, the user takes on the pivotal role of the driver, wielding complete control over the car through driving tools. In this context, the car becomes the avatar, referred to as an *"avacar"* in the research literature [181]. Henceforth, whenever we use the term "car," we refer to the avatar. We intend to equip the car with human-centered intelligence, enabling it to interact with the user in a responsive and adaptive manner.

One fundamental question that sparks curiosity is: how can this intelligent avatar substantially enhance the user experience within virtual environments?

To answer this query, the capabilities of the intelligent avatar will be leveraged to address a prevalent challenge in VR applications known as cybersickness. By employing this intelligent avatar, we aim to mitigate cybersickness and create a more comfortable and immersive VR experience for users. Figure 4.1 shows an overview of the application.

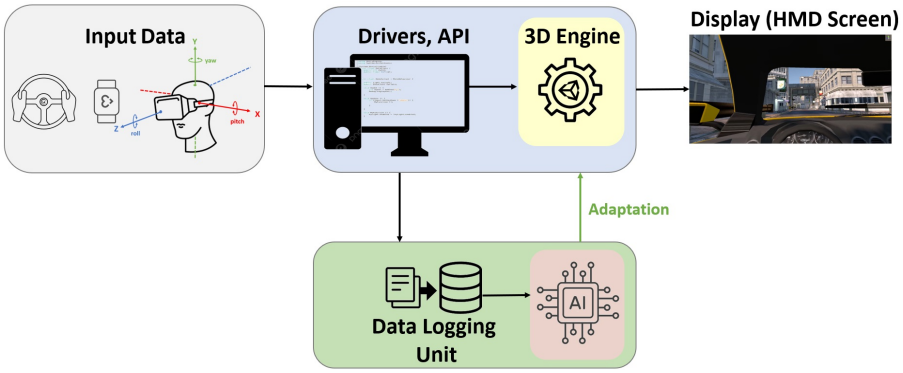


Figure 4.1: Application Overview.

However, the real-time processing of data and the ability to auto-adapt pose significant challenges for specific virtual applications, such as driving simulators. These simulators require the capability to render detailed environments, simulate precise vehicle dynamics, and process user inputs in real-time. These computational tasks can strain the central processing unit (CPU), particularly when running demanding driving simulator applications or performing multiple tasks concurrently, such as simulating complex traffic scenarios and providing immediate feedback. While hardware advancements have partially addressed high CPU usage by introducing more powerful CPUs and GPUs designed for simulation and gaming, these resources may not be accessible to everyone. Therefore, developing

optimized software and platforms becomes crucial in maximizing the utilization of available equipment, ensuring efficient performance even with less powerful hardware configurations.

Every Adaptive Virtual Reality Systems (AVRS) is built upon three key features. One key feature of AVRS is performance measures, encompassing various aspects such as task performance, physiological data, and kinematic/kinetic data. Task performance measures evaluate the user's proficiency and success in completing specific tasks within the virtual environment. Physiological data, such as heart rate, skin conductance, or brainwave activity, provides insights into the user's emotional state, stress levels, or cognitive load. Kinematic/kinetic data captures the user's movements and interactions, enabling assessment of physical performance and biomechanical factors. By integrating these performance measures, AVRS can gather real-time data about the user's engagement, physical response, and cognitive abilities, allowing for adaptive adjustments during the VR experience.

Another crucial feature of AVRS is the adaptive logic component, powered by rule-based systems, AI algorithms such as classification algorithms, and optimization methods. These AI algorithms enable VR applications to analyze and interpret user inputs intelligently, paving the way for more advanced and dynamic interactions within virtual environments. Rule-based systems utilize predefined rules to determine adaptive actions based on specific conditions or user characteristics. On the other hand, optimization methods aim to find the optimal configuration of the virtual environment based on performance measures and user preferences. By employing these adaptive logic techniques, AVRS can dynamically adapt various elements of the virtual reality experience, such as content presentation, difficulty levels, or environmental conditions, to enhance user satisfaction and engagement.

Additionally, AVRS incorporates a range of adaptive variables to enhance the virtual reality experience further. These variables encompass the simulated environment, controlled elements, display features, and scenario difficulty. Each variable can be dynamically adjusted to suit the user's preferences and performance, enhancing the overall VR experience.

The implemented AVRS utilized physiological/behavioral data and questionnaires as performance measures. A classification algorithm (AI) served as the adaptive logic, while a simulated environment (car acceleration) functioned as the adaptive variable. Figure 4.2 showcases these three key features.

This chapter will explore the theoretical foundation of cybersickness which was chosen as a safety index and the intricacies of developing this intelligent avatar, encompassing the integration of user

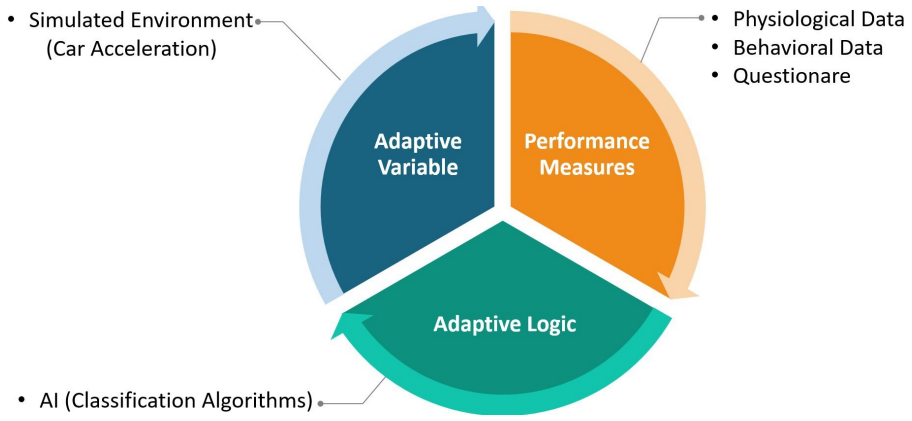


Figure 4.2: Adaptive Virtual Reality Systems (AVRS) Features.

states, and implementing human-centered intelligence. Additionally, we will delve into how this intelligent avatar can effectively tackle the challenges associated with VR applications, particularly in the context of cybersickness mitigation. By doing so, we aim to contribute to the ongoing advancement of avatar technology and enrich the overall VR experience for users.

4.1 Artificial Intelligence(AI) and Machine Learning for VR Platform

The rapid advancements in Artificial Intelligence (AI) and Machine Learning (ML) have unlocked new possibilities for enhancing VR experiences. AI and ML algorithms have become instrumental in powering various aspects of VR applications, from realistic avatar rendering to adaptive user interfaces and personalized content delivery.

At the core of AI's integration with VR is the ability to process and analyze the vast amounts of data generated by users' interactions within the virtual environment. Through machine learning techniques, VR systems can learn from user behaviors, preferences, and interactions to create highly immersive and responsive experiences tailored to individual users.

One of the primary roles of AI in VR is to enable natural and intuitive interactions between the user and the virtual world. This is achieved through computer vision, natural language processing, and gesture recognition, which allow the VR system to understand user inputs and respond accordingly. By seamlessly interpreting user commands and actions, AI-powered VR can create a sense of presence, making the virtual experience more engaging and enjoyable.

4.1.1 Overview of common AI/ML Applications in VR

Integrating AI and ML in VR has led to the development of various applications that enhance the user experience and enable new capabilities. The following sections will explore some key areas where the convergence of AI and ML has profoundly impacted VR applications and experiences.

- **Personalization and Adaptive Learning**

  ◦ *User Profiling.* AI and ML algorithms can analyze a wealth of user data to create detailed user profiles, including behavior, preferences, physiological responses, and even emotional states. These comprehensive user profiles enable VR systems to tailor the experience to individual needs, preferences, and learning styles, enhancing engagement, effectiveness, and overall user satisfaction.

  ◦ *Adaptive Content.* In educational VR applications, AI can adjust the difficulty level and type of content based on the learner's progress and performance. This ensures that the learning experience is neither easy nor difficult, maintaining optimal engagement and learning efficiency.

- **Real-time Interaction and Feedback:**

  ◦ *Gesture and Voice Recognition.* AI-powered gesture and voice recognition systems enable more natural and intuitive interactions within VR environments. These systems can interpret user commands and actions in real time, providing immediate feedback and enhancing the sense of immersion.

  ◦ *Emotion Recognition.* AI can detect the user's emotional state and adjust the VR environment accordingly by analyzing facial expressions, voice tone, and physiological signals (e.g., heart rate, skin conductance). This can be particularly useful in therapeutic settings, where the environment can be modified to reduce stress or anxiety.

- **Enhanced Realism and Immersion:**

  ◦ *Procedural Content Generation.* AI algorithms can generate complex and dynamic virtual environments on the fly based on predefined rules and user interactions. This allows for

the creation of vast, detailed, and varied virtual worlds without requiring extensive manual design.

◦ *Physics and Behavior Simulation.* ML models can simulate realistic physical interactions and behaviors of virtual objects and characters. This includes accurate collision detection, fluid dynamics, and character animations that respond naturally to user actions.

- **Health and Rehabilitation:**

  ◦ *Therapeutic VR.* AI-driven VR applications are used in physical and mental health treatments. For example, in physical rehabilitation, AI can monitor patient movements and provide real-time feedback to perform exercises correctly. In mental health, VR environments can be adjusted to create calming or stimulating experiences based on the patient's needs.

  ◦ *Monitoring and Assessment.* AI can continuously monitor physiological signals (e.g., heart rate and skin conductance) to assess the user's state and progress. This data can be used to adjust the therapy in real-time or to provide detailed reports to healthcare providers.

- **Training and Simulation:**

  ◦ *Skill Training.* AI-powered VR simulations are used for training in various fields, such as medical procedures, military operations, and industrial tasks. These simulations can adapt to the trainee's skill level, provide real-time feedback, and create realistic scenarios that enhance learning and retention.

  ◦ *Performance Analysis.* ML algorithms can analyze performance data to identify strengths and weaknesses, providing insights that help improve training programs. This can include tracking reaction times, decision-making processes, and task execution accuracy.

4.1.2 Mathematical Foundations of ML: Supervised, Unsupervised, and Reinforcement Learning

Machine learning encompasses various techniques and algorithms that enable systems to learn from data and make predictions or decisions without being explicitly programmed. The mathematical foundations of machine learning can be broadly categorized into three main areas: supervised

learning, unsupervised learning, and reinforcement learning. Each of these areas has its mathematical underpinnings and formulations.

I. Supervised Learning Supervised learning involves learning a mapping function from input data (features) to output data (labels or targets) based on a set of labeled training examples. The goal is to build a model that accurately predicts new, unseen data.

The mathematical foundation of supervised learning is based on the concept of empirical risk minimization (ERM). The objective is to find a function f that minimizes the expected risk or loss over the joint distribution of inputs and outputs:

$$\min_{f \in \mathcal{F}} \mathbb{E}_{(X,Y)\sim P} l(f(X), Y) \tag{4.1}$$

Here, $\mathcal{F}$ is the hypothesis space (set of possible functions), l is the loss function (e.g., mean squared error, cross-entropy), and P is the underlying data distribution. Since the true distribution P is unknown, the empirical risk minimization principle is used, where :

$$\min_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^{n} l(f(x_i), y_i) \tag{4.2}$$

This formulation is the basis for various supervised learning algorithms, such as linear regression, support vector machines (SVMs), neural networks, and logistic regression. Here, we briefly outline the logistic regression model, as it plays a crucial role in the AI components of our development.

Logistic Regression. Logistic regression is a popular algorithm for binary classification tasks. The goal is to learn a linear function that maps the input features to the probability of belonging to the positive class:

$$f(x) = \sigma(w^\top x + b) \tag{4.3}$$

where σ is the sigmoid function. The objective is to minimize the cross-entropy loss:

$$\min_{w,b} \frac{1}{n} \sum_{i=1}^{n} [-y_i \log f(x_i) - (1 - y_i) \log(1 - f(x_i))] \tag{4.4}$$

This optimization problem is typically solved using iterative methods like gradient descent or Newton's method.

II. Unsupervised Learning Unsupervised learning involves finding patterns or structures in data without labeled responses. The goal is to learn the underlying structure of the data. Common tasks in unsupervised learning include clustering (K-Means Clustering), dimensionality reduction (Principal Component Analysis (PCA)), and density estimation (Gaussian Mixture Models (GMM)). These algorithms help uncover insights from unlabeled data, making them vital in various applications.

III. Reinforcement Learning (RL) Reinforcement Learning (RL) is a type of machine learning where an agent learns to make decisions by performing actions in an environment to maximize cumulative reward. The mathematical foundation of RL is based on Markov Decision Processes (MDPs) and the Bellman equation.

4.1.3 Stream Learning for VR

Stream learning, also known as online learning or incremental learning [5], is a paradigm in machine learning that operates in real time. Models are trained and updated as data arrives sequentially rather than on a fixed dataset. In other words, the critical distinction of stream learning is that the model does not have access to the entire dataset at once but processes data instances one by one or in small batches. This allows the model to be updated immediately based on the latest observations, reducing the latency between data generation and model update. This incremental processing makes stream learning well-suited for dynamic environments where data distributions, such as virtual reality (VR) applications, can change over time.

One of the primary benefits of stream learning is its ability to handle concept drift - the phenomenon where the underlying data-generating process changes over time. Stream learning algorithms can adapt to these changes and maintain high performance by continuously updating the model, whereas batch-trained models may quickly become outdated. This is particularly crucial in VR, where user behaviors, environmental conditions, and other factors can evolve rapidly.

Stream learning approaches are often more computationally and memory-efficient than retraining a batch model from scratch. Since only the current data instance or mini-batch is used to update the

model, the memory footprint can be kept relatively small, making stream learning models suitable for deployment on resource-constrained VR platforms. The online nature of the updates also allows for faster model iterations and responsiveness to changing user needs.

Furthermore, stream learning allows for adding new functionality or personalization to the VR system without the need to retrain the entire model. This allows for more agile and user-centric development of AI-powered VR experiences.

4.1.3.1 Challenges of Traditional Batch Learning in Dynamic VR Environments

Traditional batch learning methods face significant challenges when applied to dynamic VR environments. These environments, with their continuous data generation, real-time interaction requirements, and evolving user behaviors, present obstacles that traditional methods struggle to overcome.

Traditional batch learning models are ill-equipped to handle real-time data processing. Their requirement that the entire dataset be available before training is simply impractical in dynamic VR environments where data is continuously generated. This limitation delays model updates and hinders the ability to provide real-time feedback and adapt to new information, leading to outdated or irrelevant training scenarios in VR simulations.

In dynamic VR environments, user behaviors, preferences, and interactions can evolve, leading to concept drift, which refers to changes in the underlying data distribution over time. Traditional batch learning models, once trained, need to adapt to these changes. This results in degraded model performance over time as the model becomes less relevant to the current data. For instance, a VR recommendation system may fail to adapt to changing user preferences, leading to poor user experience. Additionally, batch-learning models often require significant computational resources and memory to process large datasets. This is especially problematic in VR environments where data can be high-dimensional and voluminous. The high resource requirements make it challenging to train and update models frequently. This can lead to inefficiencies and increased latency in VR applications, such as real-time rendering and interaction.

Furthermore, updating a batch learning model typically involves retraining the model from scratch with the entire dataset, including new data. This process is time-consuming and computationally expensive. The inflexibility in updating models makes it difficult to incorporate new data quickly.

In dynamic VR environments, this can result in outdated models that need to reflect the latest user interactions and environmental changes. As the volume of data grows, batch-learning models need help to scale effectively. The need to store and process large datasets can become a bottleneck. Scalability issues can hinder the deployment of VR applications that require continuous learning from large-scale data, such as social VR platforms or large multiplayer VR games.

VR environments often generate sparse and noisy data due to the variability in user interactions and sensor inaccuracies. Batch learning models may need help handling such data effectively, leading to poor performance. This can affect the accuracy and reliability of VR applications, such as motion tracking and gesture recognition.

4.1.3.2 Stream Learning Algorithms and Techniques

Some stream learning algorithms include Stochastic Gradient Descent (SGD), Incremental Decision Trees, Online Bayesian Learning, Incremental Clustering, Fuzzy Decision Trees, and Federated Learning. For instance, SGD updates model weights incrementally based on each data point, making it efficient for large datasets. Incremental Decision Trees grow by adding new branches without needing to rebuild the entire tree, while Online Bayesian Learning updates probabilities as new data arrives. Incremental Clustering adjusts clusters dynamically, and Fuzzy Decision Trees allow for soft classifications where data points can belong to multiple classes. Federated Learning enables decentralized model training across multiple devices while keeping data localized, enhancing privacy, and reducing data transfer. Together, these algorithms facilitate the development of responsive and scalable models in ever-changing environments.

Stochastic Gradient Descent (SGD): One of the fundamental algorithms for stream learning is Stochastic Gradient Descent (SGD). In SGD, the model parameters are updated after processing each data point, rather than the entire batch. The update rule for SGD is:

$$\theta_{t+1} = \theta_t - \eta \nabla_\theta L(\theta_t; x_t, y_t) \tag{4.5}$$

where:

- θ_t are the model parameters at time t.
- η is the learning rate.

- $\nabla_\theta L(\theta_t; x_t, y_t)$ is the gradient of the loss function concerning the parameters, evaluated at the current data point (x_t, y_t).

4.1.3.3 Handling Concept Drift and Model Updates in Stream Learning

As mentioned in section 4.1.3.1, drift is a critical challenge in stream learning. Concept drift refers to changes in the statistical properties of the target variable over time, which can degrade the performance of predictive models if not properly addressed. The goal of a drift detector is to send a signal when the drift is present. An effective drift detector seeks to maximize the number of true positives (TP) while minimizing the number of false positives (FP). It must also be computationally efficient to handle the continuous data flow in real-world applications. In VR applications, as user behaviors, preferences, and system conditions evolve in VR environments, the relationship between input features and target variables can shift, rendering previously trained models less accurate or even obsolete. Stream learning algorithms equipped with mechanisms to detect and adapt to concept drift demonstrate their adaptability. This capability reassures audiences about the effectiveness of these algorithms in managing changes in a timely and efficient manner. Below is a brief overview of some algorithms designed to detect and adapt to concept drift:

- **Adaptive Learning Rates:** Adjusts the learning rate dynamically to respond quickly to changes in data distribution.
- **Ensemble Learning:** Combines multiple models to improve robustness and stability against concept drift.
- **Sliding Window:** Focuses on the most recent data points, allowing the model to adapt to recent trends and changes.
- **Adaptive Forgetting Factor:** Gives more weight to recent data, enabling quicker adaptation to new patterns.
- **Meta-Learning:** Learns how to adapt to new tasks or shifts in data distribution, improving overall model flexibility.
- **Adaptive Windowing(ADWIN):** Adaptive windowing techniques adjust the size of the window based on the rate of change observed in the data. Adaptive Windowing (ADWIN) maintains a

variable-length window and detects changes by comparing the statistics of two sub-windows [23]. Let W be the current window, and W_1 and W_2 be two sub-windows of W. Change is detected if:

$$|\mu_{W_1} - \mu_{W_2}| > \epsilon \tag{4.6}$$

where μ_{W_1} and μ_{W_2} are the means of the sub-windows, and ϵ is a threshold determined by the desired confidence level.

4.2 SmartSimVR: Flexible Software Architecture

The implementation of the auto-adaptive intelligent avatar builds upon various modules. An advanced architecture called *"SmartSimVR"* was developed to have a closed-loop architecture that integrates intelligent VR applications with real-time adaptation in a single cycle (see Figure 4.3). The designed architecture was constructed upon four fundamental components:

1. **Distribution:** The integration involves distributing the architecture's modules across multiple systems, with the VR application on one system and the AI module on another. Various methods, such as TCP/IP, HTTP requests, and message queues, enable smooth data transfer between the two, ensuring efficient and reliable communication.

2. **Concurrent Processes:** The real-time auto-adaptation system relied on a well-designed multi-threaded architecture for efficient data processing. This architecture utilized concurrent processes to handle dynamic adaptive VR applications, ensuring real-time adaptation based on user inputs and environmental changes.

3. **Shared Virtual Memory System:** A shared virtual memory system enabled smooth transitions and synchronized data sharing among concurrent processes, including the AI module. This mechanism facilitates efficient coordination and collaboration between components, improving overall performance by reducing inter-process communication overhead.

4. **Stream Learning as the Adaptive Logic:** The designed architecture utilized a stream learning approach detailed in the previous section.

Figure 4.4 represents the overview of the software architecture and its four key features. The architecture designed for the auto-adapted intelligent avatar was developed with a high degree of

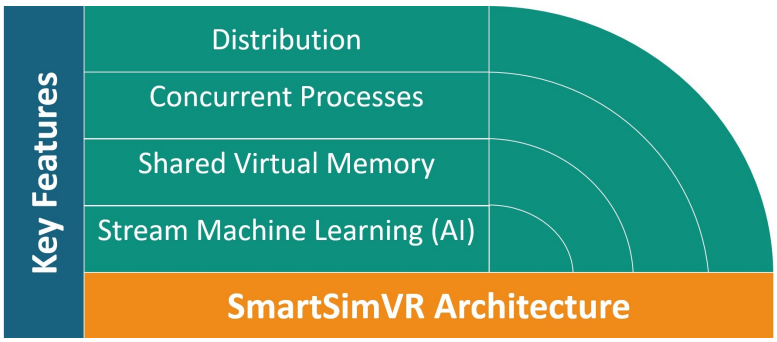


Figure 4.3: Key features of SmartSimVR.

flexibility in mind, enabling its application to a wide range of real-time auto-adapted intelligent VR applications. The modular and distributed nature of the architecture allows for the easy integration of various components, including the VR application and the AI module, across different systems. This flexibility supports the deployment of the architecture in diverse VR scenarios, as each application's specific requirements and constraints can be accommodated through adjustments to the distribution and communication methods between the modules. Furthermore, the concurrent processing capabilities and the shared virtual memory system provide a scalable and adaptable framework that can handle the dynamic demands of various intelligent VR applications, ensuring real-time responsiveness and seamless user experiences. Additionally, the architecture can integrate any number of sensors as the hardware limitation allows multiple sampling rates, allowing for the seamless incorporation of diverse data sources to enhance the adaptive capabilities of intelligent VR applications. Finally, the stream learning approach as the adaptive logic enables the architecture to be readily applied to different domains, as the learning mechanisms can be tailored to the specific needs and requirements of each auto-adapted intelligent VR application.

4.2.1 Software Implementation Details

Figure 4.5 shows the integration and communication of the various modules detailed in the previous sections of this chapter. It provides a comprehensive visualization of how these key components interact and work together within the overall system.

Figure 4.6 provides a comprehensive overview of the SmartSimVR and illustrates its data flow. The software operates as a distributed application, utilizing a TCP/IP socket for communication and data exchange between the client and server processes, ensuring efficient connectivity.

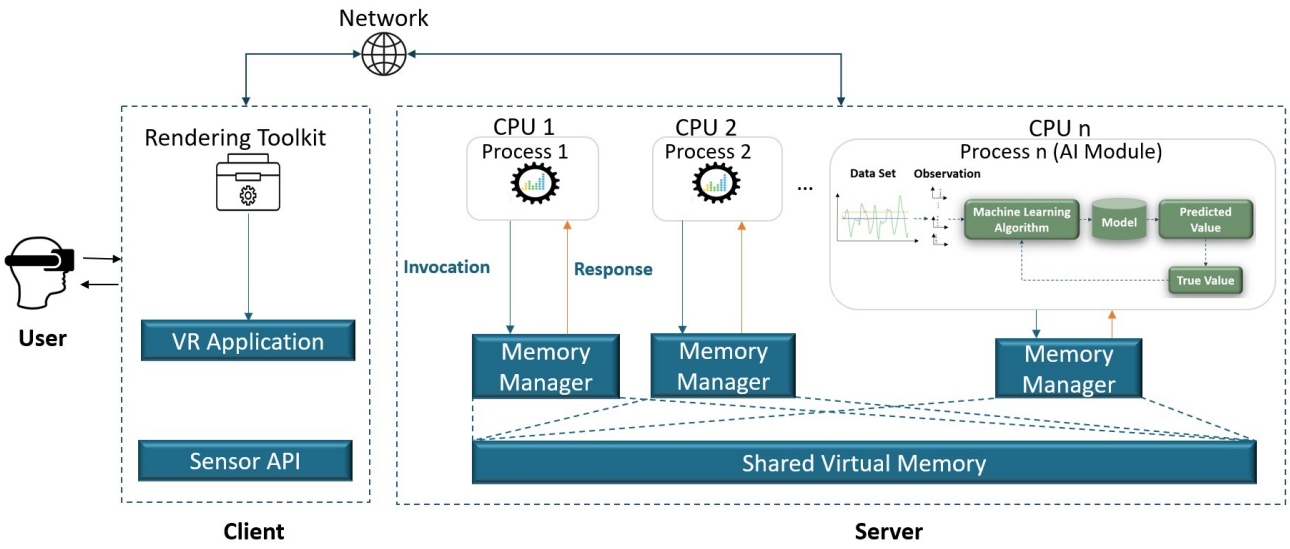


Figure 4.4: Overview of software architecture.

This system enables simultaneous data recording and processing by incorporating multiple concurrent processes. Each process is designed to fulfill specific purposes at pre-configured intervals, such as every minute. This concurrent approach ensures the architecture can efficiently handle the continuous data stream, facilitating accurate and timely processing.

I. Client Side:

The client side of the system comprises two crucial processes that play integral roles in its functionality.

- **VR Application:** The VR Application forms the system's core, as detailed in section 4.3, providing users with an immersive virtual reality experience tailored for real-time auto-adapted intelligent applications. Additionally, this process captures behavioral data, i.e., eye-tracking and head movement, during the VR experiment, as detailed in section 4.5.1.

II. Server Side:

The server side of this implementation incorporates multiple essential processes.

- **Data TCP Client:** As mentioned in section 4.5.1.2, this process receives crucial eye tracker data and head movement data from the VR application. It establishes a TCP connection and ensures the seamless transmission of these critical behavioral measurements to the Brain module for further analysis.

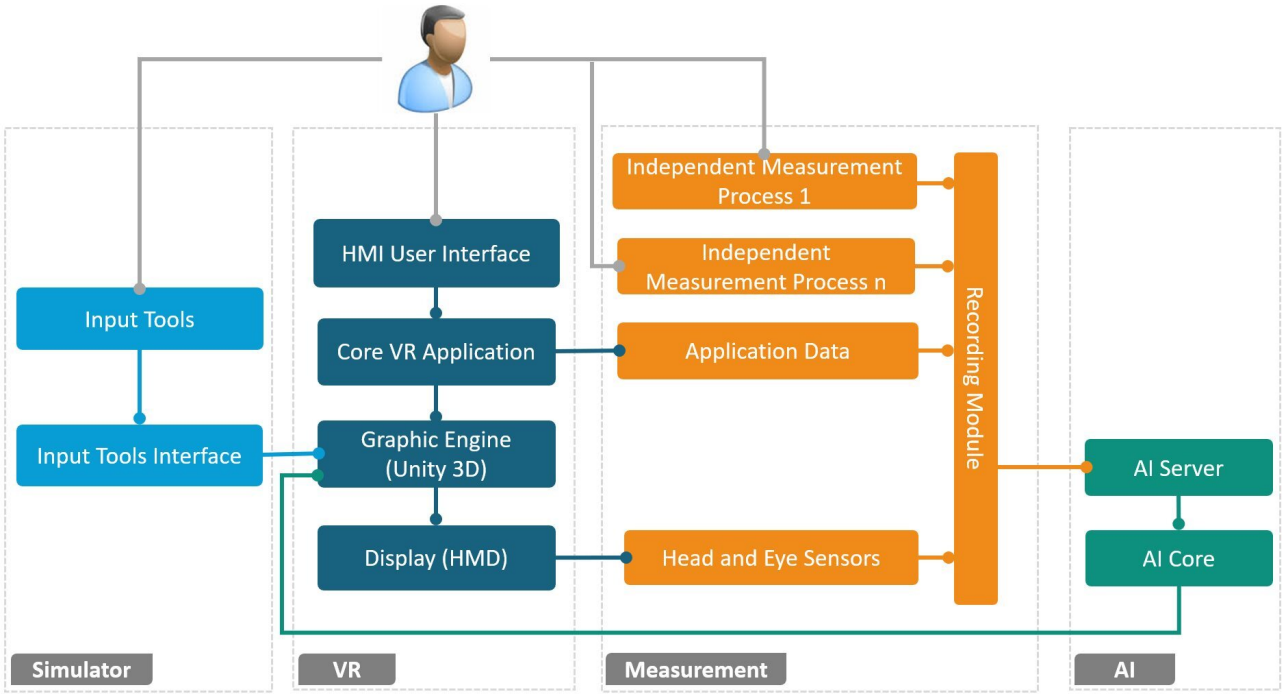


Figure 4.5: System design and collaboration between various modules and components.

- **Measurement Process:** The Measurement process focuses on receiving data from various sources, which may include a dedicated Measurement Streaming Server or directly from measurement devices. This acquired data is then available for processing and integration into the system's analysis pipeline.

- **Brain Module:** The Brain Module serves as the central component on the server side, encompassing three sub-modules:

  ◦ **Preprocessing Module:** This module plays a crucial role in data gathering from different processes, fusing, and synchronizing the time series data collected from various sensors to a unified frequency. It ensures that the data is cleaned and transformed, addressing missing values and inconsistencies. The module performs data normalization to scale the features to a common range, which is essential for effective analysis and modeling. By standardizing the data format and structure, it prepares the dataset for subsequent processing. This comprehensive approach enhances the reliability and accuracy of insights derived from the integrated data, ultimately supporting more effective decision-making in the system.

  ◦ **AI Module:** In this module, a machine learning model utilizing a stream learning ap-

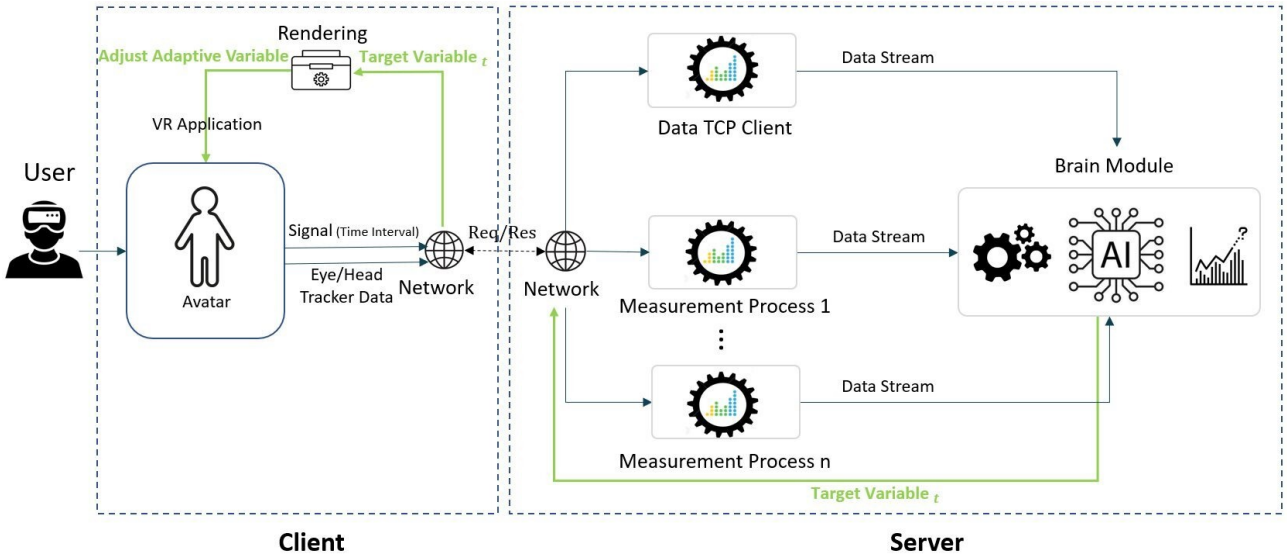


Figure 4.6: Data flow diagram of the SmartSimVR architecture, a distributed system implemented across the client and server.

proach receives and continuously trains on data. A suitable training algorithm optimizes the model's performance, allowing real-time adjustments based on the incoming data stream. The model can be applied to various tasks, such as classification or regression, depending on the nature of the incoming data and the specific objectives of the analysis.

◦ **Detection/Adaptation Module:** After sufficient training according to the prescribed protocol, the trained model receives data from the Data TCP Client and the Measurement TCP Clients at a pre-set interval. While continuing to train, the model predicts the target variable and sends a signal to the VR application for adaptation based on these predictions.

4.3 Use Case: Driving Simulator and Avatar Design

In recent years, driving simulators have emerged as powerful tools for training, research, and development in transportation. These immersive systems provide a safe and controlled environment where drivers can practice their skills, researchers can investigate driver behavior, and engineers can test and refine vehicle technologies. A critical component that enhances the realism and engagement of driving simulators is using avatars, virtual representations of vehicles, or drivers within the simulation. However, one significant challenge with driving simulators is the occurrence of cybersickness, a form of motion sickness induced by VR experiences.

Integrating avatars in driving simulators brings a new level of visual feedback, interaction, and customization possibilities. For instance, an avatar can represent the driver, allowing them to see themselves within the virtual environment, or it can embody the vehicle being simulated, providing a dynamic and responsive representation of its behavior and characteristics. The design of such avatars, mainly when the avatar represents a car, requires careful consideration of factors like realism and functionality. However, cybersickness can significantly impact driving simulators' effectiveness and user experience. As avatars play a vital role in enhancing the realism and engagement of driving simulators, their design and implementation can contribute to the occurrence and mitigation of cybersickness.

This section explores the fascinating intersection of driving simulators and avatar design, specifically focusing on car avatars. We delve into the technology and components that make driving simulators possible, highlighting the hardware and software systems that create an immersive driving experience. Additionally, we examine the significance of avatars in driving simulators and their role in enhancing user engagement and presence. To realize this vision, we implemented SmartSimVR for our use case, as depicted in Figure 4.7. All of the modules will be detailed in the following sections.

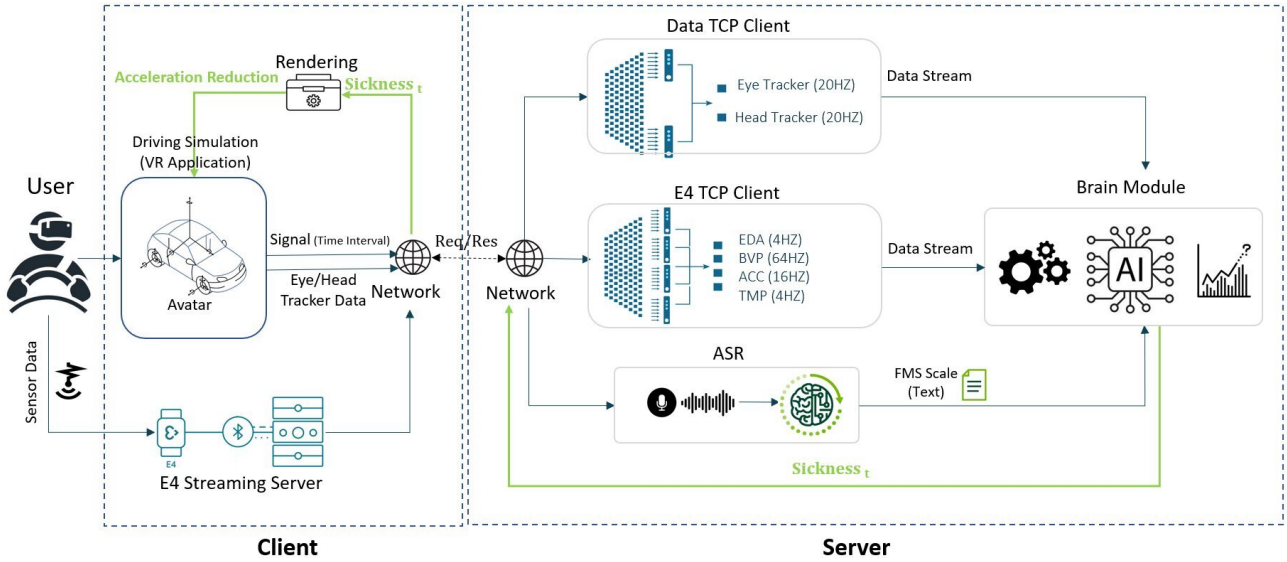


Figure 4.7: Data flow diagram and detailed implementation of customized SmartSimVR architecture for Driving Simulator use case.

We start with the design considerations for car avatars. We discuss the factors influencing car avatar design, including the physics and dynamics model of the driving simulator.

4.3.1 Driving Simulator Technology

The concept of simulation, defined as the emulation of a specific behavior through a generic imitating system [215], finds significant application in driving simulation technology. This technology is a powerful tool for replicating real-world driving scenarios within a controlled and immersive virtual environment. By accurately simulating vehicle dynamics, road conditions, and driver interactions, driving simulators provide an invaluable platform for training, research, and development in transportation.

A driving simulation's primary objective is to faithfully recreate the behavior and characteristics of real-world driving. This involves modeling the physics and dynamics of vehicles, encompassing factors such as acceleration, braking, steering, and suspension. By accurately emulating these behaviors, driving simulators allow users to experience a realistic driving environment and practice essential driving skills. Driving simulators offer many real-world scenarios, including urban streets, highways, adverse weather conditions, and emergencies. These scenarios can be precisely controlled, allowing users to practice specific maneuvers, encounter challenging driving conditions, and develop critical decision-making skills.

Driving simulators fall into two main categories: static simulators and motion simulators (see Figure 4.8).

- **Static Simulators:** These are stationary setups that replicate driving experiences without physical motion, consisting of a driving seat, steering wheel, pedals, and a visual display.
- **Motion Simulators:** These use hydraulic systems to simulate physical movements and forces, enhancing immersion and realism by reproducing sensations like inertia and G-forces.

4.3.2 Simulation Design and Architecture

This thesis aimed to enhance the user experience and mitigate cybersickness during navigation in a virtual environment by employing a human-centered intelligent car (avatar). The focus was specifically on addressing the visual-vestibular sensory conflict, which significantly contributes to cybersickness. Therefore, the simulation design did not encompass sound and motion systems. Figure 4.9 showcases the designed driving simulator architecture.

Figure 4.8: Types of driving simulator (a) Static Simulator: A stationary setup replicating a driving experience without physical motion. (b) Motion Simulator with Hydraulic Systems: Incorporates hydraulic systems to simulate movements and forces encountered during driving, enhancing realism.

Driver input is employed to compute the vehicle dynamics through the vehicle model. Subsequently, the feedback systems utilize this model to give the driver the essential cues for an immersive driving experience. By incorporating parameters such as steering angle and brake input, the car model simulates the behavior and response of the virtual vehicle. The scenario control module utilizes the defined environment with the vehicle dynamics. This integration enables the generation and output of visual cues that reflect the virtual driving scenario, providing the driver with a realistic and immersive experience. Subsequent sections will provide a comprehensive overview of the hardware and software components involved in the designed driving simulator system.

4.3.2.1 Simulation hardware

To create the cockpit mockup, we utilized a non-swiveling chair and the Logitech G25 racing wheel as driving tools, along with the Meta Quest Pro HMD as the visual display system, as demonstrated in Figure 4.10. The Logitech G25 is widely recognized for its ability to deliver a more authentic and immersive driving experience than conventional gamepad or keyboard controls.

We enhanced the driving simulator's hardware setup by integrating the Logitech G25 racing wheel with an HMD.When coupled with the Meta Quest Pro HMD, users benefit from an expanded field of view, eye and head tracking, and immersive visuals, effectively bringing the virtual driving environment

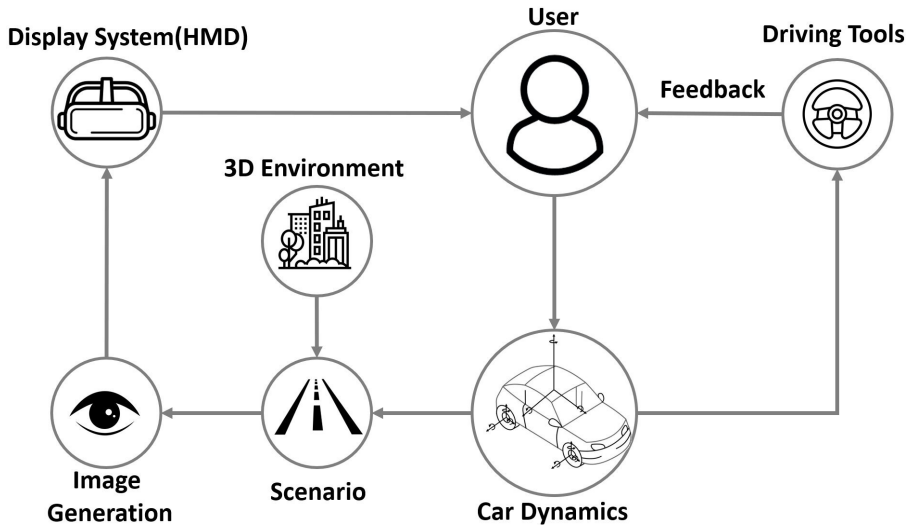


Figure 4.9: Diagram of a driving simulator architecture.

to life. The integration of these components ensures a more realistic and captivating driving simulation.

It is important to highlight that the hardware setup for a driving simulator can vary significantly based on specific requirements and budget constraints. In the case of this setup, we carefully considered the unique features of the Logitech G25 racing wheel and the Meta Quest Pro HMD and how they align with the thesis objectives before making our choice.

4.3.2.2 Physics Simulation

I. Suspension and Tires: The car consists of a rigid body with multiple wheels, typically four, two of which are powered by the engine. The wheels use a simplified suspension system resembling a spring and a damper, and the force exerted by the wheel colliders depends on the suspension configuration.

Tire friction is defined for each wheel, and Unity3D employs a two-spline curve to calculate the force at the wheel-road contact point, as shown in Figure 4.11. This force is determined by the tire slip, measuring tire movement relative to the road. When the tire has full traction, the slip is zero, but it increases during tire sliding, such as during emergency braking. The wheel friction spline is defined by two points indicating maximum slip and corresponding force, along with an asymptote representing maximum force. Evaluating the spline with the tire slip value determines the force on the wheel. It should be noted that these values are approximations derived from testing behavior during acceleration, braking, and high-speed turns rather than specific tire specifications.

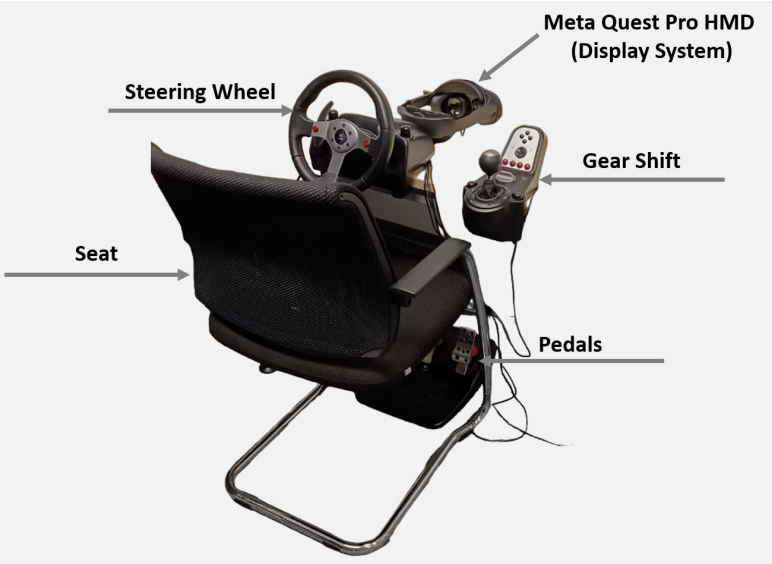


Figure 4.10: Driving simulator showcasing a non-swiveling chair, Logitech G25 steering wheel as driving tools, and Meta Quest Pro serving as the display system.

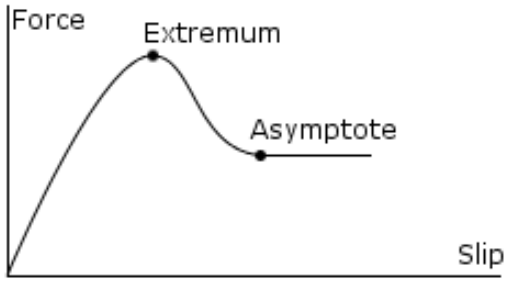


Figure 4.11: The WheelFrictionCurve in Unity3D is utilized by the WheelCollider to define the friction properties of the wheel tire.

II. Engine and Transmission: The simulated car has a combustion engine, resulting in a non-linear relationship between the torque generated and the engine speed. Typically, this relationship is quantified in terms of revolutions per minute (RPM). When the accelerator pedal is depressed, the engine's output force is contingent upon its current velocity.

In this system, the engine does not directly power the wheels. Instead, they are linked to a gearbox, which converts the engine's rotational speed into a distinct output speed based on the transmission ratio of the selected gear. This arrangement allows the driver's input to influence the engine, gearbox, and, ultimately, the wheels. The interconnectedness of these components is illustrated in Figure 4.12, demonstrating the relationship and control system between the engine, gearbox, and wheels within

this setup.

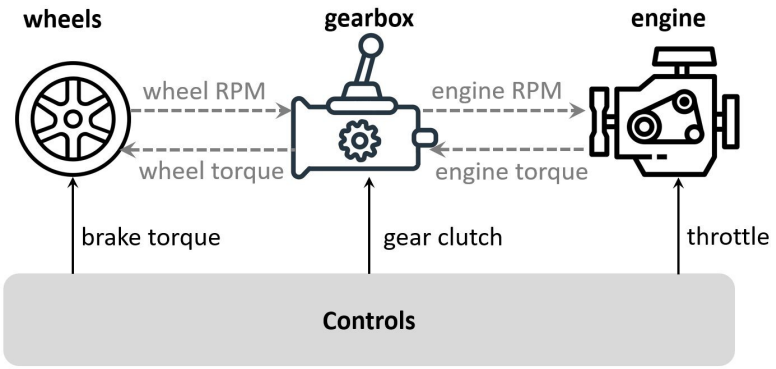


Figure 4.12: Components of the transmission model.

The idle RPM and redline RPM parameters are commonly used to define the operating range of an engine. In the given simulator's engine and transmission system, the idle RPM is set to 800, representing the minimum speed at which the engine can operate smoothly. The redline RPM is configured as 6500, indicating the maximum RPM limit beyond which it is not recommended to operate the engine to avoid potential damage.

The differential ratio parameter is critical to a car's drivetrain system. In this simulator, the differential ratio is configured as 4. The differential distributes torque between the vehicle's left and right wheels. A differential ratio of 4 is pivotal in determining how the torque is divided and transmitted to the individual wheels. This, in turn, significantly enhances the vehicle's handling and traction characteristics, making it a parameter of utmost importance.

The motor power parameter, set to 100 horsepower in the simulator, also represents the engine's power output.

These parameter settings significantly shape the engine's behavior and the drivetrain system's performance within the simulator. It's important to understand that they don't work in isolation but collectively contribute to determining the engine's operating range and torque distribution. This interconnectedness is what makes the simulation of engine performance realistic and accurate.

An HP-RPM curve was defined to calculate torque accurately, representing the relationship between the engine's horsepower output and its RPM. Figure 4.13 demonstrates the curve implemented in Unity3D for our simulator. This curve plays a crucial role in capturing the engine's power delivery across various RPM levels, serving as a fundamental characteristic of the engine. By evaluating the

curve at specific RPM ratios, the simulator can precisely determine the corresponding horsepower output of the engine, facilitating accurate torque calculations and realistic simulation of engine performance.

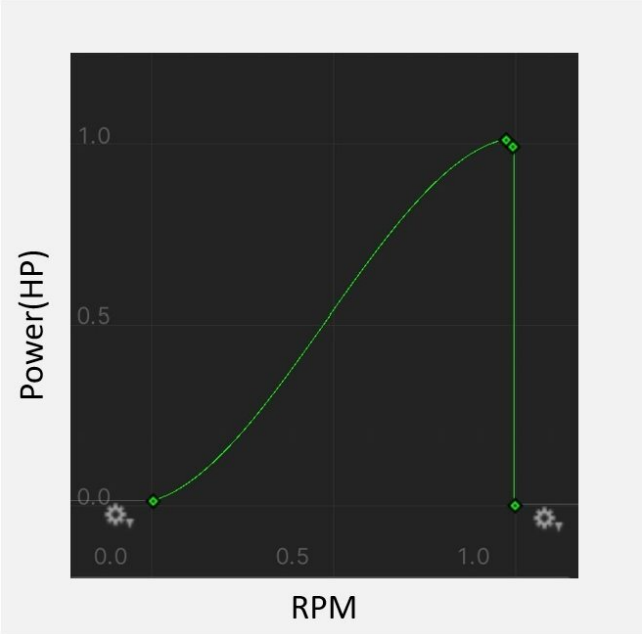


Figure 4.13: Power-RPM curve of the developed simulator's engine defined in Unity3D.

The torque is calculated using the following formula:

$$\begin{aligned}\text{torque } = &(hpToRPMCurve.Evaluate(RPM/\text{ redLine }) * \text{ motorPower } /RPM) * \text{ currentGear} \\ &* \text{ differentialRatio } * 5252\text{f} * \text{ clutch}\end{aligned} \tag{4.7}$$

The constant value 5252 is used for unit conversion from foot-pounds per minute (ft-lb/min) to horsepower (hp). This conversion factor is commonly employed in the context of power calculations in automotive engineering.

III. Steering Angle Control: To calculate the vehicle steering angle, a non-linear downward curve (see Figure 4.14) was defined to adjust the steering angle based on the vehicle's speed. This curve aims to provide a smoother steering response, particularly during high-speed cornering scenarios. As the vehicle's speed increases, the curve gradually reduces the steering angle adjustment, resulting in less aggressive steering input at higher speeds.

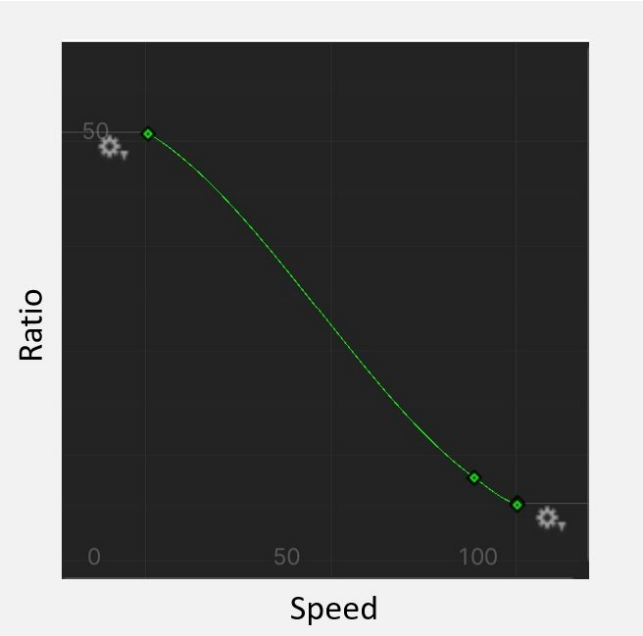


Figure 4.14: Steering ratio curve defined in Unity3D.

The steering angle is calculated using the formula:

$$\text{currentSteerAngle} = \text{steer} * \text{steeringCurve.Evaluate (speed)} \tag{4.8}$$

The *steeringCurve.Evaluate(speed)* function obtains a value from the steering curve based on the vehicle's current speed. In VR, when users encounter sudden or exaggerated steering inputs, particularly during high-speed maneuvers or cornering, the visual cues and the corresponding rotation of the virtual environment can create a more intense sensation of self-rotation. The combination of the visual scene rotating and the perceived G-forces due to the acceleration and centrifugal forces can enhance the perception of circular vection for some drivers.

Circular vection can be created by rotating around the yaw axis. The yaw axis is the vertical axis around which a person or object rotates horizontally. When a visual scene rotates around this axis, it can induce the sensation of self-rotation in the opposite direction. Circular vection can create a sensory mismatch between the visual cues and the user's vestibular system. This discrepancy can increase the likelihood of cybersickness and discomfort [162]. The steering angle is controlled using the described approach to ensure a more realistic driving experience and minimize intentional factors that may induce cybersickness.

4.3.2.3 Virtual Environment and 3D Content

I. Car and Cockpit Model: The Azerilo Model No. 1203 Unity3D asset was utilized for the 3D model of the car, offering several advantages. Specifically chosen for its optimized car body collider, accurate collision detection was ensured by this model. Moreover, the modularity of this model allowed for individual manipulation of various body parts, such as separate wheels and steering wheels. These features enabled a more realistic and customizable simulation of the car within the virtual environment to be achieved. Figure 4.15 illustrates the driver's perspective within the car's cockpit.

Figure 4.15: Interior view of the car. RPM value and gear shift indicator are displayed on the vehicle dashboard.

II. City Scene: A large-scale virtual environment was created using the Urban Construction Pack Unity3D asset. The virtual city encompassed buildings, highways, and sidewalks. Two-lane highways were employed for ease of use. Additionally, sidewalks were incorporated along both sides of the road, designed to mimic real streets with rounded corners to facilitate smooth turning at road bends.

The city scene was thoughtfully crafted as a loop to ensure control and consistency, excluding intersections or side streets. This loop design ensured all participants followed the same path throughout the simulation. The loop consisted of 8 road bends, inducing participants to experience rotational and lateral movements. Figure 4.16 displays the entire city, providing a comprehensive layout. Furthermore, Figure 4.17 presents a frame capturing the city scene and car from a third-person perspective.

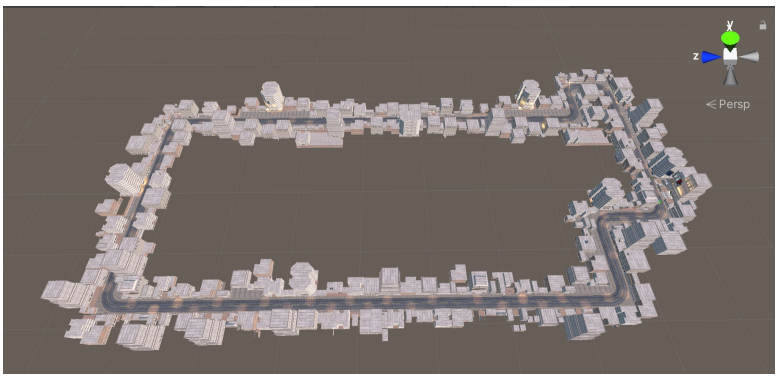


Figure 4.16: The city design showcasing 8896m of roads, sidewalks, and buildings. The navigation path encompasses a combination of rotational and lateral movements.

Figure 4.17: A frame depicting the city scene and car from a third-person perspective.

4.3.3 Multi-sensory Perception in Driving Simulators

Understanding the sensory systems that come into play is crucial for designing an effective driving simulator to create an authentic experience and effectively deceive human perception.

The visual system plays a vital role in driving simulation since driving is primarily a visual task. It accounts for a significant portion of motion perception in a three-dimensional environment [257]. While the visual system is the main player in driving simulation, the auditory system should not be overlooked. It also contributes to self-motion perception. A study [108] found that combining auditory and visual cues can enhance the illusion of motion. This is particularly important in static simulators that lack physical motion, as they rely on visual and auditory cues to create the illusion of motion and simulate the driving environment, offering a realistic driving experience through immersive visuals

and sound.

The somatosensory system plays a crucial role in motion simulators alongside the visual and auditory systems. It is responsible for tactile perceptions and proprioceptive sensors [249], which can detect force changes, position perception, and accelerations.

Research has indicated that drivers utilize the vestibular system to determine steering angles and adjust their behavior accordingly. The vestibular system recognizes linear and angular motion [150]. For example, a study in [141] demonstrated how the visual-vestibular system is modeled to understand how drivers utilize the vestibular system for steering angle determination. This work provides evidence that drivers adapt their behavior based on the cues provided through the vestibular system. Figure 4.18 summarizes the sensory systems involved in driving simulator experiences.

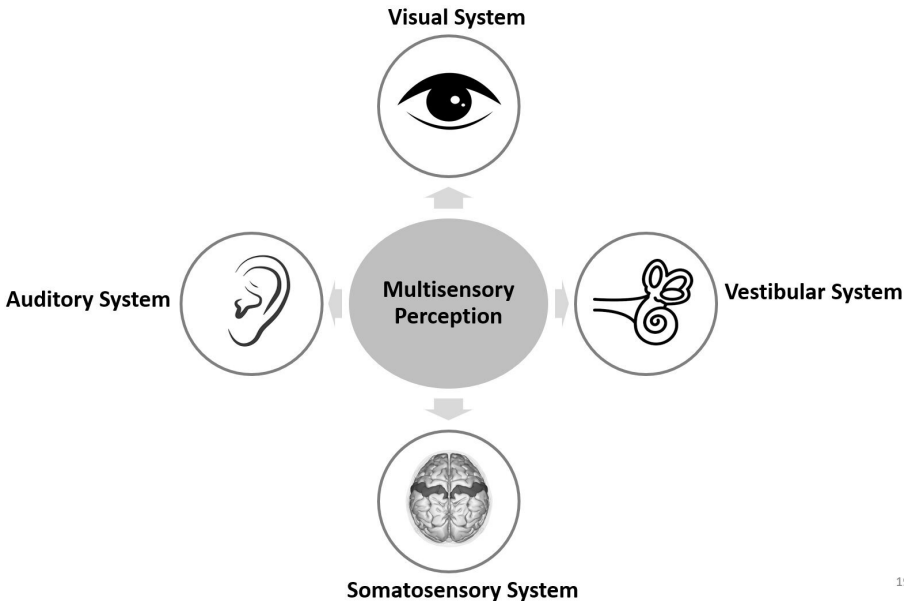


Figure 4.18: Human Perceptual System In Driving Simulator.

However, despite the efforts to design driving simulators to mimic real-world driving experiences, users often suffer from an inconvenient state known as cybersickness. This phenomenon raises important questions regarding its causes, symptoms, and impact on the overall simulation experience.

4.4 Cybersickness: The Selected Concern for Auto-Adaptive Intelligent Avatar

Cybersickness, commonly referred to as VR sickness, is a phenomenon that can occur when using VR or immersive technologies, particularly during navigation. It is characterized by symptoms similar to motion sickness, including nausea, dizziness, disorientation, headache, sweating, and fatigue. There is a notable correlation between the discomfort levels of VR sickness and user comfort [218]. Despite the development of numerous navigation interfaces to improve user comfort, cybersickness remains an inherent challenge that needs to be addressed in VR experiences.

4.4.1 Cybersickness: Understanding the Causes

The scientific community has been searching for theories to elucidate the causes and mechanisms behind cybersickness. However, due to the complex nature of this phenomenon, no single unified theory can comprehensively explain all aspects. Instead, researchers have developed and refined multiple theories that complement each other and are applicable in different contexts. In this section, we summarize the most famous theory related to cybersickness.

4.4.1.1 Sensory Conflict Theory

The sensory conflict theory, a cornerstone in the study of cybersickness, was first proposed by the pioneering work of Reason and Brand [184]. Their theory, which suggests that cybersickness results from conflicting information from different sensory organs, has significantly advanced our understanding of this phenomenon. They identified two primary sources of conflict: the disparity between visual motion conveyed through the optic nerve and the actual physical motion in the real world, and the conflict between the vestibular system's detection of orientation and acceleration [184][32].

As detailed in [248], the vestibular system in the inner ear is a complex sensory system consisting of semi-circular canals and otolith organs (see Figure 4.19). These components enable the recognition of motion. The semi-circular canals are three circular cavities filled with a fluid called endolymph. When the individual experiences angular acceleration, such as nodding or tilting their head, the hair lining the canals is deflected, generating an output proportional to the angular velocity. This information is then transmitted to the central nervous system.

In addition to the semi-circular canals, the vestibular system includes the otolith organs, comprised

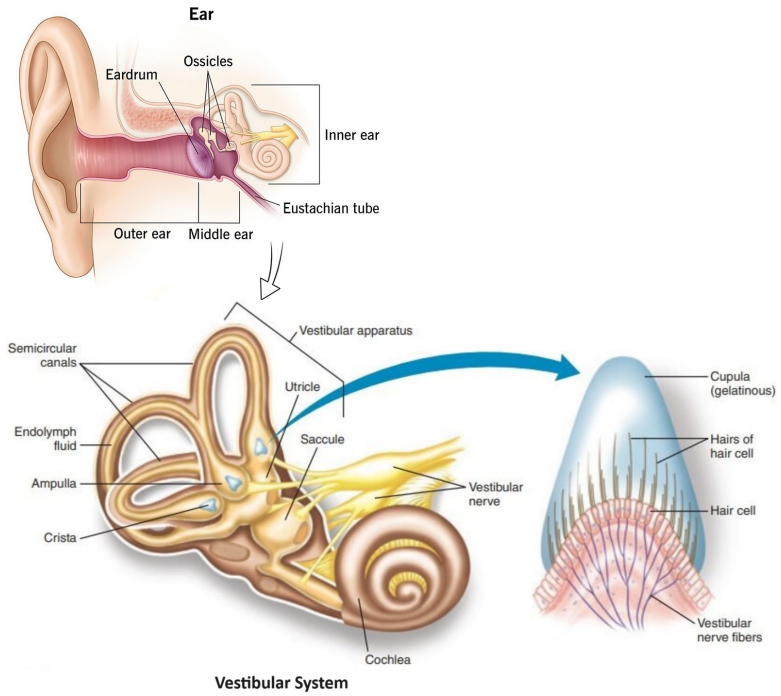


Figure 4.19: Human Vestibular System

of tiny sacs containing sensitive hairs. These hairs play a crucial role in sensing linear acceleration and gravitational forces. The utricle detects motion in the horizontal plane, while the saccule detects motion in the vertical plane. The semi-circular canals and otolith organs contribute to the perception of motion and spatial orientation, providing crucial sensory input for maintaining balance and coordinating body movements.

Additionally, Reason [183] developed a neural mismatch model, suggesting that cybersickness symptoms arise when sensory information contradicts an individual's past experiences, and the inconsistency persists, leading to a constant conflict between the sensory input and expectations.

Thereby, cybersickness can arise due to a range of conflicts encompassing the following:

1. Conflicts among sensory modalities (i.e. visual-vestibular)

For example, when a person is immersed in a moving 3D virtual environment, they receive visual cues indicating self-motion. At the same time, their vestibular system reports no self-motion since they are actually in a stationary position.

2. Conflicts within a sensory system

One notable example is the mismatch between naturally correlated motor and retinal cues to motion-in-depth in 3D virtual environments. In the physical world, accommodation cues (related to the focusing of the eyes) and disparity cues (related to the differences in the retinal images received by each eye) typically provide congruent information. When an object moves closer to a person in reality, the retinal image becomes blurred, and the disparity between the signals received by the two eyes adjusts accordingly.

However, in 3D immersive viewing, these two depth cues often conflict with each other. In head-mounted devices, the disparity cues in a 3D stereoscopic environment suggest that an object is approaching, while the focus of the retinal image relies on the fixed distance between the eye and the VR display. As a result, the cues fail to reflect variations in depth, leading to a perceptual conflict.

3. Conflicts between expected and experienced patterns

Stoffregen and Riccio [223] argued that the concept of sensory conflict relies on the presence of mediating assumptions or expectations by the individual.

Individual differences in sensory system sensitivity and adaptation play a significant role in the experience and severity of cybersickness. While sensory conflict is a prominent theory to explain cybersickness, it is important to recognize that individuals have varying degrees of susceptibility and adaptability to these conflicts. Chung and Barnett-Cowan [45] examined the relationship between susceptibility to cybersickness and sensory reweighting in VR. Participants performed a subjective visual vertical (SVV) task before and after exposure to low- and high-intensity VR games. The results showed that after the high-intensity VR experience, participants' perception of verticality shifted towards the vestibular signal. Cybersickness severity, measured using the fast motion sickness (FMS) scale, was higher after the high-intensity VR game. The change in SVV accounted for nearly half of the variance in FMS ratings, suggesting a connection between SVV shift and cybersickness severity. These findings highlight the impact of VR on sensory perception and suggest a relationship between susceptibility to cybersickness and sensory reweighting.

Regarding the driving simulator, when the visual cues presented in the simulator do not align accurately with the motion perceived by the vestibular system, it can lead to a sensory conflict.

For example, if the visual display suggests acceleration or turning, but the vestibular system detects no corresponding physical motion, this mismatch can give rise to sensations of discomfort, nausea, dizziness, and disorientation.

In addition to the visual-vestibular mismatch, inconsistent or delayed visual feedback can contribute to cybersickness. A noticeable lag or latency in the visual display can disrupt the user's perception of motion and spatial awareness. This delay can exacerbate the sensory conflict and increase the likelihood of experiencing cybersickness.

Furthermore, factors such as the intensity and frequency of motion cues, the limitations of the field of view, and the level of immersion in the simulation environment can influence the occurrence and severity of cybersickness. Similarly, if the field of view in the visual display is limited or constrained, it can create an unnatural visual experience that contributes to the onset of cybersickness.

4.4.2 Individual Factors and Their Impact on Cybersickness

Individual characteristics play a significant role in the experience of cybersickness during navigation in virtual environment. Besides technological factors related to the virtual environment, understanding the human-related characteristics is equally crucial in identifying the primary reasons behind its occurrence.

By examining individual characteristics, researchers can gain insights into the factors contributing to cybersickness susceptibility and severity.

1. Age: The impact of age on cybersickness is viewed from opposing perspectives. First, Reason and Brand [184] proposed that children's developing and adaptable sensory systems contribute to their lower susceptibility to cybersickness. Second, in older individuals, particularly those over 60 years old, the aging of the vestibular system may reduce cybersickness due to the loss of vestibular nerve fibers [21] and the aging of vestibular ganglion cells [185]. Theories such as the poisoning theory [235] and negative reinforcement [29] offer additional explanations for the age-related effects on cybersickness.

However, contrasting findings emerge. Arns and Cerney [14] conducted a study using a similar but slightly different age grouping method, incorporating participants aged 9 to 60. Their research revealed that adults aged 30 years and older, especially the elderly, exhibited greater susceptibility to simulator sickness compared to younger age groups. The severity of cybersickness symptoms among

older subjects was notable compared to the overall severity scores.

2. Gender: Gender appears to have a modest but significant effect on susceptibility to cybersickness, with women generally experiencing more severe symptoms compared to men. However, the research findings are somewhat inconsistent and influenced by various factors:

Several studies have reported that women tend to experience greater cybersickness, particularly nausea-related symptoms, compared to men [232]. This has been attributed to increased postural instability, interpupillary distance (IPD) differences, and higher susceptibility to motion sickness in women [232].

Other studies, however, have found no significant gender differences in cybersickness susceptibility [232] [117]. These discrepancies may arise from methodological issues like small sample sizes, sample gender imbalance, lack of pre-testing for individual susceptibility, and differences in virtual content/tasks.

Recent reviews highlight the need to understand better the complex interplay between gender, sensory cues, individual susceptibility, postural instability, and task parameters to develop effective strategies for mitigating cybersickness. More comprehensive studies are therefore needed to clarify the underlying mechanisms and develop personalized adaptation approaches.

3. Previous Experience: Research findings on this topic have been mixed. Some studies suggest that individuals with extensive gaming experience exhibit a reduced susceptibility to cybersickness due to their familiarity with navigating virtual environments and their developed tolerance to motion stimuli [95] [117].

Other research suggests that the relationship between gaming experience and cybersickness is not statistically significant, implying that susceptibility to cybersickness may be influenced more by individual factors, such as vestibular sensitivity, rather than gaming experience alone. An online survey by Rangelova et al. [180] revealed that gamers can be vulnerable to cybersickness, exhibiting symptoms such as nausea, fatigue, and discomfort during or after VR games. The aforementioned study emphasized this finding.

4. Postural stability: Postural stability refers to maintaining balance and stability while standing or performing other tasks. It is a broad concept encompassing various factors such as body posture control, coordination, and the ability to resist perturbations. Postural stability is typically assessed

by comparing measurements taken before and after exposure to the virtual environment. However, the relationship between postural stability and cybersickness is multifaceted and needs to be fully understood. Some studies stated that individual differences in postural stability can predict which people are more likely to experience cybersickness [224] [10].

5. Individual Differences: Factors such as motion sickness history [75], anxiety levels [129], visual acuity [246], headache, and migraines [217] can also contribute to cybersickness susceptibility. For example, individuals with a history of motion sickness or anxiety disorders may be more prone to experiencing cybersickness.

4.4.3 Cybersickness Evaluation

Researchers and developers depend on robust evaluation methods to effectively address cybersickness and develop effective mitigation strategies to assess its occurrence, severity, and potential influencing factors. These evaluation methods can be broadly categorized into subjective and objective methods [161]. In this section, we will explore both of these approaches in detail.

I. Subjective Evaluation Subjective valuation methods capture the individual's subjective experience and perception of symptoms. Questionnaires and rating scales commonly employ subjective measures to assess cybersickness, where participants complete surveys before and after engaging in VR tasks. These surveys include widely used questionnaires such as the Motion Sickness Questionnaire (MSQ) [68], Simulator Sickness Questionnaire (SSQ) [105], Fast Motion Sickness Scale (FMS) [107], and VR Sickness Questionnaire (VRSQ) [110].

By utilizing these questionnaires and surveys, researchers can gain insights into participants' subjective experiences of cybersickness and assess the severity and frequency of symptoms. These subjective evaluation methods provide valuable information for understanding the impact of VR on individuals and can aid in developing strategies to mitigate cybersickness. However, it is essential to note that subjective evaluation methods have limitations since they rely on a posteriori feedback, meaning that participants provide feedback after experiencing cybersickness symptoms. This limitation prevents immediate and real-time action to efficiently restrict or mitigate cybersickness during VR experiences.

II. Objective Evaluation Objective evaluation of cybersickness relies on physiological and behavioral measurements, complementing the subjective data obtained from questionnaires. This scientific

rigor ensures a comprehensive understanding of the condition.

One commonly used objective measure is galvanic skin response (GSR), also known as electrodermal activity (EDA), which measures the skin's electrical conductivity. Plouzeau et al. [173] have employed GSR as a physiological indicator to assess the autonomic nervous system response during VR experiences. Additionally, electroencephalography (EEG) records and analyzes brainwave patterns associated with cybersickness. Studies by Kim et al. [111], Jeong et al. [97], Liao et al. [126], and Lin et al. [127] have utilized EEG to understand the neural correlates of cybersickness and identify specific brainwave patterns associated with the condition.

Furthermore, electrocardiogram (ECG) measurements are utilized to assess the impact of VR on heart rate and cardiovascular activity. Garcia-Agundez et al. [73] have used ECG to investigate the physiological responses associated with cybersickness. By monitoring these physiological indicators and recording instantaneous signals during VR exposure, researchers can analyze and process the data to identify the extent of cybersickness experienced by participants and understand the impact of the VR task on their physiological responses.

Behavioral indicators such as postural sway, which reflect changes in body posture and balance, have emerged as crucial indicators of cybersickness. Understanding the impact of VR on balance and stability, as demonstrated in studies by Chardonnet et al. [38], is key to addressing cybersickness.

Integrating subjective and objective evaluation methods allows for a more comprehensive understanding of cybersickness, facilitating the development of effective mitigation strategies and improving VR technologies.

4.4.4 Strategies to Alleviate Cybersickness

Various strategies have been explored to mitigate cybersickness. Designing the virtual environment and visual display system to minimize potential triggers for cybersickness has been investigated extensively. Techniques such as reducing latency [220][221], increasing display refresh rates [6], minimizing visual conflicts [64], and simplifying virtual scenes to lower visually induced self-motion perception (optic flow) [133] can contribute to a more comfortable and immersive VR experience.

Limiting the field of view (FOV) presented to users, achieved through techniques like constraining peripheral vision or employing dynamic FOV reduction methods, has shown promise in alleviating

cybersickness symptoms [231]. By reducing the visual stimulus that can cause sensory conflicts, this approach could decrease the likelihood and severity of cybersickness.

In addition, implementing smooth and natural locomotion techniques in VR, such as teleportation or redirected walking [104], can effectively reduce the incidence of cybersickness. These techniques aim to minimize the visual-vestibular conflicts that often contribute to discomfort and disorientation during virtual movement.

Furthermore, decreasing the speed or acceleration in VR experiences is an effective method for mitigating cybersickness [123]. High-speed or abrupt changes in motion can increase the likelihood of experiencing cybersickness symptoms by inducing sensory conflicts. By reducing the speed of movements or gradually accelerating/decelerating, the sensory stimuli that can trigger discomfort and disorientation are minimized. It is important to note that individual susceptibility to cybersickness may vary, and the effectiveness of these strategies can differ from person to person. Therefore, finding the right balance of speed and acceleration that minimizes cybersickness while maintaining the desired user experience is a crucial consideration during the development of VR applications.

4.5 Multi-modal Measurement and Speech-Driven Self-Reporting

As discussed in section 2.2, the auto-adaptive intelligent avatar should account for the user's state. Physiological and behavioral measurements linked to cybersickness were collected in the driving simulation use case. Additionally, the user's self-reported state based on the FMS score was verbally collected as the adaptive variable, given the avatar's key adaptability feature. The following sections will detail the types of measurements collected and how they were received.

4.5.1 Behavioral Measurements

One key approach to addressing cybersickness is the use of behavioral measurements, which can provide valuable insights into the physiological and cognitive changes associated with this condition. Behavioral data, captured through techniques like eye tracking and head tracking, offers a noninvasive and real-time way to monitor the user's responses within the virtual environment.

4.5.1.1 Eye Tracking Data Capture

As mentioned in section 4.4.1.1, the conflict between the visual information presented in the virtual environment and the user's vestibular system is a primary cause of cybersickness. This sensory mismatch can disrupt the normal functioning of the eyes, which plays a crucial role in the user's spatial orientation and perception of self-motion.

The Simulator Sickness Questionnaire (SSQ), a widely used tool for assessing cybersickness, significantly emphasizes ocular symptoms. A substantial portion of the SSQ questions are dedicated to eye-related symptoms, such as eye strain, blurred vision, and difficulty focusing, further highlighting the close connection between eye function and the experience of cybersickness.

By closely monitoring the eye's behavioral measurements, more profound insights can be gained into the physiological and cognitive mechanisms underlying cybersickness.

I. Eye Movement In this work, we focused on the key values measured through eye tracking to assess cybersickness:

- *Eye Velocity:* The speed of eye movements, important for understanding how quickly users can shift their gaze between points of interest.
- *Eye Angular Velocity:* This measures the rate of change in the angle of the eye during movements, providing insights into the dynamics of eye motion.
- *Eye Position:* Refers to the location of the eyeball within the socket, crucial for tracking visual targets and maintaining focus.
- *Eye Rotation:* The angular movement of the eye around three spatial axes—horizontal, vertical, and torsional—essential for assessing how users orient their gaze in response to stimuli.

4.5.1.2 Head Tracking Data Capture

Within the broader spectrum of behavioral measurements, head tracking data, which captures the user's spatial orientation and balance, can further complement the information provided by eye tracking.

Sudden changes in head position and rotation can indicate a user's struggle to maintain stable self-perception and spatial awareness, which are critical factors in developing cybersickness [193]. Head movements involve proprioceptive feedback, which provides information about the position and movement of the body. Inconsistent or delayed feedback in VR can contribute to the sensory conflict, exacerbating cybersickness.

I. Head Movement: Different types of head movements can have varying impacts on cybersickness:

- *Rotational Head Movements:* The head's rotational movements along the three primary axes are shown in Figure 4.20. These rotational degrees of freedom include:
    - *Pitch:* Flexion and head extension, enabling nodding motions.
    - *Roll:* Lateral flexion of the head, allowing the head to tilt from side to side.
    - *Yaw:* Axial head rotation facilitates turning or scanning movements.

  The ability to quickly rotate the head through these three degrees of freedom can create more significant sensory mismatches between the visual information in the virtual environment and the vestibular/proprioceptive cues from the physical head movements. This sensory conflict is a critical factor in the development of cybersickness symptoms, such as nausea, dizziness, and disorientation.

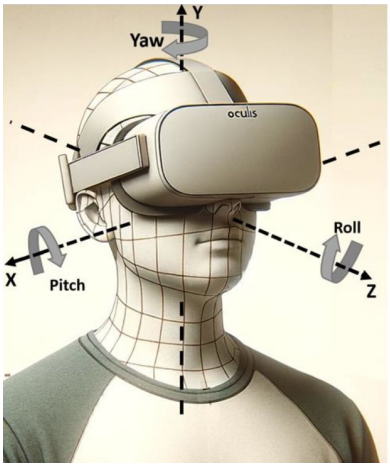


Figure 4.20: Axes of the head movements [193].

- *Translational Head Movements:* Movements that involve shifting the head position (e.g., moving forward, backward, or side-to-side) can also contribute to cybersickness, especially if the virtual environment does not accurately reflect these movements.

II. The Meta Quest Pro's Integrated Eye Tracker: Head trackers are devices or systems that monitor head movement data in space. They are essential in VR, augmented reality (AR), gaming, and human-computer interaction applications. Here are the main types of head trackers:

- *Optical Head Trackers:* These kinds of trackers use cameras to track the position and orientation of the head by detecting markers or features on the user's head or headset. They can be :
    - *Outside-In Tracking:* (cameras in the environment tracking markers on the headset). Examples include the HTC Vive and Oculus Rift.
    - *Inside-Out Tracking:* (cameras on the headset tracking the environment). Examples include the Oculus Quest and Microsoft HoloLens.
- *Inertial Head Trackers:* These use inertial measurement units (IMUs), including accelerometers, gyroscopes, and sometimes magnetometers, to track the head's motion and orientation.
- *Magnetic Head Trackers:* These use magnetic fields to determine the position and orientation of the head. A transmitter generates a magnetic field, and sensors on the headset detect the field's strength and direction. One example is Polhemus Liberty.
- *Acoustic (Ultrasonic) Head Trackers:* These use ultrasonic sound waves to determine the position and orientation of the head. Transmitters emit ultrasonic pulses, and receivers on the headset detect the time it takes for the pulses to reach them.
- *Hybrid Head Trackers:* These trackers combine multiple tracking technologies to leverage each one's strengths and mitigate its weaknesses. For example, combining optical and inertial tracking can provide high accuracy and reduce drift.

In this thesis, the head tracker embedded inside the Meta Quest Pro headset, an inside-out optical tracking system, has been used. The recorded head movement signals involve six degrees of freedom: head position along the X, Y, and Z axes, measured in centimeters, and head orientation along with

pitch, roll, and yaw axis. This approach provides freedom of movement and does not require external tracking infrastructure, making it a convenient solution for this immersive virtual reality application. The Meta Quest Pro headset features IR cameras for head and eye tracking, as shown in Figure 4.21.

Figure 4.21: The Meta Quest Pro headset collects behavioral measurements, including eye-tracking and head-tracking data. The Data TCP Client is a TCP client that connects to the core VR application to receive real-time behavioral and application data recordings.

Figure 4.22 illustrates the eye tracker and head tracker configuration within the application. Data TCP Client retrieves behavioral and application data from the core VR application.

4.5.2 Physiological Measurements

Physiological measurements play a crucial role in assessing and predicting cybersickness. By monitoring and analyzing physiological signals, valuable insights can be gained into the underlying mechanisms driving cybersickness, ultimately leading to more effective mitigation strategies.

The human body's physiological reactions to the sensory discrepancies encountered in virtual environments are a direct window into the neurological and biological processes. Measures such as electrodermal activity, cardiovascular responses, and postural instability can provide objective indicators of an individual's level of discomfort and susceptibility to cybersickness. For example, increases in skin conductance have been associated with the onset of cybersickness, reflecting the activation of the sympathetic nervous system in response to the mismatch between visual and vestibular cues.

Furthermore, the ability to track physiological changes in real time opens up the possibility of developing adaptive systems that can dynamically adjust the virtual experience to minimize cybersickness.

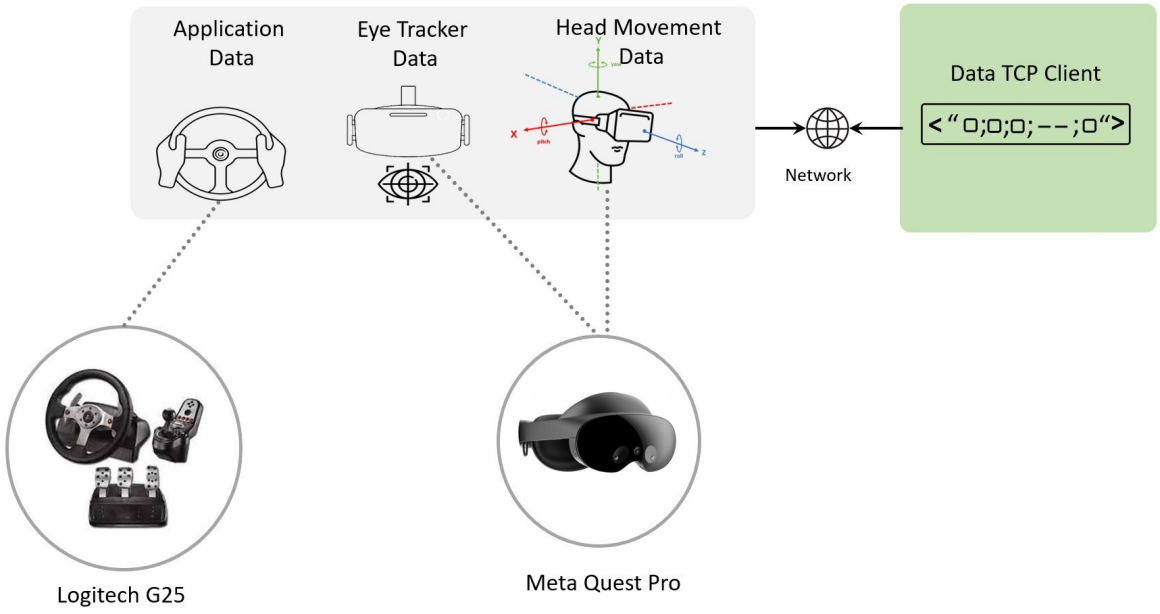


Figure 4.22: Meta Quest Pro headset.

By monitoring an individual's physiological state, VR applications could detect the early stages of cybersickness and implement mitigating strategies, such as adjusting the field of view, reducing visual flow, or speed reduction, before the user experiences significant discomfort. This personalized, data-driven approach to cybersickness management has the potential to significantly enhance the overall user experience and adoption of immersive technologies.

4.5.2.1 Physiological Measures Captured

This study leverages a comprehensive suite of physiological measurements to assess and predict cybersickness in VR system users. By monitoring multiple physiological signals concurrently, the research aims better to understand the body's responses to immersive virtual experiences.

The primary physiological measures captured in this study include electrodermal activity (EDA), blood volume pulse (BVP), 3-axis accelerometry (ACC), and skin temperature(TMP). These modalities were selected for their established relationships with motion sickness and their potential to provide groundbreaking insights into the onset and progression of cybersickness.

I. Electrodermal Activity (EDA) or Galvanic Skin Response(GSR): One of the primary physiological measures examined in this thesis is electrodermal activity (EDA), also known as galvanic skin response

(GSR) or skin conductance. EDA reflects the activity of the sympathetic nervous system and provides insights into the autonomic responses elicited by environmental stimuli and emotional states.

The underlying mechanism of EDA is the fluctuation of the skin's electrical conductance, which is influenced by the eccrine sweat gland activity. When the sympathetic nervous system is activated, it triggers the release of sweat, leading to increased skin conductance [190][226]. EDA is often used as a measure of emotional arousal, stress, cognitive load, and attention [190][228][125]. Changes in EDA can indicate emotional conditions such as stress, excitement, and fear.

EDA signals consist of two main components: tonic (slowly varying) and phasic (rapidly varying) components [226]. The tonic component represents the overall level of arousal, while the phasic component reflects transient changes in response to specific stimuli or events. Figure 4.23 and Figure 4.24 showcase two examples of EDA signals.

Figure 4.23: Example of a low tonic EDA signal with an insignificant phasic activity (45-minute segment).

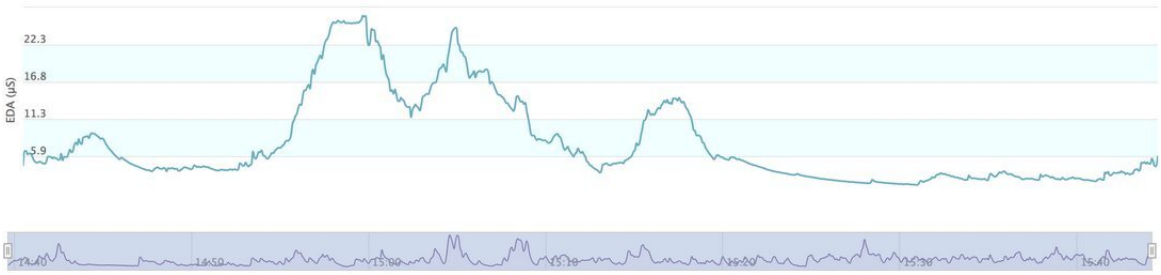


Figure 4.24: Example of a high tonic phase with significant phasic activity (1-hour long segment).

Previous research has consistently demonstrated the utility of EDA in assessing motion sickness and cybersickness [194][44][103]. Studies have shown that individuals experiencing motion sickness or cybersickness exhibit significant increases in skin conductance levels and the frequency of nonspecific skin conductance responses. Figure 4.25 demonstrates an upward trend of an EDA signal recorded

from a participant experiencing cybersickness. These EDA patterns reflect the heightened sympathetic nervous system activation associated with the distress and discomfort experienced during episodes of sensory disruption.

Figure 4.25: Example of an EDA signal recorded from a participant experiencing cybersickness.

The aim was to monitor EDA during driving, capture the autonomic nervous system's response to virtual reality exposure, and correlate it with the onset and severity of cybersickness. In conjunction with other physiological measures, EDA can contribute to a more comprehensive understanding of the multifaceted nature of cybersickness, allowing for a deeper exploration of the interplay between physiological, cognitive, and behavioral responses to immersive virtual environments.

EDA can be measured using wearable devices or sensors placed on the skin, typically on the fingers, palms, or wrists [114][125]. These sensors measure the electrical conductance or resistance of the skin. EDA signal can be analyzed in both the time and frequency domains [190][226]. Time-domain analysis focuses on the characteristics of individual EDA responses, such as amplitude, rise time, and recovery time, while frequency-domain analysis examines the spectral properties of the signal.

II. Blood Volume Pulse(BVP) Another key physiological measure captured in this study is the blood volume pulse (BVP). BVP is derived from the variations in light absorption caused by the pulsatile nature of blood flow in the microvascular bed. When the heart pumps blood, the volume of blood in the peripheral vessels changes, leading to corresponding changes in light absorption, which PPG sensors can detect [90][25][52][26]. The PPG sensor processes these volumetric changes in blood circulation using an algorithm to derive the BVP waveform. This physiological measurement provides insights into cardiovascular function and blood flow regulation, both known to be affected by motion sickness and cybersickness.

The BVP signal contains several key features that can provide valuable insights, as shown in Figure 4.26:

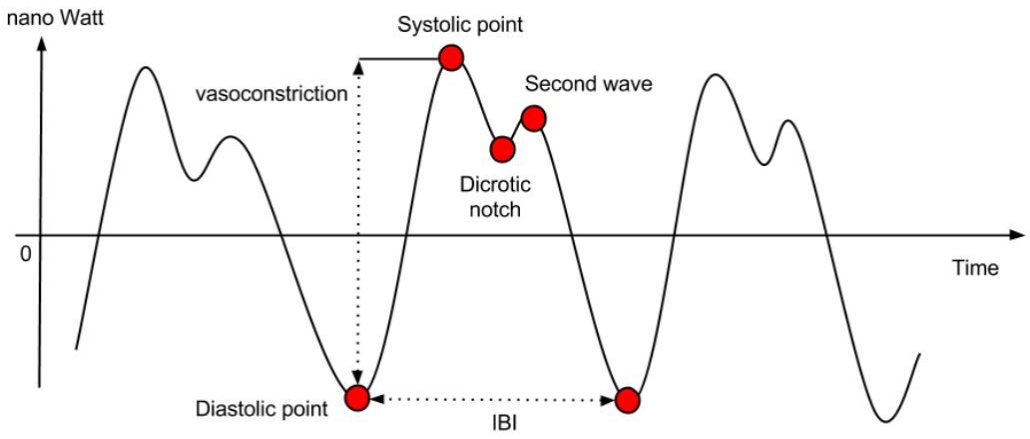


Figure 4.26: Points of interests of a BVP signal.

- *Diastolic points:* These correspond to the local minima of the PPG waveform and can be used to compute the inter-beat interval (IBI), which measures heart rate variability.
- *Systolic points:* The local maxima of the PPG waveform can be used in conjunction with the diastolic points to estimate the degree of vasoconstriction in the subject. This information can help analyze changes in cardiovascular regulation.
- *Dicrotic notch:* Under certain conditions, such as when the subject is relatively still, the dicrotic notch may be observed in the BVP waveform. This feature can study different types of cardiac diseases and cardiovascular function.
- *Dicrotic wave:* The dicrotic wave is a secondary wave that follows the primary pulse wave and is caused by the dicrotic notch. Analysis of the dicrotic wave can provide additional insights into blood flow dynamics and vascular tone.

It is important to note that the characteristics of the BVP signal can vary depending on the individual user and environmental factors. Factors such as skin pigmentation, ambient lighting, and sensor positioning can all influence the observed BVP waveform. Therefore, it is crucial to ensure proper sensor placement and environmental control to obtain reliable and consistent BVP data across participants in this study. During episodes of motion sickness or cybersickness, the body's physiological responses can alter the BVP signal. For example, research has shown that individuals experiencing motion sickness often exhibit increased heart rate, decreased heart rate variability, and changes in the amplitude and shape of the BVP waveform (see Figure 4.27).

Figure 4.27: Example of a BVP signal recorded from a participant experiencing cybersickness.

These cardiovascular changes are thought to be driven by the activation of the sympathetic nervous system, which helps the body cope with the perceived threat or distress caused by the sensory mismatch experienced in virtual environments. By monitoring the BVP signal, the dynamic changes in the cardiovascular system can be tracked and correlated with the onset and progression of cybersickness.

III. Accelerometer (ACC): A 3-axis accelerometer measures the rate of change of velocity (acceleration) along three orthogonal axes, typically labeled as the x-axis, y-axis, and z-axis. This allows it to detect movement and orientation in three-dimensional space [138][120]. These measurements can comprehensively understand an object's or person's motion and orientation in 3-dimensional space.

The three-axis configuration allows the accelerometer to measure dynamic accelerations, such as those caused by movement or vibration, and static accelerations, like the constant force of gravity. By analyzing the acceleration data across all three axes, various motion parameters can be derived, including linear acceleration, tilt, orientation, and rotation.

The raw data from a 3-axis accelerometer is a starting point, but it often requires digital signal processing to filter noise and extract meaningful information. Techniques like Fast Fourier Transform (FFT) and digital filtering play a crucial role in this process, enabling the analysis of frequency components and trends in the data [120].

3-axis accelerometers play a significant role in investigating the physiological mechanisms underlying cybersickness. Their ability to precisely measure acceleration forces across multiple dimensions makes them a valuable tool in this context.

When an individual experiences cybersickness, their body's balance and motion perception systems become disrupted due to the mismatch between the visual cues provided by the virtual environment

and the actual physical movement (or lack thereof) detected by the vestibular and proprioceptive systems. This sensory conflict can lead to autonomic nervous system responses, such as changes in heart rate, blood pressure, and sweat gland activity. Figure 4.28 presents a sample of the recorded 3-axis accelerometer signal from a participant experiencing cybersickness. Irregular movement patterns can be observed in this signal.

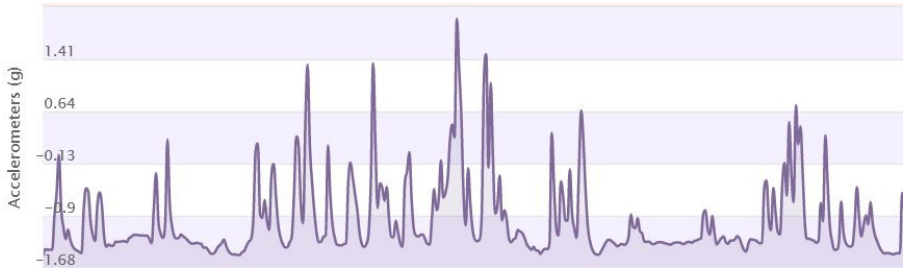


Figure 4.28: Example of an accelerometer signal recorded from a participant experiencing cybersickness.

IV. Temperature (TMP): One of the physiological responses associated with cybersickness is changes in body temperature. During episodes of cybersickness, individuals may experience thermal dysregulation, where their skin temperature fluctuates in response to the sensory conflict between the visual inputs from the virtual environment and the vestibular and proprioceptive cues from the physical world. This imbalance can lead to disturbances in the body's thermoregulatory mechanisms, primarily controlled by the hypothalamus in the brain.

Studies have shown that individuals experiencing cybersickness often exhibit increases in skin temperature [254], particularly in regions like the face and hands. This rise in temperature is believed to be a result of the body's stress response, which triggers the sympathetic nervous system to mediate increased blood flow to the skin's surface. The increased blood flow is the body's attempt to dissipate the excess heat generated by the stress response triggered by the sensory conflict.

Conversely, cybersickness may lead to changes in skin temperature, particularly in the extremities. According to Nalivaiko et al.[254], in subjects with initially higher finger temperatures (32-36°C), cybersickness either did not affect finger temperature or caused a transient decrease of 1.5-2°C. This phenomenon is associated with changes in cutaneous thermoregulatory vascular tone, suggesting a link between cybersickness and the body's thermoregulatory response.

By monitoring skin temperature changes using specialized sensors, valuable insights into the autonomic nervous system's response to cybersickness can be gained. These temperature fluctuations, serving as potential biomarkers, can be used to detect and quantify the severity of cybersickness, potentially leading to the development of more effective countermeasures and mitigation strategies.

4.5.2.2 Measurement Devices

Physiological measurement devices are crucial for assessing human responses in cybersickness research. Commonly used devices include electroencephalography (EEG) for brain activity, electromyography (EMG) for muscle tension, and skin conductance sensors for emotional responses. Heart rate monitors and thermometers also provide valuable data on cardiovascular and temperature changes. Together, these tools offer insights into the physiological effects of cybersickness on human performance.

An Empatica E4 wristband was used in this study to collect the physiological indicators. This wristband is a widely used multi-sensor device that can capture various physiological signals. The E4 wristband has several sensors, each operating at different sampling frequencies to optimize data collection.

Specifically, the wristband includes a 4 Hz EDA sensor, which measures changes in the skin's electrical properties. Additionally, it features a 64 Hz photoplethysmography (PPG) sensor that can be used to derive the BVP signal, providing insights into cardiovascular responses.

Furthermore, the E4 wristband incorporates a 4 Hz infrared thermopile (TMP) sensor, allowing for the continuous monitoring of skin temperature fluctuations, which are associated with cybersickness. Lastly, the wristband houses a 32 Hz 3-axis accelerometer sensor, enabling the tracking of subtle head and body movements that can indicate postural instability during virtual environment navigation.

The Empatica E4 wristband offers offline analysis by recording physiological data for later review and real-time streaming via Bluetooth Low-Energy (BLE). This dual functionality allows researchers to monitor physiological responses immediately. In this study, the E4's real-time capabilities were essential for evaluating the intelligent avatar's responsiveness.

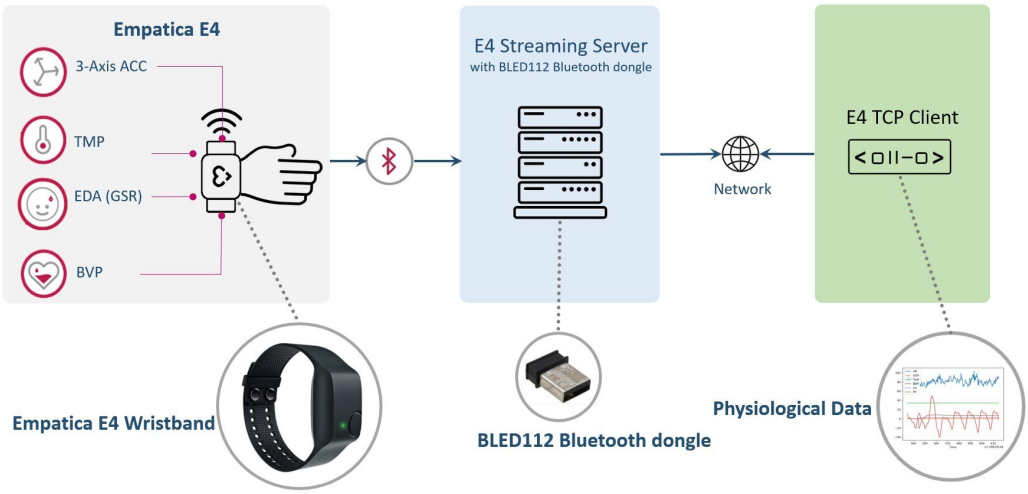


Figure 4.29: Empatica E4 configuration in real-time mode. The E4 TCP Client module was developed to receive and process the physiological data transmitted by the Empatica E4 Streaming Server in real-time.

4.5.3 Classification Indicator Powered by Reliable Speech Recognition

Automatic Speech Recognition (ASR) is a technology that utilizes sophisticated algorithms running on devices like computers or computer clusters to transform spoken language into written text [94][56]. By analyzing voice signals, ASR algorithms decipher and convert them into a coherent sequence of words or other linguistic units, enabling the accurate transcription of spoken content.

Mathematically, the goal of ASR is to accurately associate an acoustic input sequence, denoted as $X = \{x_1, x_2, \ldots, x_T\}$, of length T , with a corresponding label sequence, denoted as $L = \{l_1, l_2, \ldots, l_N\}$. The label sequence typically consists of letters or words; its length is denoted as N. Each $x_t \in \mathbb{R}^D$ represents a D-dimensional speech input vector corresponding to the t-th speech frame. The vocabulary V represents the set of labels and $l_u \in V$ denotes the label at position u in the label sequence L. We use V^* to represent the collection of all possible label sequences formed by labels from V . The primary objective of ASR is to determine the most probable label sequence $\hat{L}$, given the acoustic input sequence X. This can be expressed formally as follows:

$$\widehat{L} = \arg_{L \in \mathrm{V}^*} \max p(L \mid X) \tag{4.9}$$

Hence, the fundamental task of ASR involves developing a model capable of accurately estimating the posterior distribution $p(L|X)$.

In this thesis, one method for gathering measurements was recording the participants' voices and transcribing them to text through the ASR for further processing. This voice-based measurement was the adaptive variable. Figure 4.30 showcases the scheme of this module.

Specifically, the Fast Motion Sickness Scale (FMS) was used to assess the participants' sickness levels. During the VR simulation, the system automatically played an audio question, *"What is your score?"* every minute, and participants verbally reported their sickness level on a scale from 0 to 3. Expressly, a score of 0 indicated no sickness, a score of 1 showed initial symptoms, a score of 2 denoted moderate symptoms, and a score of 3 represented severe symptoms.

The original FMS scale, which ranged from 1 to 20, was adjusted to a more concise scale of 0 to 3 based on the trial experiment. This adjustment was made because participants needed help to report their status with high-resolution accuracy.

Because the system is real-time, obtaining the most accurate transcription of the participants' verbal responses was critical. To this end, a post-processing module was developed to refine and enhance the initial voice-to-text transcription, improving the overall accuracy and reliability of the data collected during the experiment.

In the post-processing module, the result of the speech-to-text transformation was evaluated to determine if the participant's verbal answer corresponded to a number in the valid range of 0 to 3. If this condition was not met, the time for recognition had elapsed, or if the ASR system could not reliably recognize the participant's voice, the process was repeated. The audio question was played again for the participant. This iterative process continued until the application successfully transformed the participant's verbal response into a valid numeric value on the 0-to-3 sickness scale.

4.5.4 Data Preprocessing for VR Driving Simulator

Synchronizing the heterogeneous sensor measurements and data streams is an essential general preprocessing step in this system. As discussed in section 4.5, multimodal measurements are recorded through different sensors. Collecting these multimodal measurements occurs in a separate system through various concurrent processes. The behavioral data, i.e., eye tracker and head tracker measurements, are transmitted to the simulation. Then, an independent process (i.e., the Data TCP Client) receives them over the network connection. Factors like intentional downsampling in Unity3D

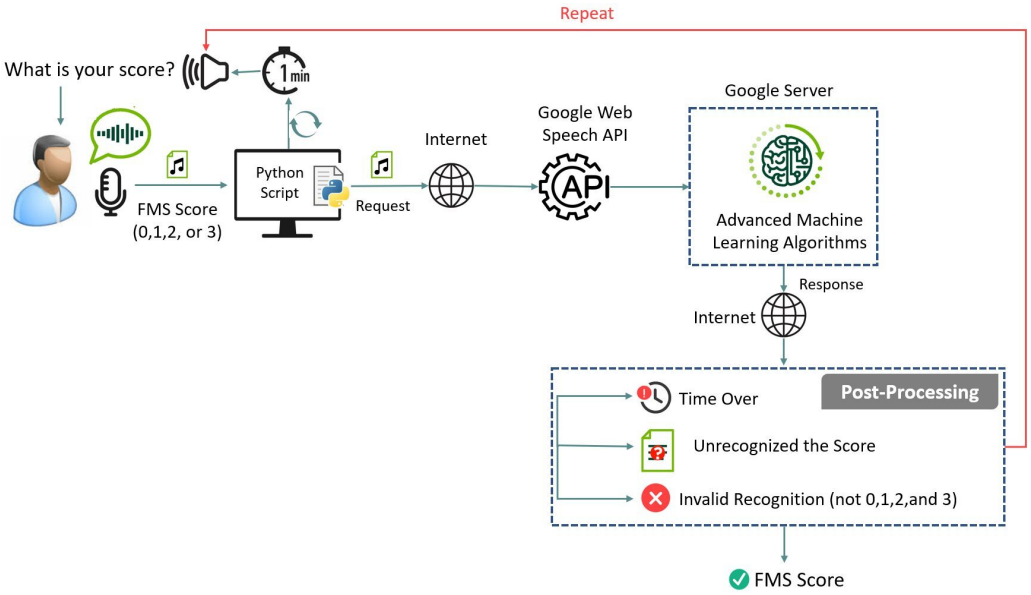


Figure 4.30: ASR module scheme. The diagram illustrates the ASR (Automatic Speech Recognition) module used in the system. Google Web Speech API was leveraged as the speech-to-text API. A post-processing module was also implemented to accurately transform the participants' verbal responses into the appropriate numeric sickness scale values.

or buffering or queuing delays can lead to a lower frequency than the fundamental frequency. The Data TCP Client receives the eye tracker and head tracker data with a frequency of 4 Hz.

Regarding the physiological data, as mentioned in section 4.5.2, the following data are transmitted to the responsible process (i.e., the E4 TCP Client):

- EDA with a frequency of 4 Hz
- BVP with a frequency of 64 Hz
- TMP with a frequency of 4 Hz
- ACC with a frequency of 32 Hz

To synchronize the representation of the multimodal data at a standard frequency of 4 Hz, a linear interpolation algorithm was implemented in the preprocessing process that should process the stream data in real-time accurately. This allowed for the alignment and resampling of the different data streams to a consistent sampling rate, facilitating the integration and analysis of the multimodal

measurements. Figure 4.31 shows a sample of the preprocessed signal resampled to a 4 Hz sampling rate.

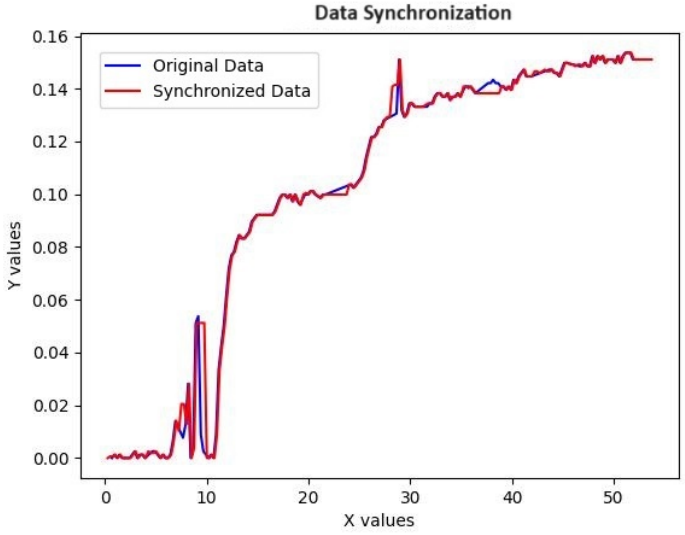


Figure 4.31: Preprocessed signal using linear interpolation to 4 Hz sampling rate.

4.5.5 Stream Learning with River: Personalized Adaptation in Driving Simulator

Data preprocessing at the level of feature engineering and transformation is a crucial step in many machine learning tasks, including stream learning, as it helps to ensure that the input data is in a suitable format for the model to learn and generalize effectively. One common preprocessing technique is feature standardization, which StandardScalar implemented in this system's AI module.

The StandardScaler transforms the input features by removing the mean and scaling the data to have unit variance. Mathematically, this is achieved by subtracting the mean of each feature and dividing by the standard deviation of that feature. This process ensures that all features are on a similar scale, with a mean of 0 and a standard deviation of 1.

$$X_i' = \frac{X_i - \mu}{\sigma} = \frac{X_i - X_{\text{mean}}}{X_{\text{std}}} \tag{4.10}$$

In the collected dataset, this transformation was applied to each feature independently. This is particularly important when working with high-dimensional data, as the features may have vastly different scales, ranges, and distributions. Standardizing the features helps ensure that no single feature dom-

inates the model's learning process and that the model can learn the underlying relationships more effectively.

Applying the StandardScaler preprocessing step ensured that the dataset was adequately conditioned for the stream learning task, which could lead to improved model performance, faster convergence, and better generalization to unseen data.

We utilized the River machine learning tool [152] to implement this system's stream learning componen. Logistic Regression was selected as the learning algorithm's base (see Figure 4.32). Logistic Regression is a well-suited choice for stream learning tasks, particularly when dealing with high-dimensional feature spaces like the one present in this dataset.

We optimized the Logistic Regression using the Stochastic Gradient Descent (SGD) algorithm for the stream learning setting, as detailed in section 4.1.3.2. The SGD optimizer was configured with a learning rate of 0.01. This combination of Logistic Regression as the base model and SGD as the optimization technique is a proven and effective approach for handling stream learning tasks, particularly in high-dimensional feature spaces.

The Adaptive Windowing (ADWIN) detection method was also used to handle the drift. The drift detector was integrated into the stream learning pipeline to ensure that the model remains adaptive and responsive to changes in the underlying data distribution.

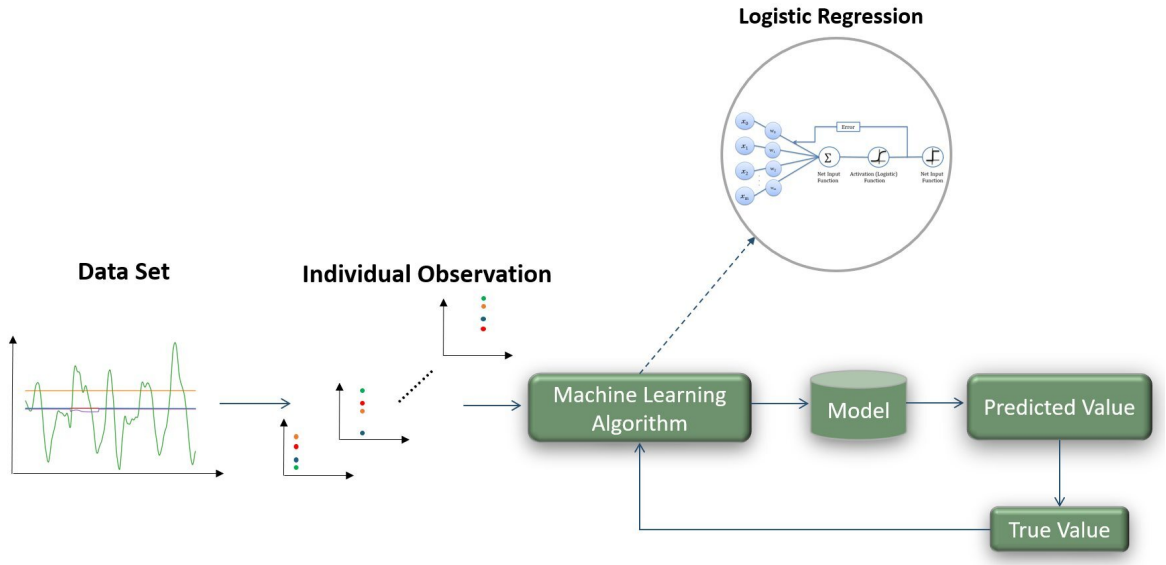


Figure 4.32: Stream learning using Logistic Regression as machine learning algorithm.

4.6 Application Performance Analysis

We conducted a comprehensive performance analysis of various performance metrics on both the client and server sides and the VR application under varying scene complexity conditions to assess the application's overall performance. For this study, we collected data over a 10-minute training session.

As illustrated in Table 4.1 and Table 4.2, we used the original/default city scene configuration for the initial analysis and measured critical metrics for the client and server processes. These included Memory Usage (MB), CPU Usage, GPU Usage, CPU temperature, and Frames Per Second (FPS). We also tracked the amount of Data Sent (MB), Data Received (MB), Packet Sent, and Packet Received to evaluate the network performance.

In the second phase of the analysis, we doubled the number of 3D objects in the city scene and repeated the performance measurements to understand the impact of increased scene complexity. Finally, we tripled the city scene complexity and reran the performance tests to assess the application's behavior under the most demanding conditions. These analyses provide invaluable insights into the scalability and optimization opportunities for our VR application across different levels of scene complexity, enhancing our collective understanding and contributing to the project's success.

The performance analysis with the original scene configuration provides a baseline understanding of the system's resource utilization and network characteristics. On the server side, the individual components, such as ASR.py, DataTCPClient.py, E4TCPClient.py, and AIModule.py, show relatively low CPU usage, ranging from 0.05% to 12%. The memory usage for these server-side components is also reasonably low, between 28 MB to 226 MB. However, on the client side, the IntelligentDS.exe application exhibits higher resource consumption, with 18.99% CPU usage and 40.57% GPU usage, resulting in a frame rate of 78 FPS. The CPU temperature for the client system is also notably higher at 64.41°C.

The network performance analysis reveals that the data transfer and packet exchange are significant, with the server components sending and receiving several gigabytes of data and millions of packets. The client-side IntelligentDS.exe application also shows substantial network activity, with over 4 GB of data sent and received and over 1.6 million packets exchanged.

Doubling the scene complexity has a noticeable impact on the system's performance. Table 4.3 and Table 4.4 represent the results of this analysis. On the server side, resource utilization remains

Table 4.1: Performance Analysis with Original Scene

	Artifact	Memory Usage (MB)	CPU Usage	GPU Usage	FPS	CPU Temperature
Server	ASR.py	28.58	0.05%	0	-	85.8°C
	DataTCPClient.py	80.35	0.22%	0	-	
	E4TCPClient.py	34.40	12.05%	0	-	
	AIModule.py	226.15	12%	0	-	
Client	IntelligentDS.exe	280.83	18.99%	40.57%	78	64.41°C

Table 4.2: Network Performance Analysis with Original Scene

	Artifact	Data Sent (MB)	Data Received (MB)	Packet Sent	Packet Received
Server	ASR.py	884.42	3073.93	3333914.73	4544544.78
	DataTCPClient.py	891.11	3108.16	3374192.09	4610123.59
	E4TCPClient.py	891.14	3107.16	3373315.29	4608864.67
	AIModule.py	888.76	3094.20	3358611.42	4585239.53
Client	IntelligentDS.exe	4925.22	6009.92	1683113.87	2621968.77

Table 4.3: Performance Analysis with 2x Scene

	Artifact	Memory Usage (MB)	CPU Usage	GPU Usage	FPS	CPU Temperature
Server	ASR.py	28.60	0.16%	0	-	87.16°C
	DataTCPClient.py	80.40	0.26%	0	-	
	E4TCPClient.py	34.47	11.77%	0	-	
	AIModule.py	227.78	12.09%	0	-	
Client	IntelligentDS.exe	305.84	22.37%	35.76%	47	66.33°C

Table 4.4: Network Performance Analysis with 2x Scene

	Artifact	Data Sent (MB)	Data Received (MB)	Packet Sent	Packet Received
Server	ASR.py	1651.55	6089.50	7356594.96	10693051.08
	DataTCPClient.py	1659.39	6126.73	399699.99	10761131.02
	E4TCPClient.py	1660.76	6132.32	7406312.53	10771441.83
	AIModule.py	1658.62	6122.33	7394765.7	10753738.17
Client	IntelligentDS.exe	5075.16	6047.09	1980382.37	2819064.71

relatively stable, with slight increases in CPU usage for some components. However, on the client side, the IntelligentDS.exe application experiences a more significant increase in CPU usage (22.37%) and a decrease in GPU usage (35.76%), resulting in a substantial drop in frame rate from 78 FPS to 47 FPS. The client system's CPU temperature also increases to 66.33°C.

The network performance analysis with the 2x scene complexity shows a significant increase in data transfer and packet exchange, doubling the values observed in the original scene. The server components now send and receive over 6 GB of data and handle over 10 million packets, while the client-side IntelligentDS.exe application exchanges nearly 6 GB of data and over 2.8 million packets.

Further increasing the scene complexity to 3 times the original has a more pronounced impact on the system's performance (see Table 4.5 and Table 4.6). On the server side, resource utilization remains relatively consistent with the 2x scenario, with only minor CPU and memory usage variations. However, the client-side IntelligentDS.exe application experiences a notable increase in CPU usage (23.68%) and a decrease in GPU usage (32.95%), leading to a further drop in frame rate to just 33

Table 4.5: Performance Analysis with 3x Scene

	Artifact	Memory Usage (MB)	CPU Usage	GPU Usage	FPS	CPU Temperature
Server	ASR.py	28.44	0.14%	0	-	87.21°C
	DataTCPClient.py	80.23	0.04%	0	-	
	E4TCPClient.py	33.98	12.03%	0	-	
	AIModule.py	227.15	12.09%	0	-	
Client	IntelligentDS.exe	334.41	23.68%	32.95%	33	66.37°C

Table 4.6: Network Performance Analysis with 3x Scene

	Artifact	Data Sent (MB)	Data Received (MB)	Packet Sent	Packet Received
Server	ASR.py	2407.98	9113.81	11314832.17	16702512.88
	DataTCPClient.py	2409.58	9121.68	11324071.24	16717145.84
	E4TCPClient.py	2415.11	9148.43	11354944.30	16766736.86
	AIModule.py	2415.23	9147.25	11354016.16	16764562.1
Client	IntelligentDS.exe	5226.06	6084	2273964.81	3007418.19

FPS. The client system's CPU temperature also rises to 66.37°C.

The network performance analysis with the 3x scene complexity shows a significant increase in data transfer and packet exchange compared to the 2x scenario. The server components now send and receive over 9 GB of data and handle over 16 million packets, while the client-side IntelligentDS.exe application exchanges nearly 6 GB of data and over 3 million packets.

In addition to the mentioned performance metrics, as part of our ongoing performance analysis efforts, we conducted a detailed examination of each component's response times during a 10-minute training session in the AI training process. For this study, we evaluated each artifact's total AI training response time and individual response times, including voice reception, data collection, physiological data acquisition, dataset creation, pre-processing, and model training. This analysis revealed that the ASR.py module is the most time-consuming aspect of the overall training pipeline. This can be attributed to the fact that the ASR.py module relies heavily on the Internet speed and the Google ASR service. The process involves capturing voice data, encoding it, sending requests to the Google

Web Speech API, receiving and processing responses, and handling potential errors, repeating the cycle in case of failures, and performing post-processing. Therefore, in the following analysis, we will primarily focus on reporting the response time of the ASR module and the total AI training response time, as these metrics provide the most meaningful insights into the system's performance.

Figure 4.33 depicts the analysis results for the original scene. In the original scene, the Total AI training response times range from 2.5 seconds to 9.2 seconds, with an average of around 4.5 seconds. The ASR response times range from 0.9 seconds to 7.7 seconds, with an average of around 3.5 seconds. This indicates that the overall AI training process, which includes voice reception, data collection, physiological data acquisition, dataset creation, pre-processing, and AI model training, takes a relatively short time, with the ASR component significantly contributing to the total response time.

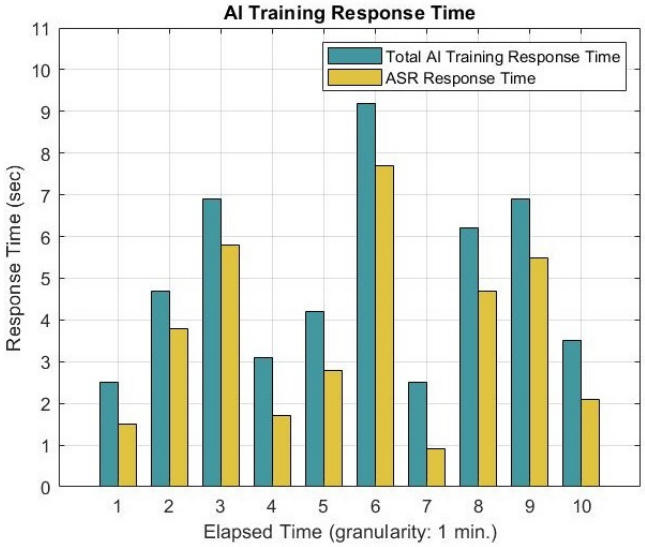


Figure 4.33: Server-side Response Time during 10 Minutes of AI Training with Original Scene: Total Response Time and ASR Response Time.

In the 2x scene, the Total AI training response times are generally higher, ranging from 5.53 seconds to 12.22 seconds, with an average of around 8 seconds (see Figure 4.34). The ASR response times also increased, ranging from 4.34 seconds to 10.59 seconds, with an average of around 6 seconds. This indicates that scaling up the system by a factor of 2 has resulted in a noticeable increase in the response times for the Total AI training and ASR processes.

According to Figure4.35, in the 3x scene, the Total AI training response times are further increased,

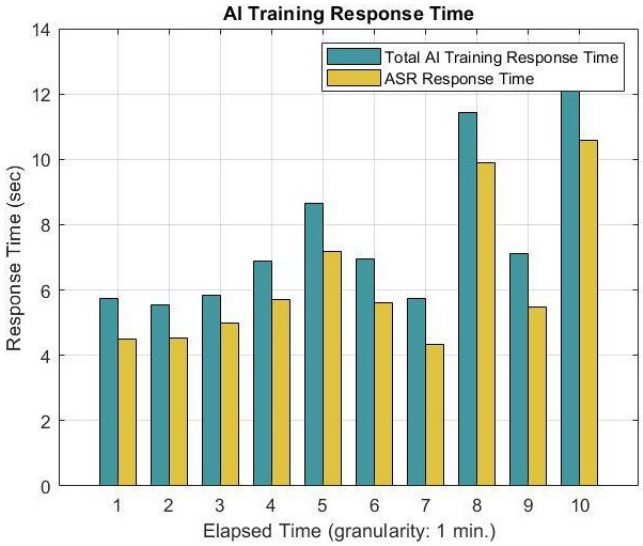


Figure 4.34: Server-side Response Time during 10 Minutes of AI Training with 2x Scene: Total Response Time and ASR Response Time.

ranging from 5.42 seconds to 15.83 seconds, with an average of around 10 seconds. The ASR response times also increased, ranging from 4.42 seconds to 14.94 seconds, with an average of around 9 seconds. This confirms that scaling up the system by a factor of 3 has led to a significant increase in the response times for both the Total AI training and ASR processes, likely due to increased resource demand and potential bottlenecks in the system. The performance differences between the scenes can be attributed to the efficiency of resource allocation, the scalability of the underlying algorithms and infrastructure, and potential optimization opportunities. Moreover, the finding that the total training time for a 1-minute data packet is less than 1 minute is a strong indicator of the system's potential to operate at a high level of responsiveness. This is a critical feature for applications that require true real-time processing. It suggests that the overall architecture and design of the training pipeline are well-suited to handle the incoming data and perform the necessary computations in a timely manner.

However, the analysis also highlights the ASR module as a potential bottleneck, with response times ranging from 0.9 seconds to 7.7 seconds and contributing significantly to the total training time. This variability in ASR performance, which is likely influenced by factors such as Internet speed and the reliability of the Google ASR service, could undermine the system's ability to maintain consistent, low-latency operation. By optimizing the ASR module, perhaps through the implementation of more robust error handling, the exploration of alternative speech recognition solutions, or utilizing another

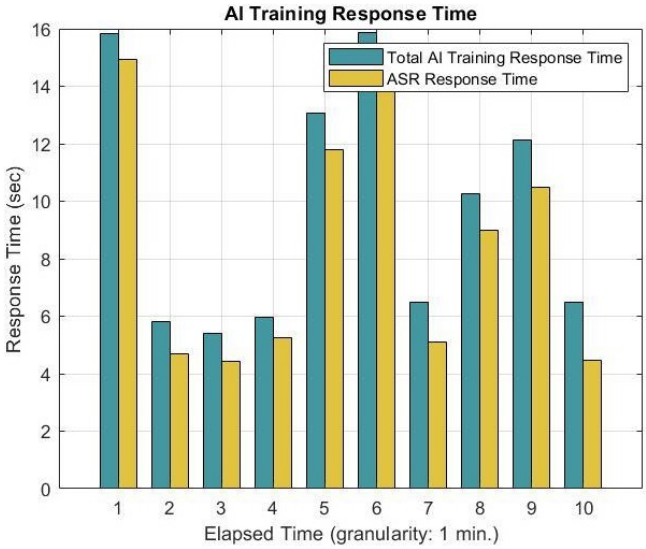


Figure 4.35: Server-side Response Time during 10 Minutes of AI Training with 3x Scene: Total Response Time and ASR Response Time.

method as a classification indicator, the granularity of the time intervals could potentially be decreased even further, allowing the system to approach actual real-time performance.

Overall, the findings from this performance analysis provide valuable insights that can guide future development efforts, focusing on enhancing the efficiency and responsiveness of the AI training process. By addressing the identified areas of improvement, particularly the ASR component, the system's ability to deliver timely and accurate results can be significantly improved, ultimately leading to a more robust and capable solution.

4.7 Summary

This chapter presented the whole implementation of the trustworthy intelligent avatar, detailing its components and outlining the associated use case for this work.

We highlighted the importance of stream learning algorithms that can adapt to the dynamic nature of intelligent avatars, discussing the challenges posed by traditional batch learning and the techniques used to handle concept drift and update models in real-time.

Next, we provided an overview of the "SmartSimVR" software architecture developed for the adaptive intelligent avatar, detailing its key features.

The driving simulator was chosen as our use case, with the car serving as the avatar. We explored the designed VR application and the driving simulator technology employed in the research, focusing on the simulation's multi-sensory aspects, hardware and software architecture, and the physics simulation necessary for creating a realistic driving experience.

One primary aim of the intelligent avatar is to address cybersickness, a significant challenge in VR environments. This chapter discussed the causes of cybersickness, including sensory conflict theory and individual factors affecting susceptibility, as well as methods for evaluating and measuring cybersickness.

We outlined a multimodal measurement approach to analyze user behavior and responses within the driving simulator environment, emphasizing physiological and behavioral data streams relevant for understanding and mitigating cybersickness. This included capturing eye tracking, head tracking, and various physiological signals, along with automatic speech recognition to collect FMS score questionnaires from participants.

Finally, we conducted a comprehensive performance analysis of the application, evaluating various metrics on both the client and server sides of the VR application under different scene complexity conditions. This assessment aimed to gauge the overall performance and efficiency of the application.

The next chapter delves into the experimental design and protocol, followed by the presentation of data analysis and results.

Chapter 5

Experiment Design, Data Analysis, and Results

Content

Given the nature of the experiment, which involves studying cybersickness in an end-to-end dynamic pipeline environment, a precise and applicable protocol was essential. To this end, we conducted two trial studies, which were critical in refining the procedure and protocol for the final experiment. The pilot phases allowed us to distinguish the bottlenecks during the experiment, identify areas for improvement, and make necessary revisions to ensure the final experimental protocol was robust and well-suited to investigate the target use case effectively. In this chapter, we delve into two key aspects of the experiment - the data collection procedure and the participant selection.

The central hypothesis guiding this research is that the use of an auto-adapted intelligent avatar in our use case, which is a driving simulator, can decrease cybersickness and, consequently, improve user safety and trustworthiness.

To assess this hypothesis, we meticulously analyze the collected data from both studies to evaluate the impact of the intelligent avatar on cybersickness reduction. To ensure a comprehensive understanding, the signals have been examined across four distinct domains: time domain, frequency domain, time-frequency domain, and topological analysis. Considering the data from these varied perspectives can help gain a more holistic view of the effects.

Alongside these signal processing techniques, the statistical analysis will also be conducted. This statistical approach quantifies the significance and reliability of the observed trends and patterns in the data. Furthermore, a correlation analysis has been performed between the SSQ scores and the relevant physiological and behavioral signals. This analysis aims to uncover any significant relationships between the subjective reports of simulator sickness and the objective measures recorded during the experiment.

This multi-faceted approach to signal analysis is expected to yield deeper insights and inform more accurate modeling and prediction capabilities within the scope of this work. Should our hypothesis prove, the results of these comprehensive analyses have the potential to significantly enhance user safety and comfort in immersive virtual environments.

5.1 Hardware Setup

The experimental setup (see Figure 5.1) was established with specific hardware and software configurations. The virtual application (client side), developed using Unity3D, was deployed on a high-

performance workstation running Windows 10 Pro 64bit version 22h2. The workstation had an 11th Gen Intel Core i7-11800H processor, 32GB of RAM, and an NVIDIA GeForce RTX 3080 Laptop GPU graphics card. These hardware specifications provided sufficient computing power and graphics capabilities to render the immersive driving simulator environment in real-time.

On the server side, Python 3.8 was utilized as the backend. The server ran on a dedicated machine powered by an Intel Core i7-7700HQ processor, 16GB of RAM, and a high-speed solid-state drive. The server machine was connected to the client workstation via a local area network (LAN) with 543 MB Ethernet connections to ensure fast and reliable communication.

The simulation was presented using a Meta Quest Pro HMD, which featured a display resolution of 1800 x 1920 pixels per eye. The HMD had a refresh rate of 90Hz and a field of view (FoV) of 106 degrees. To collect the physiological indicators, we utilized an Empatica E4 wristband worn by participants. This wristband has sensors with different sampling frequencies, including a 4 Hz electrodermal activity (EDA) sensor, a 64 Hz PPG sensor for blood volume pressure (BVP), a 4 Hz infrared thermopile (TMP) for temperature measurement, a 32 Hz 3-axis accelerometer sensor. The data was transmitted to the Empatica server computer via Bluetooth during the navigation experiment.

Additionally, the Logitech-G25 was employed as the driving tool. Figure 5.1 represents the hardware and software components of the implemented system architecture.

Figure 5.2 showcases the deployment diagram, which is a type of Unified Modeling Language (UML) diagram that illustrates the physical deployment of artifacts (such as software components or files) onto nodes (physical or virtual machines or devices). The deployment diagram helps visualize and communicate the physical infrastructure hosting the software system, including the relationship between the software components and the hardware they will run on. This can provide valuable insights into hardware requirements, network connectivity, scalability, and resilience of the overall system architecture.

5.2 Experimental Protocol and Data Collection Procedure

The study involved two driving simulation sessions (see Figure 5.3) that were assessed with a within-subjects design. The first experiment, i.e., Intelligent Avatar, showcased a unique feature—an

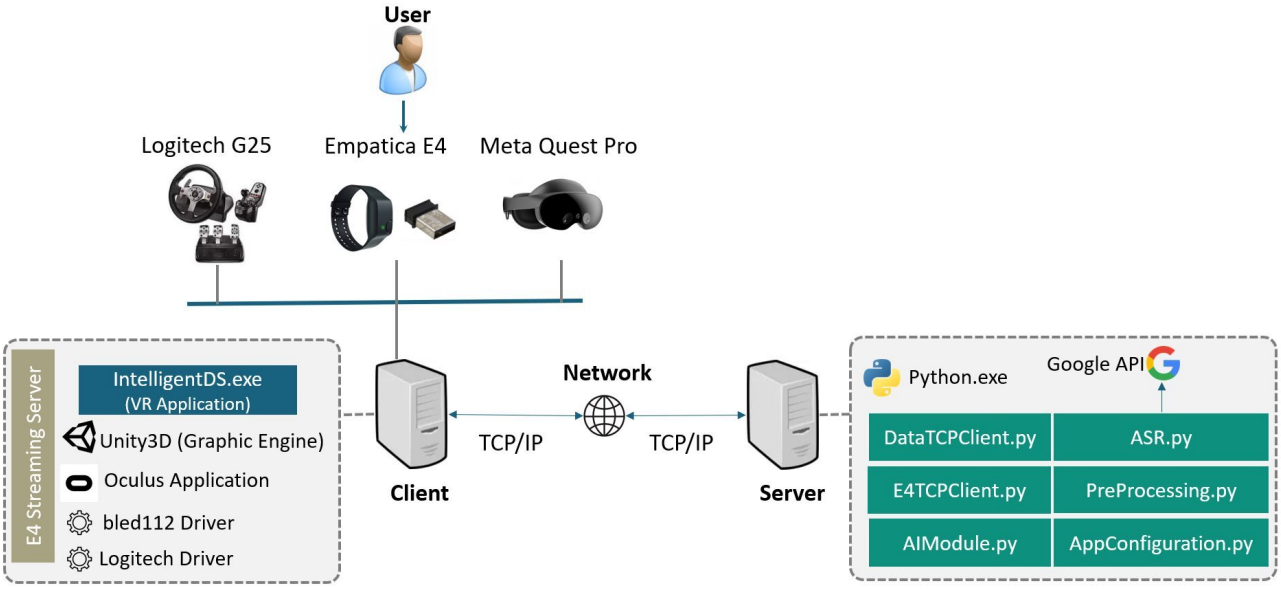


Figure 5.1: Hardware and software components of the implemented distributed system architecture.

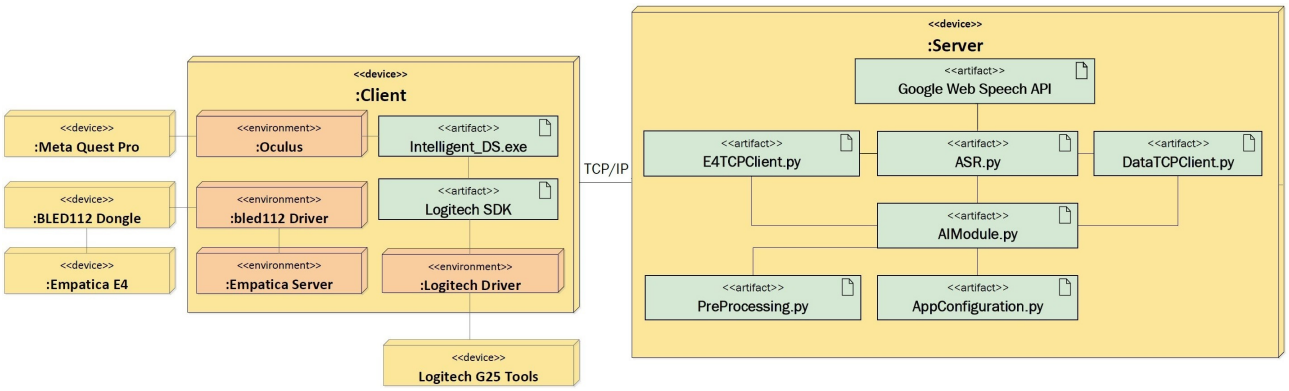


Figure 5.2: Deployment diagram of implemented system.

intelligent avatar that employed real-time AI-powered cybersickness detection and dynamic adaptation of the virtual environment. This adaptive condition was designed to monitor participants' physiological and behavioral indicators of cybersickness and automatically adjust the simulation parameters, e.g., car acceleration, to mitigate the onset of cybersickness symptoms.

The second experiment, i.e., Non-Intelligent Avatar, used a typical avatar that incorporated no AI-based adaptation. In this control condition, the simulation parameters remained static and unresponsive to participants' cybersickness state.

To minimize carryover effects, each participant completed both the Intelligent and Non-intelligent

driving simulation conditions, with a minimum washout period of 10 days between the two sessions. This extended interval ensured that participants were not overly accustomed or habituated to the virtual environment. Notably, the order in which participants experienced the two conditions was fully randomized, a crucial step to control for potential ordering or sequencing effects and ensure the validity of the study design.

The protocol and instructions provided to participants were identical across the Intelligent and Non-intelligent driving simulation conditions. This allowed for directly comparing cybersickness outcomes between adaptive and non-adaptive virtual environments.

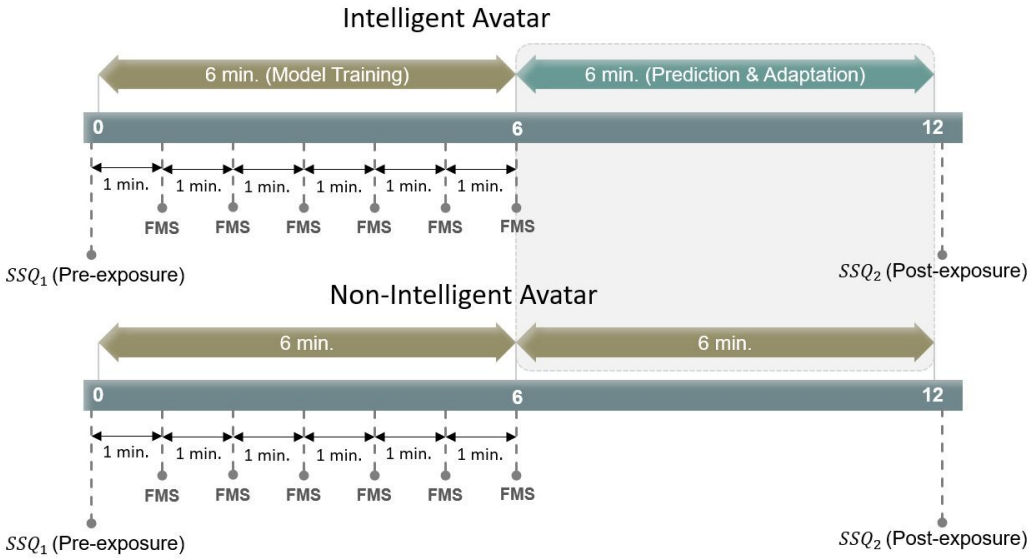


Figure 5.3: Intelligent and Non-Intelligent Avatar study protocol.

Before the commencement of the study, participants received a concise introduction to the application. This introduction was provided in three pictogram materials (Figures A.2, A.3, and A.4), including driving tools, i.e., wheel, gear shift, accelerator, brake, a self-reporting interface, and definitions of scientific terms used in the questionnaire. The FMS score was used as a classification indicator as discussed in section 4.5.3. The correlation between the scales and symptoms was explained to the participants using the provided pictogram materials before the start of the experiment, as demonstrated in Table 5.1. These correlations were considered based on the symptom coefficients in SSQ [24].

Participants were given the necessary information to understand the application and its compo-

Table 5.1: Correlation between cybersickness symptoms and adjusted FMS.

State	Description (Symptoms and Detail)	Score
State 0 (OK)	No symptom	0
State 1 (Initial) (Not OK)	Fatigue Headache Eye strain	1
State 2 (Moderate) (Not OK)	General discomfort Difficulty concentrating Stomach awareness Increase of salivation Sweating Burping	2
State 3 (Severe) (Not OK)	Difficulty focusing Fullness of head Blurred vision Vertigo Dizziness (eye open) Dizziness (eye closed) Nausea	3

nents. If participants experienced severe symptoms of cybersickness, they were allowed to discontinue the experiment if they encountered any severe symptoms. Their well-being and comfort were prioritized, and they were assured they had the freedom to withdraw from the experiment if necessary.

Subsequently, all participants were asked to provide their informed consent by reading and signing a consent form. Additionally, they were required to complete a general information form through Google form. This form captured their gaming experience and prior usage of VR devices and aimed to identify any potential factors that could impact the experiment's outcomes, such as health issues. No health concerns that could affect the experiment were reported in response to this questionnaire. (The participant training and collecting of the filled consent form and general information form were only conducted during the first experiment.)

Afterward, as is shown in Figure 5.4, the steps of subjective and objective data collection were initiated in the following sequence:

- Participants first completed a pre-exposure questionnaire, the Simulator Sickness Questionnaire (SSQ) on Google form, to assess their baseline health status and susceptibility to motion sickness.

- Next, all necessary devices, including the head-mounted display worn by the participants, were adequately set up and calibrated.
- Participants were then given a 3-minute rest period, seated in a non-swiveling chair behind the driving simulator controls. This intentional break allowed them to mentally and physically prepare for the upcoming virtual reality exposure, a vital aspect of the study.
- Following the preparatory rest period, participants were fully immersed in the virtual driving environment and instructed to navigate the simulated city streets.
- After completing the virtual driving task, participants were asked to fill out the post-exposure SSQ form. This step was crucial in assessing any changes in their cybersickness symptoms, providing valuable data for the study.

Figure 5.4: Experiment procedure steps.

Due to the documented effectiveness of herbal remedies in alleviating the symptoms of motion sickness, the participants who experienced symptoms during the experiment were provided with ginger tea [252] or mint tea [66], and were also offered chocolate, as a means of supporting their well-being upon completion of the study.

The experimental protocol placed particular emphasis on the duration of the experiment and the participants' self-reported state as classification indicators.

- **Self-Reporting State:**

  The FMS scale was used to assess the participants' sickness levels. During the VR simulation, the system automatically played an audio question, "What is your score?" every minute, and participants verbally reported their sickness level on a scale from 0 to 3, based on Table 5.1. Expressly, a score of 0 indicated no sickness, a score of 1 indicated initial symptoms, a score of 2 denoted moderate symptoms, and a score of 3 represented severe symptoms. The original FMS

scale, which ranged from 1 to 20, was adjusted to a more concise scale of 0 to 3 based on the trial experiment. This adjustment was made because participants needed help to report their status with high-resolution accuracy.

- **Exposure Time:** The duration of the experiment was flexible, ranging from 10 to 12 minutes. This included 4-6 minutes for training and an additional 6 minutes dedicated to detection and adaptation. Throughout this time-frame, relevant physiological and behavioral data were captured and recorded from the participants.

If a participant reported a non-zero score on the FMS for three consecutive minutes, the system immediately halted the pure training phase. It then transitioned to the detection and adaptation phase, which continues the training process with real-time monitoring and adjustment of car acceleration, demonstrating the system's responsiveness to participant feedback.

Conversely, if a participant did not report any FMS score greater than 0 within the initial 6 minutes, the program would be completely stopped, bypassing the detection/adaptation phase. This decision was made to ensure the well-being of the participants and to dynamically determine the experiment's duration, highlighting the system's adaptability.

This division was implemented because trial participants experiencing sickness or discomfort could not continue the experiment for more than 3 or 4 minutes. By allowing for variations in exposure time, the experiment aimed to capture accurate data while considering the well-being of the participants.

5.3 Participants

As an initial step, a trial study was conducted to establish the most effective and reliable protocol and identify potential study limitations. Six participants (M:5, F:1), most of whom were colleagues from the IMI laboratory, took part in this trial.

For the main study, 30 participants were recruited from various nationalities, consisting of 29 university students and 1 engineer (M:15, F:15). To recruit the final participant pool, a recruitment flyer (as shown in Appendix A (Figure A.1)) was designed and distributed across social media platforms and within the university community.

Our participants, with an average age of 23.43 years and a standard deviation of 4.12 years,

were thoroughly screened for good health and absence of auditory disorders. They all had normal or corrected-to-normal vision. Among them, 11 had prior experience using virtual reality controllers, 8 had experience with driving controllers, and 18 had experience with game controllers, ensuring a rigorous and reliable study.

Regarding motion sickness, 12 participants reported occasionally experiencing sickness during travel, with specific instances including car travel (5 participants), boat travel (4 participants), ship/ferry travel (2 participants), and roller coaster rides (1 participant).

During the non-intelligent driving simulation, we prioritized the comfort and safety of our participants. Four experienced severe cybersickness, leading them to discontinue the experiment due to acute discomfort. One participant did not experience any symptoms in either study and reported only a FMS score 0. Our primary focus was on participants who could complete both sessions for at least 10 minutes and experienced some level of discomfort, enabling us to examine potential physiological changes associated with cybersickness. Figure 5.5 illustrates the experimental setup for the participants, while Figure 5.6 depicts the server-side processes during the experiment.

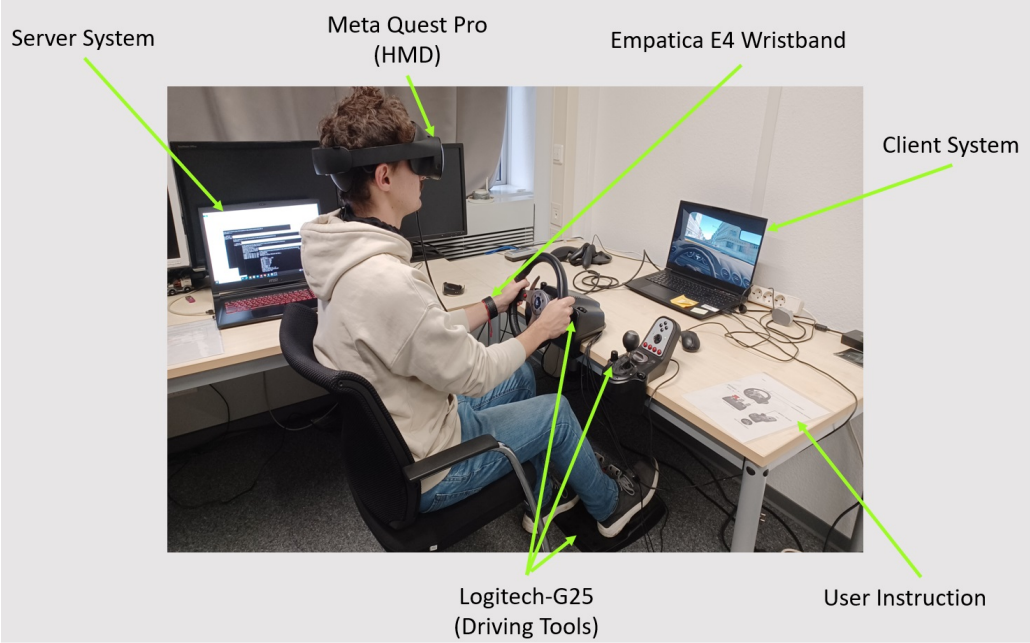


Figure 5.5: Experimental setup for the study with Meta Quest Pro head-mounted display (HMD) and Empatica E4 wristband attached to a participant.

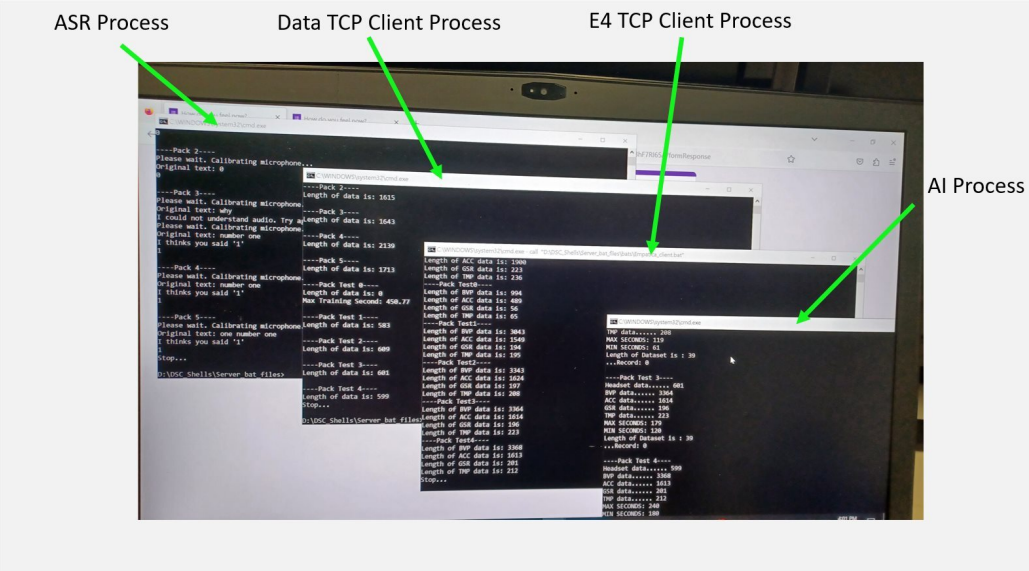


Figure 5.6: Processes running on the server system during the experiment.

5.4 Assessing Model Performance During Training

As discussed in the previous chapter, the assessment of the training model focused on the binary classification of physiological and behavioral data, utilizing self-reported FMS scores as a classification indicator. The model was trained incrementally using a stream learning approach, allowing it to adapt dynamically to new data collected every minute. Each participant contributed a dataset with dimensions of $\mathbb{R}^{240\times25}$ where 240 representing the recorded data points per minute and 25 indicating the number of physiological and behavioral signals (independent variables) captured. This incremental training strategy, leveraging these 240 recorded data points at a time, enabled the model to learn patterns associated with varying FMS scores effectively. The model's performance was evaluated through metrics such as accuracy and F1 score, which were logged after each update to monitor its learning progression and overall effectiveness in predicting cybersickness.

Analyzing these evaluation metrics provides valuable insights into the performance of the models trained for each participant during the intelligent avatar experiment. The box plot visualization (see Figure 5.7) illustrates the distribution of these metrics.

- **Accuracy Metric:**

  Accuracy is a metric used to evaluate the performance of a classification model. It is defined as

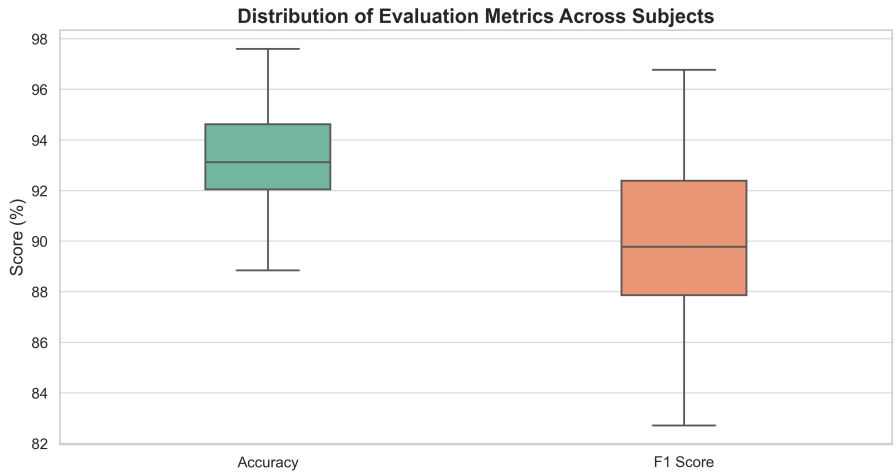


Figure 5.7: Box plot showing the distribution of model performance across subjects. Performance was evaluated using accuracy and F1 score metrics.

the ratio of correctly predicted instances to the total instances in the dataset. Mathematically, it can be expressed as:

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Instances}} \tag{5.1}$$

Accuracy provides a straightforward measure of how often the classifier is correct, but it may only partially capture model performance, especially in cases of class imbalance.

According to Figure 5.7, the median accuracy across all participants was an impressive 93.11%, indicating that the models effectively and accurately classified the physiological and behavioral data across the participant group. This high median suggests that the individual-trained models generally performed well, reflecting their impressive ability to adapt to the unique data characteristics contributed by each participant.

The first quartile (Q1) was 92.04%, meaning 25% of participants achieved accuracy below this threshold. In contrast, the third quartile (Q3) was 94.62%, indicating that 75% of participants performed at or below this accuracy level. The relatively narrow interquartile range (IQR) of 2.58% signifies a consistent performance across most participants. This suggests that the individual models maintained a similar level of effectiveness in their predictions and resulted in reliable performance across diverse data sets.

- **F1 Score Metric:**

F1 Score is a metric that combines precision and recall to provide a single measure of a model's accuracy that considers both false positives and false negatives. It is beneficial in situations where the class distribution is imbalanced. The F1 score is the harmonic mean of precision and recall, defined as:

$$\text{F1 Score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}} \tag{5.2}$$

where Precision is the ratio of true positive predictions to the total predicted positives and Recall (or Sensitivity) is the ratio of true positive predictions to the actual positives.

The median F1 score was recorded at 89.78%, which provides a balanced view of the models' precision and recall in identifying true positives while minimizing false positives and negatives. This high median F1 score reinforces the individual models' capability to accurately predict cybersickness in users, highlighting their robustness in real-time monitoring scenarios.

The first quartile for the F1 score stood at 87.86%, indicating that a quarter of the participants had F1 scores below this level. Conversely, the third quartile was 92.39%, suggesting that the upper 25% of participants achieved scores above this threshold. The IQR for the F1 score is 4.53%, reflecting a slightly wider variability compared to the accuracy metrics. This variation may indicate that while the models performed well on average, there were instances where specific participants faced more significant challenges in accurate classification, potentially due to individual differences in their physiological and behavioral data.

Overall, the accuracy and F1 score metrics distribution across all participants demonstrates strong performance in predicting cybersickness. The high median values and relatively narrow interquartile ranges indicate the effectiveness of the models trained for each user. However, the observed variability in the F1 score suggests the need for further exploration into factors contributing to individual performance differences. This understanding could inform future model enhancements, ultimately improving predictive capabilities across all users in real-world applications.

5.5 Time Domain Analysis

As discussed in section 4.5, 10 distinct data modalities comprising 26 signals spanning behavioral and physiological signals were measured as objective data. With this comprehensive dataset, time domain analysis using the mean, standard deviation, and root-mean-square (RMS) can better characterize the participants' responses. The mean values of each signal can indicate the general level of activity, arousal, or physiological state. For example, the mean eye velocity, head position, and accelerometer readings can reveal baseline visual attention patterns and gross motor behavior. Meanwhile, the mean EDA, BVP, and temperature levels can reflect autonomic nervous system activation. Comparing the means of these signals between experimental conditions can highlight differences in overall physiological and behavioral engagement.

The standard deviations can illuminate the dynamic variability within each signal. Higher variability in eye movements, head position, or cardiovascular measures may correspond to more reactive, flexible responses to the experimental stimuli. The RMS value estimates the overall magnitude or power of the signal over time. For instance, higher RMS in the accelerometer or eye movement data can indicate more intense physical activity or oculomotor responses.

Analyzing this multimodal dataset's means, standard deviations, and RMS can provide a comprehensive time-domain profile of the participants' psychophysiological states and information-processing strategies. This rich set of temporal statistics can be a powerful starting point for further analysis and interpretation of the complex relationships between brain, body, and behavior.

5.5.1 Mean and Standard Deviation

The key findings of our study, as presented in Table B.1, highlight the significant impact of the intelligent avatar on participants' behavioral and physiological responses during the virtual driving task. The intelligent avatar condition, compared to the non-intelligent condition, was associated with notably lower oculomotor, head, and postural parameters, as well as reduced physiological arousal.

Eye movements aligned with different axes (see Figure 5.8) represented these differences, with the intelligent avatar study showing lower mean eye velocity along the x-axis (intelligent: $M = -1.2e-5, SD = 0.008$ vs. non-intelligent: $M = 6.8e-5, SD = 0.010, t = 1.7, p < 0.05$) and y-axis (intelligent: $M = -5.7e-5, SD = 0.003$ vs. non-intelligent:$M = 6.9e-5, SD = 0.004, t = 1.5, p < 0.05$), as

well as lower mean eye angular velocity along the x-axis (intelligent: $M = -1.8e-4, SD = 0.024$ vs. non-intelligent: $M = 5.1e-4, SD = 0.038, t = 0.824, p < 0.05$) and y-axis (intelligent: $M = -3.4e-4, SD = 0.049$ vs. non-intelligent: $M = 5.3e-4, SD = 0.072, t = 0.797, p < 0.05$).

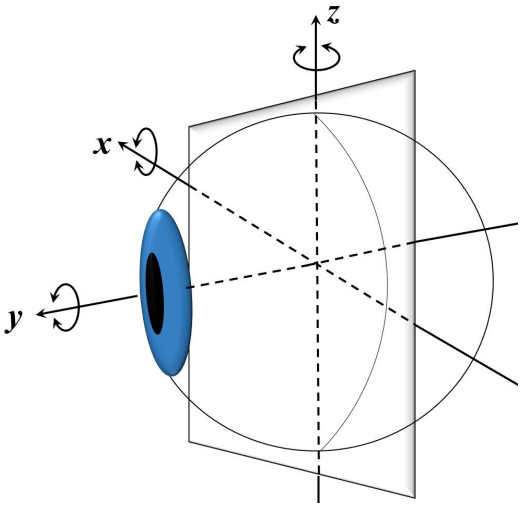


Figure 5.8: Spherical Geometry of the Eye: Illustrating the Three Planes of ocular movement.

Similarly, the Intelligent avatar study showed lower mean head position along the x-axis (intelligent: $M = -0.082, SD = 0.013$ vs. non-intelligent: $M = -0.051, SD = 0.013, t = 1.964, p < 0.05$) and lower mean head rotation along the z-axis (intelligent: $M = 0.004, SD = 0.018$ vs. non-intelligent: $M = 0.026, SD = 0.018, t = 3.073, p < 0.05$), indicating more constrained head movements compared to the Non-Intelligent avatar study. The Intelligent avatar study also demonstrated lower mean acceleration along the z-axis (intelligent: $M = 21.511, SD = 10.431$ vs. non-intelligent: $M = 29.893, SD = 13.555, t = 2.029, p < 0.05$) and lower mean EDA values (intelligent: $M = 1.204, SD = 0.198$ vs. non-intelligent: $M = 2.276, SD = 0.480, t = 2.119, p < 0.05$), which could be indicative of more regulated physical and physiological responses, further suggesting a lower susceptibility to cybersickness.

Figure 5.9 and Figure 5.10 display a sample of the eye velocity in the x, y, and z directions recorded from a single participant during the non-intelligent and intelligent avatar studies. Remarkably, the standard deviation and variability of the eye velocity were lower during the intelligent avatar experiment (Figure 5.10) compared to the non-intelligent avatar study. This suggests that the intelligent avatar interface provided the participant with a more comfortable and manageable user experience.

In the non-intelligent avatar experiment, the participant's adaptation phase lasted for 2.5 minutes. However, the participant discontinued the experiment before completing the entire procedure due to

experiencing cybersickness, as illustrated in Figure 5.9. In contrast, during the intelligent avatar experiment (Figure 5.10), the participant completed the full 6-minute duration, demonstrating improved tolerance and adaptability.

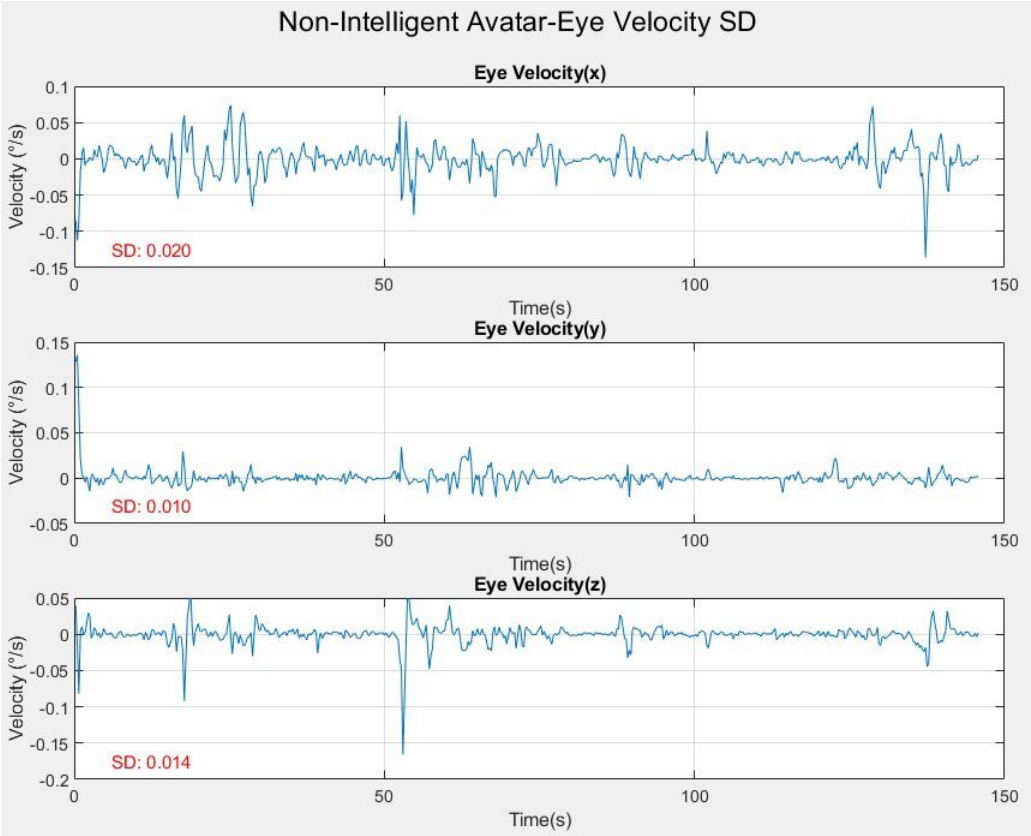


Figure 5.9: Eye Velocity for a participant during Non-Intelligent Avatar experiment.

These reduced standard deviations in oculomotor and physiological parameters are highly relevant to cybersickness. The more stable, less variable oculomotor responses and physiological signals observed in the intelligent avatar condition suggest that this interaction was less taxing on the user's sensory integration and adaptation processes, potentially mitigating the risk of cybersickness.

The lower variability in measures like eye velocity, angular velocity, and acceleration indicates more consistent, predictable visual-vestibular coupling, which is a critical factor in reducing cybersickness. Similarly, the dampened electrodermal and thermal responses point to a lower level of autonomic arousal and stress, further supporting the notion that the intelligent avatar condition was less disruptive to the user's sensory and physiological equilibrium. These findings highlight the potential of the

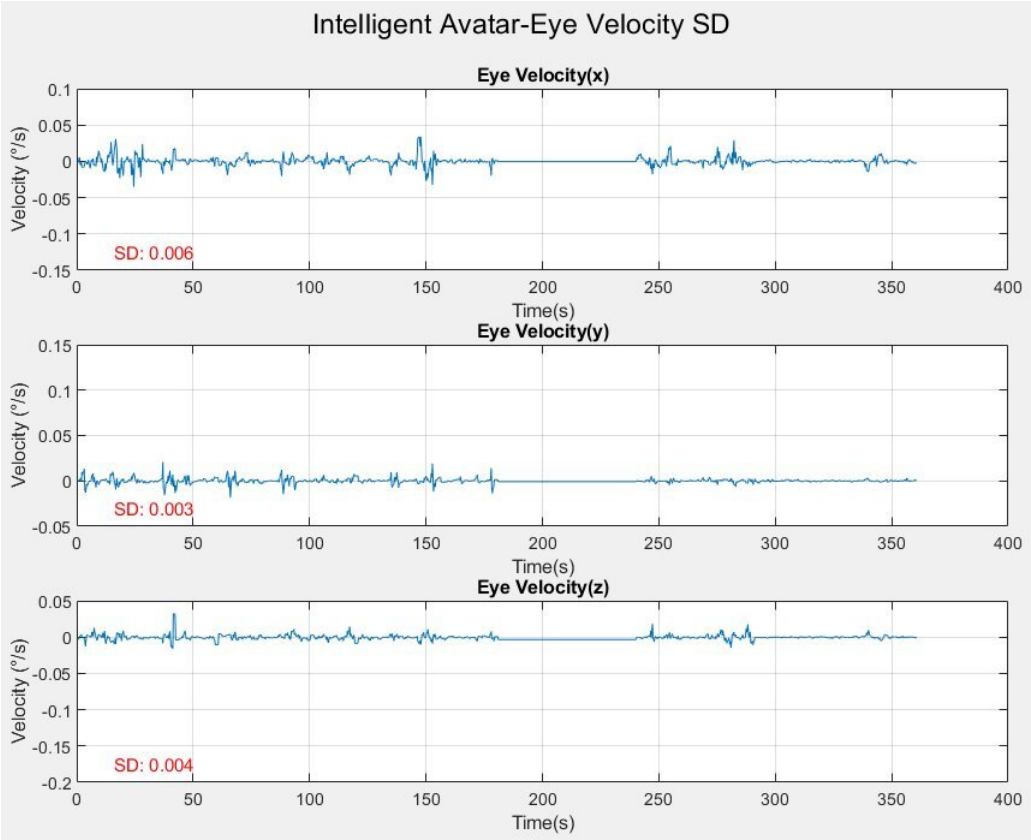


Figure 5.10: Eye Velocity of the same participant during Intelligent Avatar experiment.

intelligent avatar design to enhance the virtual experience by minimizing cybersickness-related issues.

5.5.2 Root Mean Square (RMS)

Root Mean Square (RMS) is a statistical measure that represents the square root of the mean of the squared values of a variable. It is commonly used to quantify the magnitude or intensity of a time-varying signal or parameter. The formula for calculating RMS is:

$$RMS = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}} \tag{5.3}$$

Where x represents the individual data points and n is the total number of data points.

The RMS values in Table B.2 comprehensively assess the magnitude and consistency of various oculomotor and physiological parameters.

For eye velocity along the x-axis, the intelligent avatar condition had a significantly lower RMS (intelligent: $M = 0.007, SD = 0.003$ vs. non-intelligent: $M = 0.010, SD = 0.005, t = 1.975, p < 0.05$).

Similarly, the RMS of eye velocity along the y-axis was lower in the intelligent avatar condition (intelligent: $M = 0.003, SD = 0.001$ vs. non-intelligent: $M = 0.004, SD = 0.002, t = 2.054, p < 0.05$).

In terms of eye angular velocity, the intelligent avatar condition had significantly lower RMS values for eye angular velocity around the x-axis (intelligent: $M = 0.026, SD = 0.014$ vs. non-intelligent: $M = 0.039, SD = 0.022, t = 2.177, p < 0.05$), y-axis (intelligent: $M = 0.049, SD = 0.018$ vs. non-intelligent: $M = 0.072, SD = 0.047, t = 2.040, p < 0.05$), and z-axis (intelligent: $M = 0.022, SD = 0.009$ vs. non-intelligent: $M = 0.030, SD = 0.014, t = 2.057, p < 0.05$) compared to the Non-Intelligent avatar condition.

Additionally, the Intelligent avatar condition had significantly lower RMS values for eye position along the y-axis (intelligent: $M = 1.130, SD = 0.037$ vs. non-intelligent: $M = 1.157, SD = 0.040, t = 2.272, p < 0.05$).

Finally, the x-axis of eye rotation showed a significant difference, with the intelligent avatar condition (intelligent: $M = 0.040, SD = 0.020$ vs. non-intelligent: $M = 0.065, SD = 0.033, t = 2.986, p < 0.05$).

This indicates that the intelligent avatar effectively dampened the magnitude and variability of these key oculomotor parameters, contributing to a more stable and coherent visual experience for the participants.

Furthermore, the results show that the intelligent avatar was better able to stabilize the vertical position of the head (intelligent: $M = 1.129, SD = 0.033$ vs. non-intelligent: $M = 1.153, SD = 0.043, t = 2.007, p < 0.05$) and control its rotation around the horizontal axis(intelligent: $M = 0.042, SD = 0.022$ vs. non-intelligent: $M = 0.064, SD = 0.032, t = 2.663, p < 0.05$).

In terms of physiological data, the RMS value of EDA was also significantly lower in the intelligent avatar condition (intelligent: $M = 1.230, SD = 1.558$ vs. non-intelligent: $M = 2.319, SD = 1.883, t = 1.229, p < 0.05$), suggesting a more consistent and less variable physiological response from the participants.

Figure 5.11 represents the EDA data of a single participant during both sessions. The RMS values are overlaid on the signals, indicating that the RMS for the intelligent avatar experiment (RMS = 1.31 μS) is lower than the RMS for the non-intelligent avatar experiment (RMS = 5.45 μS). This substantial reduction in the RMS indicates that the participant exhibited lower overall physiological

arousal when interacting with the intelligent avatar system than the non-intelligent avatar.

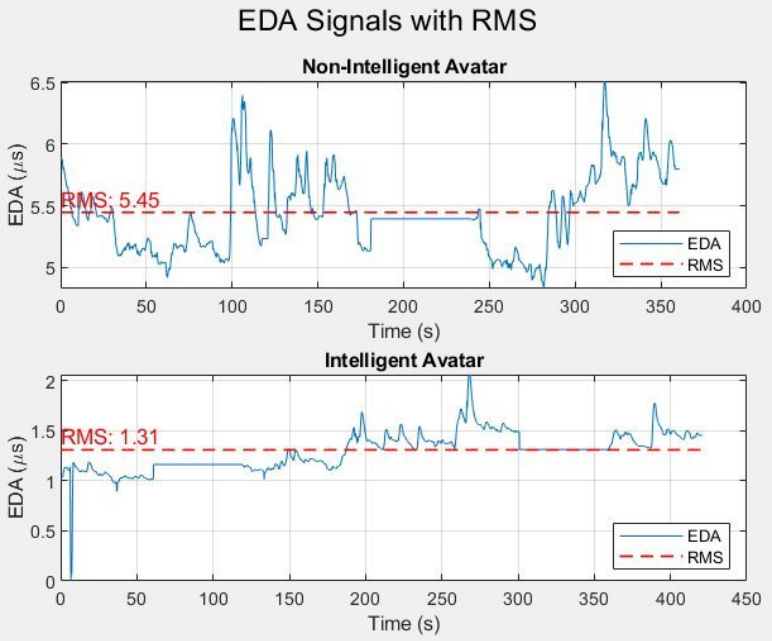


Figure 5.11: EDA of the same participant during both experiments.

The comprehensive time domain analysis provided valuable insights into the participants' behavioral and physiological responses to the intelligent avatar during the virtual driving task. However, the time domain approach, which focuses on summary statistics, may not be sufficient to characterize the complex and time-varying nature of the signals fully. The measured signals, including eye movements, head position, and physiological measures, are inherently non-stationary, and their underlying mechanisms may not be fully captured by the time domain analysis alone.

5.6 Frequency Domain Analysis

Frequency-domain analysis techniques were employed to examine the signals' spectral properties. Mean frequency, power spectral density (PSD) estimation, and entropy calculations were used to elucidate the rhythmic and information-theoretic aspects of the participants' responses to the intelligent avatar.

PSD analysis quantified signal power distribution across different frequency bands, revealing the dominant frequency components and potential neural or physiological rhythms modulated by the intelligent avatar interface. Mean frequency calculations provided an index of the central tendency of

the spectral content, indicating shifts in the overall frequency composition of the signals. Additionally, entropy measures were applied to assess the complexity and information content of the physiological and behavioral time series, capturing changes in the predictability and regularity of the participants' responses.

Integrating these frequency domain metrics with the previously reported time domain findings can provide a more comprehensive understanding of the participant's cognitive and psychophysiological processes.

5.6.1 Power Spectral Density (PSD)

Power Spectral Density (PSD) is a fundamental concept in frequency domain analysis that describes how the power of a signal or time series is distributed over different frequencies. It provides valuable insights into a signal's frequency characteristics that are not easily discernible in the time domain, making it an essential tool in various fields, including signal processing, vibration analysis, and communication systems.

PSD analysis can help identify the dominant frequencies or frequency bands that contain most of the signal's power. This information is crucial for understanding the frequency-dependent characteristics of the system or process under study.

The PSD, denoted as $S_x x(f)$, is the Fourier transform of the autocorrelation function $R_x x(\tau)$ of a wide-sense stationary (WSS) random process $x(t)$:

$$S_{xx}(f) = \int_{-\infty}^{\infty} R_{xx}(\tau) e^{-j2\pi f\tau} d\tau \tag{5.4}$$

where f is the frequency, and $j = \sqrt{-1}$. According to the analysis, as presented in Table B.3, the PSD values for eye velocity in both the x (intelligent: $M = 2.97e-5, SD = 1.740e-5$ vs. non-intelligent: $M = 5.47e-5, SD = 5.109e-5, t = 2.076, p < 0.05$) and y (intelligent: $M = 4.29e-6, SD = 3.591e-6$ vs. non-intelligent: $M = 8.45e-6, SD = 8.282e-6, t = 2.064, p < 0.05$) directions, as well as eye angular velocity in the x (intelligent: $M = 3e-4, SD = 2.271e-4$ vs. non-intelligent: $M = 6.3e{-}4, SD = 6.373e{-}4, t = 2.069, p < 0.05$) and y (intelligent: $M = 0.0012, SD = 8.122e{-}4$ vs. non-intelligent: $M = 0.0032, SD = 0.0041, t = 2.136, p < 0.05$) directions, were significantly lower in the intelligent avatar condition compared to the non-intelligent avatar condition. Also, eye position in the y direction (intelligent: $M = 0.634, SD = 0.037$ vs. non-intelligent: $M = 0.662, SD = 0.049, t =$

$2.065, p < 0.05$) and eye rotation in the x direction (intelligent: $M = 0.001, SD = 9.051e-4$ vs. non-intelligent: $M = 0.003, SD = 0.002, t = 3.061, p < 0.05$) were significantly lower in the intelligent avatar condition than in the non-intelligent.

These findings have significant implications for the field of human-computer interaction, as they suggest that the intelligent avatar may lead to more controlled and refined eye movements. The reduced PSD values imply a decrease in the overall variability and energy content of the eye velocity, angular velocity, horizontal rotation signals and vertical gaze, which may be related to changes in visual attention, information processing, and oculomotor control when engaging with the intelligent avatar. This could potentially revolutionize the design of virtual interfaces, making them more intuitive and less taxing on the user's cognitive resources.

Furthermore, the PSD value for the head position in the y direction (intelligent: $M = 0.63, SD = 0.035$ vs. non-intelligent: $M = 0.66, SD = 0.051, t = 2.271, p < 0.05$) and head rotation in x direction (intelligent: $M = 0.001, SD = 0.001$ vs. non-intelligent: $M = 0.003, SD = 0.002, t = 2.746, p < 0.05$) where also significantly lower in the intelligent avatar condition. These findings suggest that the participants exhibited less variability in their head positions and rotation when interacting with the intelligent avatar, potentially reflecting more focused attention and reduced scanning behavior. The reduced PSD values indicate that the intelligent avatar may have elicited a more stable and controlled postural strategy from the participants during the driving task.

Additionally, the PSD value for EDA was significantly lower in the intelligent avatar condition compared to the non-intelligent avatar condition (intelligent: $M = 1.033, SD = 2.490$ vs. non-intelligent: $M = 4.615, SD = 6.707, t = 2.294, p < 0.05$). This finding suggests that the intelligent avatar elicited a reduced physiological arousal response from the participants, as indicated by the decreased power in the EDA signal. The lower EDA PSD value may be associated with decreased autonomic activation, which could be related to changes in emotional engagement, cognitive load, or stress levels when interacting with the intelligent avatar.

Our comprehensive analysis of the experiment's variables in other measures, did not show statistically significant differences in PSD values between the two conditions. This thorough examination indicates that the presence of the intelligent avatar did not significantly influence the frequency-domain characteristics of these parameters.

Figure 5.12 and Figure 5.13 present the PSD of eye angular velocity along the x, y, and z axes for a single participant during the non-intelligent and intelligent avatar sessions, respectively. Across all three spatial dimensions, the mean PSD values during the intelligent avatar experiment are noticeably lower than those in the non-intelligent avatar experiment. This suggests that the participant exhibited more stable and controlled eye movements when interacting with the intelligent avatar system than the non-intelligent avatar. The reduced PSD magnitudes imply that the participant could more precisely track and focus on the avatar's movements and behaviors under the intelligent condition.

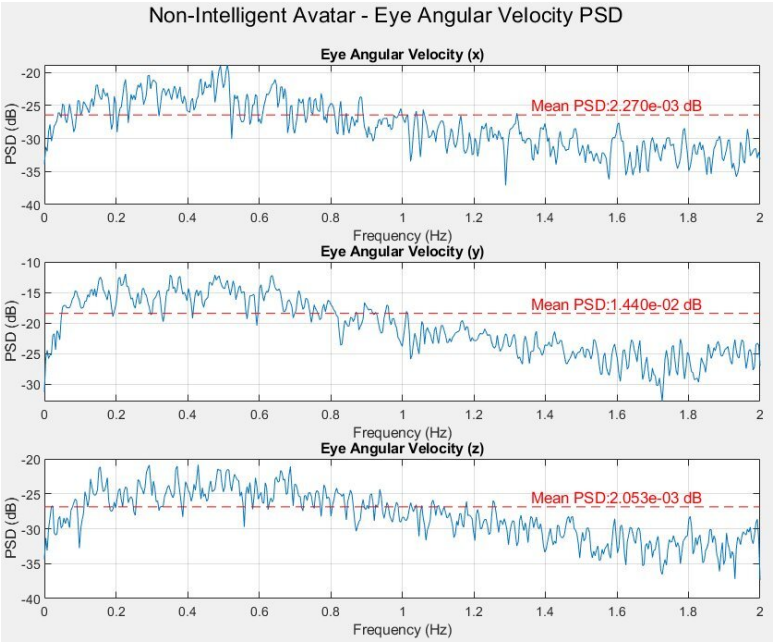


Figure 5.12: PSD of Eye Angular Velocity for a Participant during Non-Intelligent Avatar Experiment.

5.6.2 Spectral Entropy

Spectral Entropy in the frequency domain measures the uncertainty or randomness of a signal's frequency components. It quantifies power distribution across different frequencies, providing insights into the signal's complexity and predictability. A higher spectral entropy indicates a more uniform power distribution across frequencies, suggesting a more complex or unpredictable signal. Conversely, a lower spectral entropy implies a more concentrated or structured frequency content, which may be characteristic of periodic or deterministic signals.

To get the spectral Entropy, a Discrete Fourier Transform (DFT) is performed on the signal to

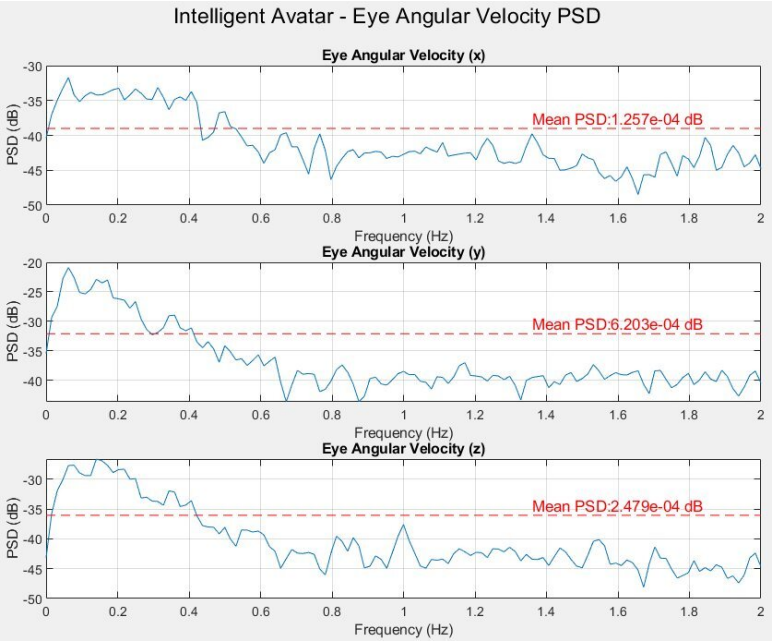


Figure 5.13: PSD of Eye Angular Velocity for the same Participant during Intelligent Avatar Experiment.

obtain the frequency components. The PSD is computed by squaring the amplitude spectrum and scaling it by the number of frequency bins. Then, the power spectrum should be normalized so that the sum of the power over all frequencies equals 1.

$$SE = -\sum_{i} P_{\text{norm}}(f_i) \log(P_{\text{norm}}(f_i)) \tag{5.5}$$

where $P_{\text{norm}}(f_i)$ is the normalized power at frequency f_i.

According to Table B.4, eye velocity x(intelligent: $M = 8.445, SD = 0.844$ vs. non-intelligent: $M = 8.958, SD = 0.897, t = 2.084, p < 0.05$), y(intelligent: $M = 8.671, SD = 0.755$ vs. non-intelligent: $M = 9.237, SD = 0.775, t = 2.451, p < 0.05$), and z(intelligent: $M = 8.617, SD = 0.789$ vs. non-intelligent: $M = 9.096, SD = 0.793, t = 2.139, p < 0.05$), eye angular velocity x(intelligent: $M = 8.705, SD = 0.589$ vs. non-intelligent: $M = 9.281, SD = 0.783, t = 2.753, p < 0.05$), y(intelligent: $M = 8.311, SD = 0.838$ vs. non-intelligent: $M = 8.932, SD = 0.950, t = 2.299, p < 0.05$), and z(intelligent: $M = 8.563, SD = 0.850$ vs. non-intelligent: $M = 9.080, SD = 0.867, t = 2.132, p < 0.05$), and eye rotation y(intelligent: $M = 4.81, SD = 1.273$ vs. non-intelligent: $M = 5.669, SD = 1.359, t = 2.301, p < 0.05$) exhibited statistically significant decreases in spectral entropy when interacting with the intelligent avatar compared to the non-intelligent avatar. The decreases in spectral

entropy for these eye-related measurements suggest that the eye movements and rotations became more structured, ordered, and less random when interacting with the intelligent avatar. Eye behavior's reduced randomness and increased organization are often associated with lower levels of cybersickness, as they indicate more stable visual processing and gaze control.

Also, head rotation in y(intelligent: $M = 4.807, SD = 1.272$ vs. non-intelligent: $M = 5.663, SD = 1.357, t = 2.3, p < 0.05$) and z(intelligent: $M = 2.149, SD = 1.405$ vs. non-intelligent: $M = 3.3, SD = 2.099, t = 2.038, p < 0.05$) directions were more stable in intelligent avatar condition which contribute to reduced susceptibility to cybersickness compared to non-intelligent avatar.

Furthermore, the decreases in spectral entropy for acceleration y(intelligent: $M = 0.756, SD = 0.608$ vs. non-intelligent: $M = 1.391, SD = 1.304, t = 2.020, p < 0.05$) and z(intelligent: $M = 1.445, SD = 0.991$ vs. non-intelligent: $M = 2.696, SD = 1.993, t = 2.575, p < 0.05$), EDA(intelligent: $M = 0.216, SD = 0.205$ vs. non-intelligent: $M = 0.397, SD = 0.285, t = 2.363, p < 0.05$), and BVP(intelligent: $M = 8.77, SD = 0.861$ vs. non-intelligent: $M = 9.41, SD = 1.186, t = 2.183, p < 0.05$) during the interaction with the intelligent avatar suggest that these physiological signals exhibited more orderly and less random patterns. This more ordered physiological response suggests improved emotional regulation and a more relaxed user experience, which can be beneficial in mitigating the physiological symptoms associated with cybersickness, such as nausea, dizziness, and discomfort.

For other measurements, such as eye position, head position, and temperature, no statistically significant differences in spectral entropy were observed between the non-intelligent avatar and intelligent avatar conditions.

5.6.3 Mean Frequency

Mean frequency is a measure used in the frequency domain to describe the average frequency of a signal's power spectrum. It provides a single value representing the central tendency of the signal's power frequency distribution. This metric is used for biomedical signal analysis to identify dominant frequency components.

Mathematically, mean frequency, denoted as freq, is the average frequency of the power spectrum of a time-domain signal. It is calculated using the signal's power spectral density (PSD).

$$\text{freq} = \frac{\sum_i f_i P(f_i)}{\sum_i P(f_i)} \tag{5.6}$$

where f_i are the frequency components and $P(f_i)$ is the power at frequency f_i.

Table B.5 presents the data analysis using mean frequency. Again, the results of this analysis indicate more stable and controlled eye movements while interacting with the intelligent avatar. They revealed that eye velocity x(intelligent: $M = 0.681, SD = 0.092$ vs. non-intelligent: $M = 0.748, SD = 0.130, t = 2.023, p < 0.05$) and y(intelligent: $M = 0.756, SD = 0.111$ vs. non-intelligent: $M = 0.879, SD = 0.118, t = 2.620, p < 0.05$), eye angular velocity x(intelligent: $M = 0.765, SD = 0.105$ vs. non-intelligent: $M = 0.856, SD = 0.155, t = 2.134, p < 0.05$), y(intelligent: $M = 0.653, SD = 0.122$ vs. non-intelligent: $M = 0.750, SD = 0.141, t = 2.321, p < 0.05$), and z(intelligent: $M = 0.741, SD = 0.136$ vs. non-intelligent: $M = 0.837, SD = 0.141, t = 2.285, p < 0.05$), and eye rotation y(intelligent: $M = 0.095, SD = 0.051$ vs. non-intelligent: $M = 0.131, SD = 0.055, t = 2.150, p < 0.05$), and z(intelligent: $M = 0.034, SD = 0.025$ vs. non-intelligent: $M = 0.072, SD = 0.074, t = 2.150, p < 0.05$) all demonstrated statistically significant decreases in mean frequency during interaction with the intelligent avatar, as opposed to the non-intelligent avatar. These findings are significant as they shed light on the impact of intelligent avatars on human physiological and cognitive responses.

Similarly, the mean frequency values for head rotation y(intelligent: $M = 0.090, SD = 0.051$ vs. non-intelligent: $M = 0.125, SD = 0.055, t = 2.087, p < 0.05$), z(intelligent: $M = 0.034, SD = 0.025$ vs. non-intelligent: $M = 0.071, SD = 0.072, t = 2.151, p < 0.05$), and w(intelligent: $M = 2.38e-6, SD = 2.277e-6$ vs. non-intelligent: $M = 7.72e-6, SD = 1.153e-5, t = 2.032, p < 0.05$) also decreased when interacting with the intelligent avatar, pointing to more controlled and less random head movements. The reduced mean frequency in head rotation signals suggests a more focused and purposeful engagement with the intelligent avatar, crucial for maintaining spatial orientation and reducing the sensory mismatch that can contribute to cybersickness.

Furthermore, the decreases in mean frequency observed for acceleration y(intelligent: $M = 0.008, SD = 0.005$ vs. non-intelligent: $M = 0.019, SD = 0.016, t = 3.025, p < 0.05$) and z(intelligent: $M = 0.021, SD = 0.020$ vs. non-intelligent: $M = 0.049, SD = 0.059, t = 2.077, p < 0.05$) and EDA(intelligent: $M = 6.42e-4, SD = 4.508e-4$ vs. non-intelligent: $M = 0.001, SD = 8.485e-4, t = 3.076, p < 0.05$) indicate a more regulated physiological state when using the intelligent avatar. Reduced physiological arousal and better regulation of bodily responses are associated with lower levels of cybersickness, as physiological instability is often a contributing factor.

For other measurements, such as eye position, head position, eye rotation (x), and head rota-

tion (x), the mean frequency values did not show statistically significant differences between the two avatar conditions, suggesting that the complexity or randomness of these signals remained relatively consistent.

While frequency-domain analysis, such as the use of PSD, spectral entropy, and mean frequency measures, can provide valuable insights into the underlying characteristics of a dataset, there are potential limitations to relying solely on this approach. The frequency-domain analysis focuses on the distribution of the signal's energy across different frequencies, but it does not capture the temporal dynamics and evolution of the signal over time. This can be particularly problematic when analyzing data that exhibits non-stationary or time-varying properties, as the frequency-domain representation may only partially capture the rich complexity of the signal.

5.7 Time-Frequency Domain Analysis

Analyzing physiological signals and behavioral indicators in the time-frequency domain offers distinct advantages over traditional time or frequency-domain analysis. These approaches allow for the simultaneous examination of the signal's temporal and frequency-domain characteristics, providing a more detailed and comprehensive data exploration. The dynamic and evolving nature of physiological responses and behavioral patterns can be better captured and understood through time-frequency analysis, as it can reveal critical transient events or localized changes in the signal that conventional frequency-domain methods may overlook.

The spectrogram and wavelet transform methodologies were employed to leverage these benefits and gain deeper insights into the experimental data. These time-frequency analysis techniques are renowned and extensively utilized in the field.

5.7.1 Spectrogram

The spectrogram is a widely used tool in signal processing that allows us to analyze a signal's time-varying frequency content. It visually represents how the signal's frequency components change over time, with time on the horizontal axis and frequency on the vertical axis. The magnitude or power of the Fourier transform is typically represented using colors or shading, where brighter colors indicate higher signal power at a particular frequency and time, as presented in Figure 5.14. Several

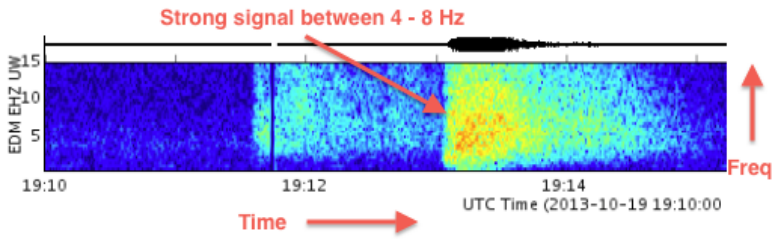


Figure 5.14: Brighter colors in the spectram represents higher signal power at a particular frequency and time.

computational methods are available to generate spectrograms, each with its characteristics and applications. Some common approaches include optical spectrometers, a bank of band-pass filters, the Short-Time Fourier Transform (STFT), and the wavelet transform.

In this thesis, the STFT-based and wavelet-based spectrogram techniques have been utilized to perform the time-frequency analysis of the experimental data. These powerful tools enable the researchers to uncover valuable insights into the temporal and spectral properties of the signals, potentially revealing patterns and features that may not be readily apparent in the raw time-domain or frequency-domain representations.

5.7.2 Short-Time Fourier Transform (STFT)

The Short-Time Fourier Transform (STFT) divides the signal into overlapping short segments and applies the Fourier Transform to each segment. This captures the signal's spectral components at different time intervals.

$$\text{STFT} = X[m,k] = \sum_{n=0}^{N-1} x[n] \cdot w[n - mR] \cdot e^{-j2\pi kn/N} \tag{5.7}$$

where

- $X[m,k]$ represents the STFT coefficient at time index m and frequency index k.
- $x[n]$ is the input signal.
- $w[n-mR]$ is the window function applied to the signal. It is typically a window of length N_w centered around mR with R being the hop size or overlap between consecutive windows.
- N is the FFT length, which determines the frequency resolution of the spectrogram.

- $e^{-j2\pi kn/N}$ represents the complex exponential term that provides the frequency modulation.

The STFT computes a set of complex values, X[m,k], which represent the signal's frequency content at different time-frequency bins. The magnitude or power of these complex values is often used to visualize the spectrogram.

5.7.3 Power of Spectrogram:

The power of a spectrogram quantifies the distribution of power or energy across both time and frequency in a signal. By examining the power values in the spectrogram, we can identify regions of high or low power that correspond to significant or negligible frequency content in the signal. This enables us to analyze the temporal and spectral characteristics of the signal more comprehensively, aiding in the detection of essential features, patterns, or changes in the underlying signal.

$$P(m,k) = |X(m,k)|^2 \tag{5.8}$$

where

- $P(m,k)$ represents the power of the STFT coefficient at time index m and frequency index k.
- $X(m,k)$ is the STFT coefficient at time index m and frequency index k.

According to Table B.5, the mean power of the spectrogram for eye velocity in the x(intelligent: $M = 9.02e-6, SD = 8.6e-6$ vs. non-intelligent: $M = 1.63e-5, SD = 1.7e-5, t = 2.042, p < 0.05$) and y (intelligent: $M = 1.40e-6, SD = 1.4e-6$ vs. non-intelligent: $M = 3.70e-6, SD = 3.8e-6, t = 2.050, p < 0.05$) directions, as well as eye angular velocity in the x(intelligent: $M = 1.08e-4, SD = 1.0e-4$ vs. non-intelligent: $M = 2.40e-4, SD = 2.6e-4, t = 2.006, p < 0.05$) and y(intelligent: $M = 3.52e-4, SD = 3.5e-4$ vs. non-intelligent: $M = 9.36e-4, SD = 9.8e-4, t = 2.111, p < 0.05$) directions, was significantly higher in the non-intelligent avatar condition compared to the intelligent avatar condition.

The higher spectral power, indicating higher mean frequencies and/or greater variability in the frequency domain, could have significant implications. It suggests that more frequent, rapid, or erratic eye movements and angular adjustments, as participants visually explored the non-intelligent avatar, could be a key factor in increased cybersickness. This finding could pave the way for more effective strategies to reduce cybersickness in virtual reality experiences.

The higher mean power of the EDA spectrogram in the non-intelligent avatar condition (intelligent: $M = 1.02e-3, SD = 3.0e-4$ vs. non-intelligent: $M = 3.35e-3, SD = 1.1e-3, t = 2.129, p < 0.05$) suggests increased physiological arousal and sympathetic nervous system activation. As mentioned, heightened physiological arousal has been associated with cybersickness, as it can reflect the body's stress response to the sensory conflicts and disorientation experienced during virtual reality exposure.

For the remaining measures, the differences in the mean power of the spectrogram between the two avatar conditions were not statistically significant. This reiterates that the frequency-domain characteristics of these other physiological and behavioral signals were relatively similar between the two avatar interaction scenarios, providing a clear picture of the overall findings.

Figure 5.15 and Figure 5.16 display a sample of the spectrogram of eye angular velocity in the x-axis recorded from a single participant during the non-intelligent driving simulator and intelligent avatar studies. In the non-intelligent avatar experiment, the participant's adaptation phase lasted for 2.5 minutes; however, due to experiencing cybersickness, he/she discontinued the experiment before completing the entire procedure, as illustrated in Figure 5.15.

In contrast, during the intelligent avatar experiment (Figure 5.16), the participant completed the entire 6-minute duration, demonstrating improved tolerance and adaptability. Remarkably, not only did the participant complete the experiment, but the average power of the spectrogram was also lower than that observed in the non-intelligent avatar study. This notable difference in completion time and spectrogram power highlights the participant's enhanced performance and engagement within the intelligent avatar environment.

The participant's successful completion of the intelligent avatar study signifies the potential of this simulation method to provide a more comfortable and engaging experience for individuals, leading to improved overall performance and outcomes.

5.7.4 Continuous Wavelet Transform (CWT)

The Continuous Wavelet Transform (CWT) is a powerful technique for decomposing a signal into wavelets, which are small oscillations highly localized in time. Unlike the Fourier Transform, which decomposes a signal into infinite-length sines and cosines, losing time-localization information, the CWT utilizes scaled and shifted versions of a time-localized mother wavelet as its basis functions.

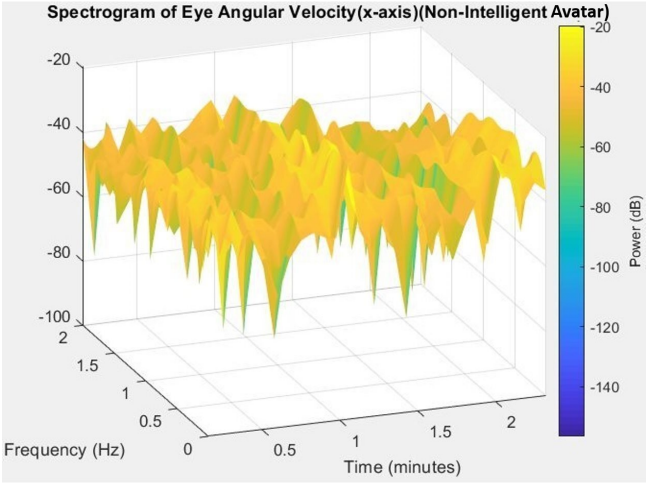


Figure 5.15: 3D Spectrogram of a participant's Eye Angular Velocity (x-axis) in Non-Intelligent Avatar.

This allows the CWT to provide a time-frequency representation of a signal with excellent time and frequency localization.

The CWT is particularly well-suited for analyzing non-stationary signals, such as physiological and behavioral signals, as it effectively captures their changing properties over time.

The definition of the CWT is as follows:

$$T(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\psi^* \left(\frac{t-b}{a}\right) dt \tag{5.9}$$

where

- a is scale (or dilation) parameter,
- b is location of wavelet,
- ψ is wavelet function,
- x is signal.

The wavelet function, which is also called the mother wavelet, is a function that is continuous in both time and frequency. It is the fundamental building block from which scaled and translated basis functions are derived. The mother wavelet can be either complex or real, and it often includes an adjustable parameter that allows for control over the properties of the localized oscillation.

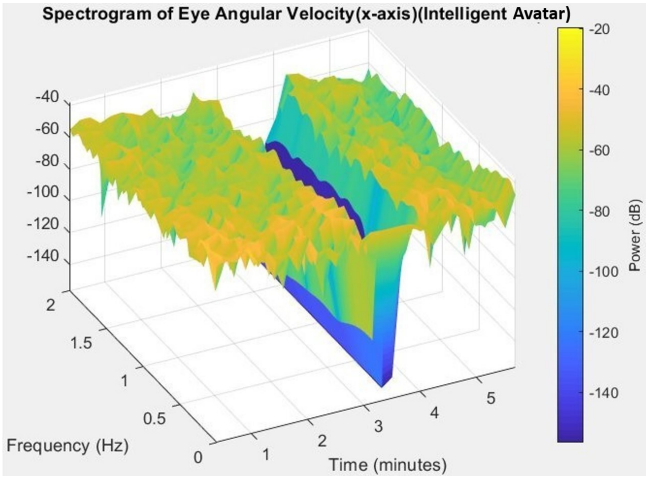


Figure 5.16: 3D Spectrogram of the same participant's Eye Angular Velocity (x-axis) in Intelligent Avatar.

Morlet Wavelet: The Morlet wavelet is the most commonly used in CWT analysis as a wavelet function. It is a Gaussian-windowed complex sinusoid that is defined in both the time and frequency domains as follows:

$$\psi(t) = \frac{1}{\sqrt{\pi}} e^{i\omega_0 t} e^{-\frac{t^2}{2\sigma^2}} \tag{5.10}$$

where

- $\psi(t)$ is Morlet wavelet function,
- ω_0 is the central frequency of the wavelet,
- σ is the standard deviation of the Gaussian envelope.

Figure 5.17 illustrates the real and complex components of a Morlet wavelet with a ω_0 set to 6.

The Morlet wavelet was utilized in the data analysis to examine the data across a range of scales from 1 to 64, with a step size 0.5. Additionally, the parameters ω_0 and σ were set with values 6 and $\frac{6}{2\pi}$, respectively. This approach facilitated the identification of structures at various frequencies, enabling a comprehensive data analysis.

Wavelet energy was utilized as a metric to assess energy or power distribution across various scales within each signal to compare the intelligent and non-intelligent objective signals. This metric highlights the relative contributions of different scales to the signal's overall energy.

$$\text{Wavelet Energy} = \sum |C(i,j)|^2 \tag{5.11}$$

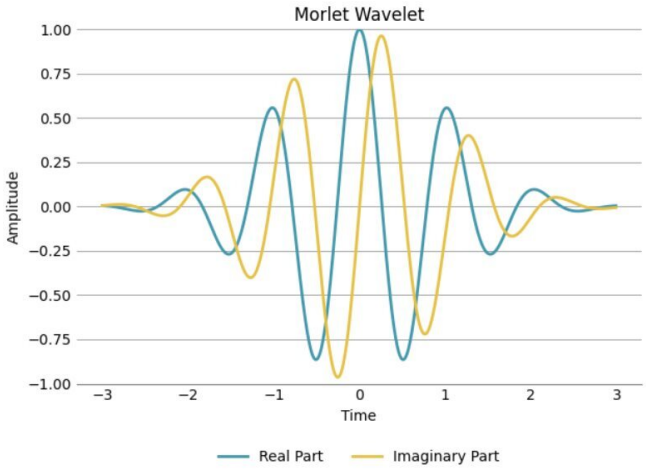


Figure 5.17: The Morlet wavelet with an initialization of $\omega_0 = 6$.

$C(i, j)$ is wavelet coefficient at scale i and position j.

The objective of applying the wavelet transform to the collected physiological and behavioral data was to uncover temporal patterns, identify relevant frequency bands, and explore the dynamic relationships between physiological responses and the occurrence of cybersickness.

Based on the data analysis using wavelet energy provided in Table B.7, significantly lower wavelet energy in the intelligent avatar condition compared to the non-intelligent avatar condition for the following measures:

- Eye Velocity
  - x(intelligent: $M = 0.210, SD = 0.089$ vs. non-intelligent: $M = 0.439, SD = 0.208, t = 2.016, p < 0.05$)
  - y(intelligent: $M = 0.019, SD = 0.008$ vs. non-intelligent: $M = 0.047, SD = 0.019, t = 2.080, p < 0.05$)
  - z(intelligent: $M = 0.052, SD = 0.023$ vs. non-intelligent: $M = 0.138, SD = 0.049, t = 2.113, p < 0.05$)
- Eye Angular Velocity
  - x(intelligent: $M = 1.543, SD = 0.650$ vs. non-intelligent: $M = 4.185, SD = 1.736, t = 2.193, p < 0.05$)

  - y(intelligent: $M = 10.412, SD = 4.794$ vs. non-intelligent: $M = 25.056, SD = 11.947, t = 2.240, p < 0.05$)

- Eye Position

  - x(intelligent: $M = 1.110, SD = 0.960$ vs. non-intelligent: $M = 1.852, SD = 1.448, t = 2.236, p < 0.05$)
  - y(intelligent: $M = 67.402, SD = 59.686$ vs. non-intelligent: $M = 70.323, SD = 62.188, t = 2.049, p < 0.05$)

- Eye Rotation

  - y(intelligent: $M = 9.844, SD = 8.419$ vs. non-intelligent: $M = 19.680, SD = 15.488, t = 2.272, p < 0.05$)

- Head Position

  - x(intelligent: $M = 1.050, SD = 0.906$ vs. non-intelligent: $M = 1.832, SD = 1.417, t = 2.256, p < 0.05$)
  - y(intelligent: $M = 67.403, SD = 59.686$ vs. non-intelligent: $M = 70.323, SD = 62.188, t = 2.048, p < 0.05$)

- Head Rotation

  - x(intelligent: $M = 2.089, SD = 1.788$ vs. non-intelligent: $M = 2.630, SD = 2.063, t = 2.187, p < 0.05$)
  - y(intelligent: $M = 10.322, SD = 8.872$ vs. non-intelligent: $M = 18.869, SD = 14.852, t = 2.019, p < 0.05$)
  - w(intelligent: $M = 52.083, SD = 46.053$ vs. non-intelligent: $M = 52.476, SD = 46.462, t = 2.392, p < 0.05$)

- Acceleration

  - x(intelligent: $M = 332.569 \times 10^3, SD = 246.672 \times 10^3$ vs. non-intelligent: $M = 645.714 \times 10^3, SD = 471.469 \times 10^3, t = 2.672, p < 0.05$)

- y(intelligent: $M = 126.280 \times 10^3, SD = 92.058 \times 10^3$ vs. non-intelligent: $M = 287.560 \times 10^3, SD = 215.842 \times 10^3, t = 2.140, p < 0.05$)
- z(intelligent: $M = 593.302 \times 10^3, SD = 455.958 \times 10^3$ vs. non-intelligent: $M = 1315.046 \times 10^3, SD = 975.146 \times 10^3, t = 2.367, p < 0.05$)

- EDA
  - (intelligent: $M = 203.696, SD = 173.553$ vs. non-intelligent: $M = 757.456, SD = 616.653, t = 2.561, p < 0.05$)
- BVP
  - (intelligent: $M = 1050.880 \times 10^3, SD = 1786.581 \times 10^3$ vs. non-intelligent: $M = 3444.041 \times 10^3, SD = 11918.320 \times 10^3, t = 2.156, p < 0.05$)

The reduced wavelet energy for these parameters suggests decreased variability and more controlled/stable eye, head, and physiological responses. Since increased variability in these signals is associated with cybersickness, the lower wavelet energy in the intelligent avatar condition indicates a potential reduction in cybersickness susceptibility.

TMP was the only measure that did not show a significant difference between conditions, suggesting that the presence of the intelligent avatar did not substantially affect this parameter.

The results indicate that the intelligent avatar condition elicited more stable and controlled eye, head, and physiological responses than the non-intelligent one. These findings suggest that the intelligent avatar helped reduce cybersickness symptoms by promoting more stable and less variable oculomotor and physiological responses.

Figure 5.18 (a) depicts the wavelet transform obtained during the non-intelligent avatar experiment for a participant and its corresponding energy distribution (Figure 5.18 (b)). Similarly, Figure 5.19 (a) showcases the wavelet transform acquired during the intelligent avatar experiment for the same participant and its corresponding energy distribution (Figure 5.19 (b)). The figure vividly demonstrates the noticeable disparity in wavelet coefficient energy between the two studies' wavelet transforms. This disparity highlights the significant differences observed in the characteristics of the wavelet transform energy, further emphasizing the distinctive nature of the non-intelligent and intelligent avatar sessions.

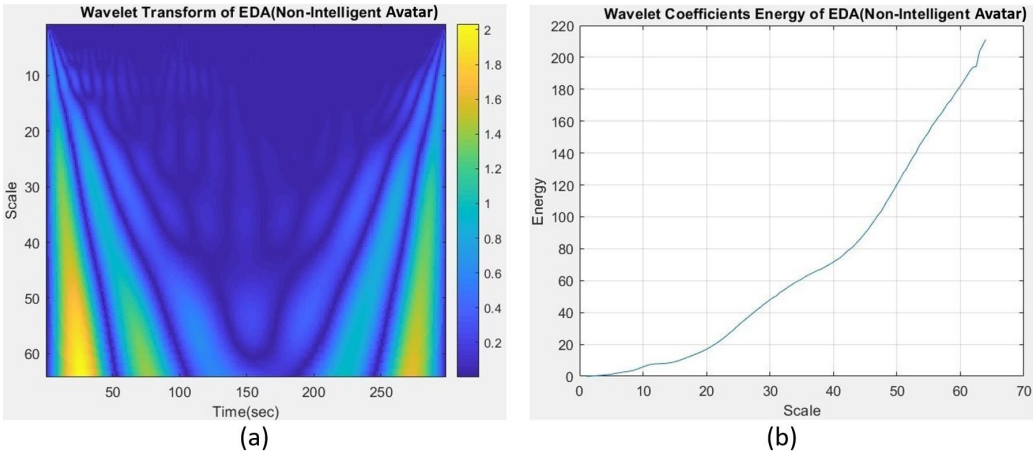


Figure 5.18: Sample of Wavelet Transform (a) and its coefficient energy (b) of a participant's EDA signal during the non-intelligent avatar experiment.

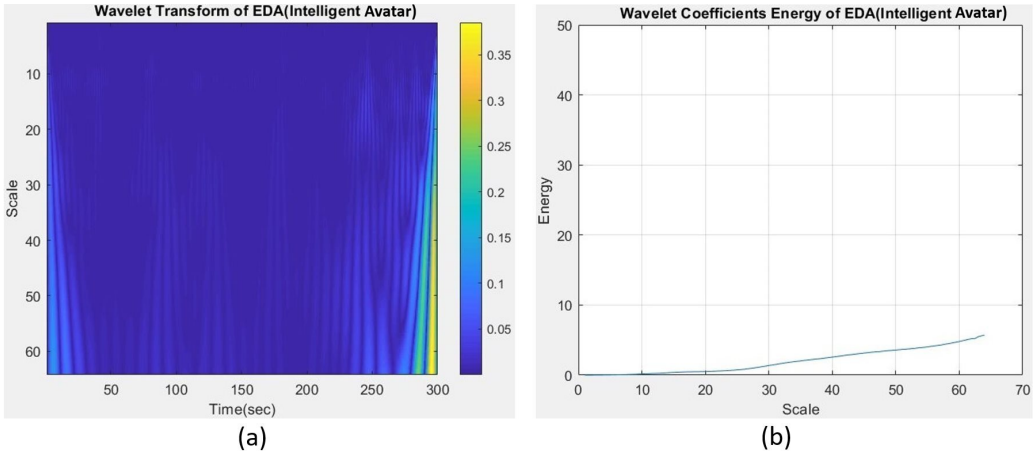


Figure 5.19: Sample of Wavelet Transform (a) and its coefficient energy (b) of the same participant's EDA signal during the intelligent avatar experiment.

5.8 Advantages of Wavelet Transform as the Superior Analysis

Based on the analyses (Table B.1 to Table B.7), wavelet energy representation was more effective in capturing differences between physiological and behavioral signals in the non-intelligent and intelligent avatar studies. This analysis identified significant differences in 72% of the signals, surpassing other analyses (see Figure 5.20). These results clearly demonstrate the enhanced sensitivity of wavelet energy in detecting variations associated with avatar conditions.

It is important to note that traditional time-domain metrics like mean, standard deviation, and RMS, while useful, provide a relatively coarse and summary-level characterization of the signal proper-

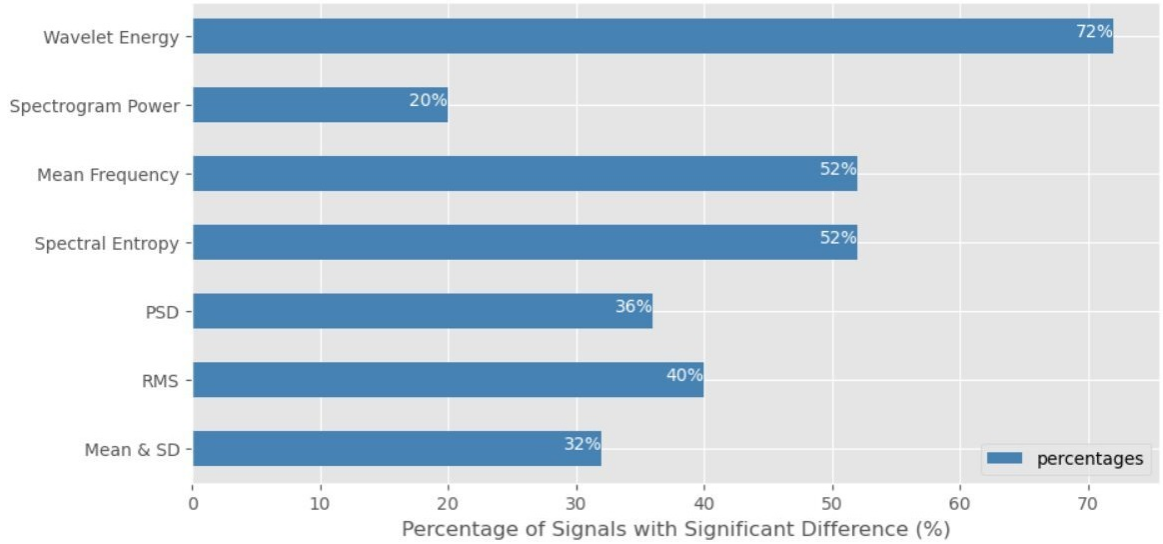


Figure 5.20: Comparison of different analysis metrics in their capability to detect significant changes in recorded physiological and behavioral measurements between intelligent and non-intelligent avatar conditions.

ties. These simple statistics capture the central tendency and dispersion of the time-domain waveform. However, they fall short in capturing the rich, time-varying dynamics inherent in the physiological and behavioral signals, highlighting the need for a more advanced approach.

Unlike traditional metrics, wavelet energy analysis takes a multi-scale, adaptive approach to quantifying the signal characteristics. By decomposing the waveform into constituent frequency bands and tracking the energy distribution across these scales over time, the wavelet analysis is able to reveal much more precise information about the non-stationary signal properties. This includes identifying transient events, localized frequency changes, and other time-varying phenomena that would be obscured or averaged out in the basic time-domain statistics, demonstrating its superiority.

Moving to the frequency domain, metrics like PSD, spectral entropy, and mean frequency provide a global, stationary representation of the signal's spectral composition. However, these frequency domain analyses also need more time-localization capabilities to capture the evolving spectral characteristics of non-stationary signals.

On the other hand, the wavelet energy spectrum retains the time dimension, allowing it to track how signal energy distribution across frequencies changes over time. This time-frequency representation is crucial for identifying and differentiating our study's dynamic, context-dependent spectral signatures

associated with the distinct avatar conditions.

Compared to the time-frequency analysis based on the STFT, such as the power of the spectrogram, the wavelet energy approach offers several key advantages. The STFT uses a fixed-size window, which means the time-frequency resolution is constant across all frequencies. However, this approach has limitations in capturing non-stationary signals' disparate time and frequency characteristics, highlighting the need for a more advanced approach like wavelet analysis.

The STFT-based spectrogram can also suffer signal distortion, particularly for high-frequency components, due to the inherent trade-off between time and frequency resolution. Wavelet analysis can better preserve the original signal characteristics, leading to a more faithful and informative time-frequency representation.

In this study, the superior performance of the wavelet energy features across the time domain, frequency domain, and time-frequency domain analyses suggests that this approach was able to extract more discriminative and meaningful information from the physiological and behavioral signals. The wavelet-based metrics likely captured subtle, context-dependent patterns that the more conventional analysis techniques missed. This allows us to better differentiate the responses to the non-intelligent and intelligent avatar conditions.

5.9 Subjective Data Analysis

In addition to the objective physiological and oculomotor measurements, the subjective user experience was also evaluated through the Simulator Sickness Questionnaire (SSQ). For each experiment, we computed the difference between the pre-exposure and post-exposure SSQ scores, which provides an overall indication of the change in simulator sickness symptoms experienced by participants.

$$SSQ = \mathrm{SSQ}_{\mathrm{post}} - \mathrm{SSQ}_{\mathrm{pre}} \tag{5.12}$$

As shown in Figure 5.21, the results reveal substantial differences in the SSQ score changes between the two experimental conditions. In the intelligent avatar experiment, we observed notable reductions in the critical sub-scales of the SSQ compared to the non-intelligent avatar condition. Specifically, the Nausea sub-scale score decreased by 43.4% ($t = 2.06, p < 0.05$) in the intelligent avatar condition compared to the non-intelligent avatar. This substantial reduction in nausea-related symptoms indicates

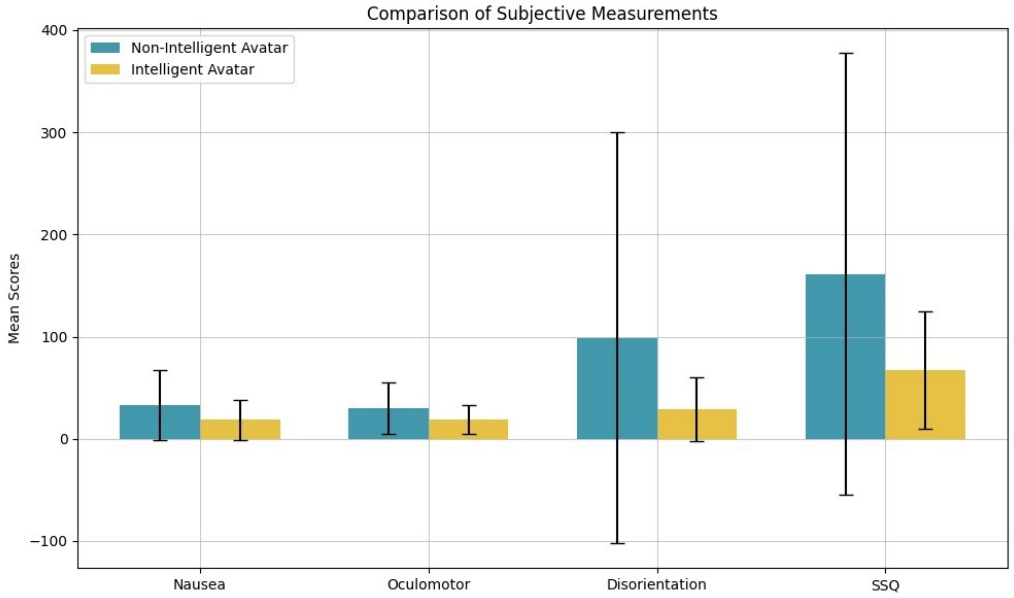


Figure 5.21: Comparison of Subjective Measurements between Non-Intelligent and Intelligent Avatar Sessions

that the intelligent avatar design was much less provocative and better able to maintain user comfort and stability.

Similarly, the Oculomotor sub-scale score decreased by 35.1% ($t = 2.01, p < 0.05$) in the intelligent avatar experiment. This suggests that the intelligent avatar was less taxing on the visual-vestibular system, leading to reduced eye strain and blurred vision issues.

Moreover, the Disorientation sub-scale showed a significant 70.6% ($t = 1.99, p < 0.05$) reduction in the intelligent avatar condition. This substantial decrease in vertigo and spatial disorientation symptoms highlights the potential of the intelligent avatar to provide users with a more coherent and stabilizing sensory experience.

When looking at the overall SSQ score, which aggregates all simulator sickness symptoms, the intelligent avatar experiment had a 44.3% decrease compared to the non-intelligent avatar. This significant 58.5% ($t = 2.15, p < 0.04$) reduction in the total SSQ score strongly indicates the intelligent avatar's effectiveness in mitigating cybersickness across multiple domains.

The magnitudes of these decreases, ranging from 33% to nearly 50%, demonstrate the substantial improvements in user comfort and well-being that can be achieved by implementing intelligent avatar

technology. These results suggest that such advanced virtual representations have great potential to enhance the overall user experience in immersive digital environments and overcome critical barriers to widespread VR/AR adoption.

5.10 Convergence of Objective and Subjective Measures of Intelligent Avatar Performance

Based on our comprehensive, objective data analysis, focusing specifically on the wavelet energy metrics, the results demonstrate significant decreases across a wide range of physiological and behavioral measures in the intelligent avatar experiment compared to the non-intelligent avatar.

Section 5.8 discusses how wavelet energy analysis has proven to be a robust and adaptable tool for examining human-avatar interactions' complex, non-stationary dynamics. Our comprehensive data analysis in the time, frequency, and time-frequency domains has identified several objective measures that could potentially serve as signatures of cybersickness. Notably, we observed a significant decrease in eye velocity (x and y), eye angular velocity (x and y), and EDA for the intelligent avatar condition compared to the non-intelligent avatar in all analyses. These findings suggest that these physiological signals could be sensitive indicators of the cybersickness experience.

Our correlation analysis has further strengthened the robustness of our findings. By calculating the correlation between the wavelet energy of the eye velocity, eye angular velocity, and EDA signals with the overall SSQ scores, we have uncovered strong positive correlations.

Specifically, our findings indicate that eye velocity in both the x and y dimensions exhibits a robust positive correlation with SSQ scores. As illustrated in Figure 5.22, for eye velocity in the x-axis, the regression equation $SSQ = -8.653 + 385.91 \times EyeVelocity(x)$ was obtained, with a highly significant correlation coefficient of $r = 0.97$.

Similarly, for the y-axis eye velocity, the relationship $SSQ = 23.567 + 1620.1 \times EyeVelocity(y)$ was observed, with an equally strong correlation of $r = 0.95$. These results suggest that changes in gaze behavior, as captured by eye velocity, are closely linked to the subjective experience of cybersickness. The data visualization in Figure 5.23 confirms the correlation between the measured variables.

Extending the analysis to eye angular velocity, we again found compelling relationships with the SSQ scores. The correlation equation $SSQ = 25.950 + 21.568 \times EyeAngularVelocity(x)$ was derived

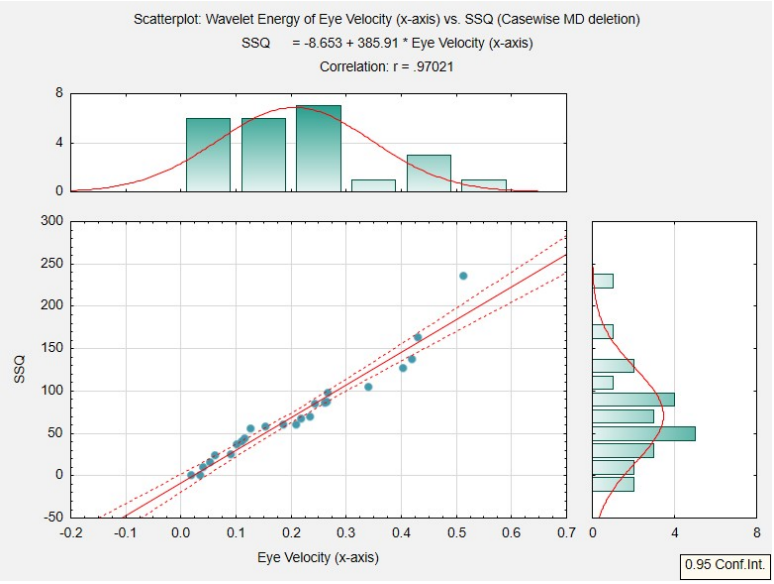


Figure 5.22: Correlation between the SSQ score and wavelet energy of Eye Velocity(x) during Intelligent Avatar experiment.

for the x-axis angular velocity, with a correlation coefficient of $r = 0.78$(see Figure 5.24).

Similarly, for the y-axis angular velocity, the equation $SSQ = 11.282 + 4.6871 \times EyeAngularVelocity(y)$ was obtained, with an even stronger correlation of $r = 0.98$. These findings indicate that alterations in eye movements, as reflected in both velocity and angular velocity, serve as sensitive biomarkers of cybersickness. Figure 5.25 depicts the regression line that models the relationship between the variables.

Furthermore, the analysis of EDA also revealed a significant positive correlation with SSQ scores, as expressed by the equation $SSQ = 54.564 + 0.2489 \times EDA$, with a correlation coefficient of $r = 0.80$ (see Figure 5.26). This suggests that changes in sympathetic nervous system arousal, as measured by EDA, are closely tied to the subjective experience of cybersickness. Taken together, these robust correlations between objective physiological measures and the subjective cybersickness reports provide converging evidence for the validity and reliability of these markers as indicators of Intelligent Avatar performance.

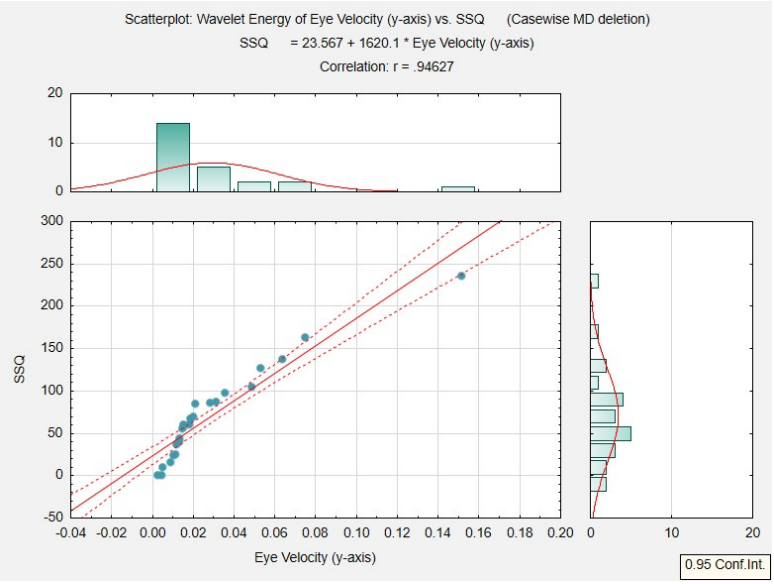


Figure 5.23: Correlation between the SSQ score and wavelet energy of Eye Velocity(y) during Intelligent Avatar experiment.

5.11 Topological Data Analysis (TDA)

In the era of big data, extracting meaningful insights and understanding the underlying structure of complex datasets has become a paramount challenge. Topological Data Analysis (TDA) is a relatively new field that has emerged from various works in applied (algebraic) topology and computational geometry, with persistent homology playing a central role [265][60]. While geometric approaches to data analysis have a long history, TDA indeed took shape as a field with the pioneering works of Edelsbrunner et al. [59] and Zomorodian and Carlsson [266] in persistent homology laid the foundation for TDA. It was initially popularized by Carlsson[36] and has its roots in the fields of topology [89], linear algebra [225], and graph theory [247].

At its core, TDA aims to provide rigorous mathematical, statistical, and algorithmic methods to infer, analyze, and exploit the complex topological and geometric structures underlying data. Instead of operating solely in the Euclidean space, TDA accommodates data represented as point clouds in more general metric spaces, enabling a more flexible and comprehensive analysis.

The key idea driving TDA is that topological features, such as connected components, loops, voids, and higher-dimensional structures, can capture essential information about the data's global and local structure. By systematically measuring and quantifying these topological features, TDA enables the

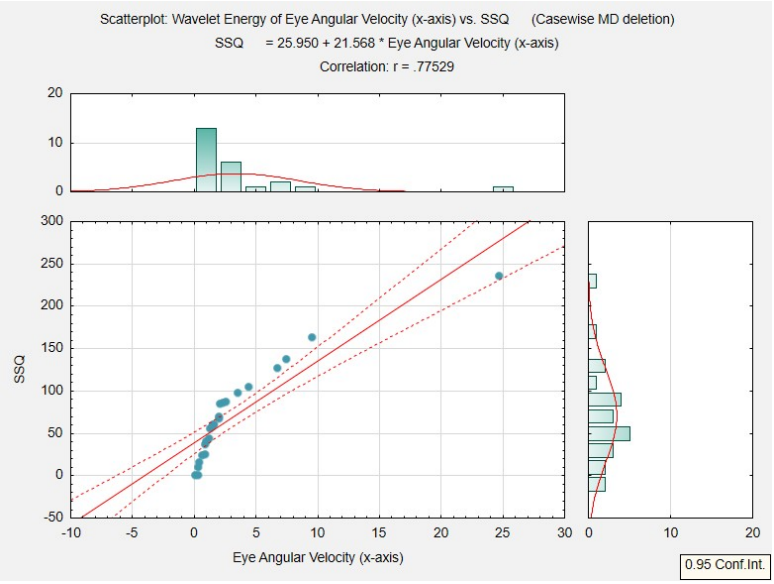


Figure 5.24: Correlation between the SSQ score and wavelet energy of Eye Angular Velocity(x) during Intelligent Avatar experiment.

extraction of robust and meaningful insights that might otherwise be obscured by noise or traditional statistical measures.

Over the past years, significant progress has been made in developing efficient algorithms, data structures, and software libraries, making TDA more accessible to researchers and practitioners. Notable examples include the Gudhi library, available in C++ and Python [139], and its R software interface [65], which provides user-friendly implementations of TDA techniques, giotto-tda [230] as a Python library to integrate TDA and machine learning, and TTK [233] an open-source library available in both Python and C++ and also visualization to the end user.

Despite being a rapidly evolving field, TDA has already demonstrated its potential in various domains, including biology [8][179][164], neuroscience [209][81][76], image analysis [208][22][15], chemistry [216], and materials science [234][118][157]. TDA empowers researchers and data scientists to uncover novel insights, make more informed decisions, and gain a deeper understanding of complex datasets by offering a complementary approach to traditional data analysis methods.

Persistent homology is used within TDA to capture and quantify the topological features present in the data. It focuses on tracking the birth and death of these features as a parameter, such as distance or scale, is varied. It provides a way to understand the evolution and persistence of topological structures

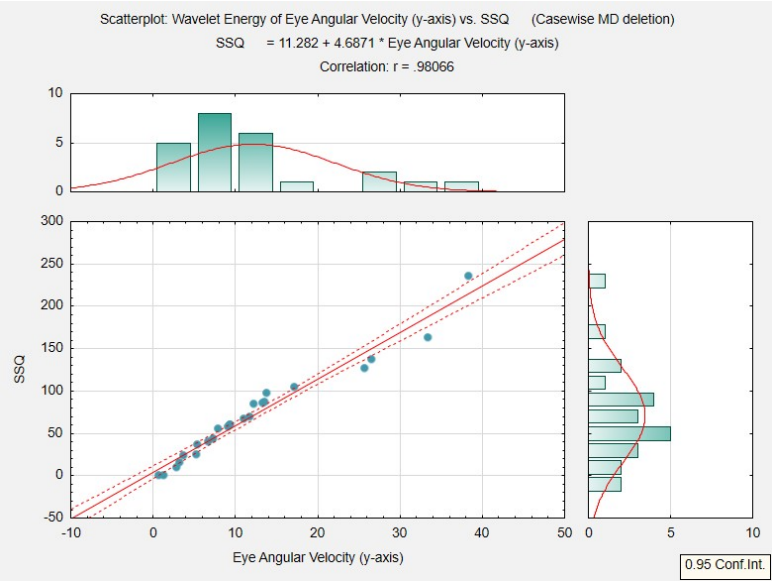


Figure 5.25: Correlation between the SSQ score and wavelet energy of Eye Angular Velocity(y) during Intelligent Avatar experiment.

across different scales or thresholds. By analyzing the persistence of these features, researchers can gain insights into the robustness and significant patterns within the data.

Persistent homology is a fundamental tool within TDA. However, TDA extends beyond persistent homology and encompasses a broader array of techniques and methodologies. TDA leverages not only persistent homology but also other complementary tools to analyze, interpret, and extract valuable insights from the acquired topological information. In essence, while persistent homology plays a pivotal role, TDA incorporates a more comprehensive suite of approaches to explore and understand the topological properties of complex datasets.TDA's ability to leverage persistent homology and capture the underlying topological structure of data opens up new avenues for analysis and interpretation in diverse fields.

5.11.1 Introduction to TDA

Here, we delve into the diverse applications of TDA, showcasing its broad utility in various domains.

TDA offers a powerful approach to analyzing and visualizing complex, high-dimensional data sets [39]. Traditional statistical techniques often rely on assumptions of linearity, normality, and independence, which may not hold in the case of complex data sets. Conversely, TDA embraces

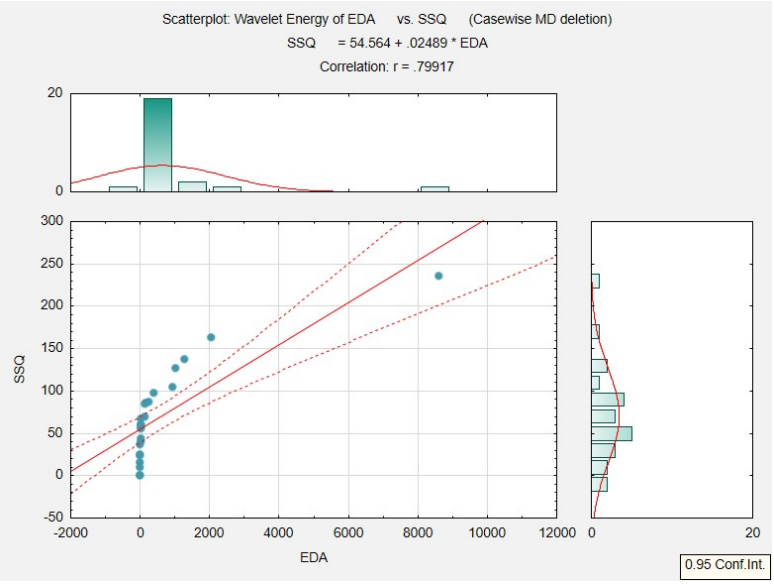


Figure 5.26: : Correlation between the SSQ score and wavelet energy of EDA during Intelligent Avatar experiment.

the inherent complexity of the data and captures its topological properties. It allows for a more comprehensive understanding of the data by considering the relationships and interactions among data points.

One of the most reassuring advantages of TDA is its ability to handle high-dimensional data sets. As the number of variables or dimensions increases, the complexity of the data grows exponentially, making traditional analysis methods less effective. TDA addresses this challenge head-on by focusing on the essential topological features of the data, disregarding irrelevant details. By reducing the dimensionality of the data to its topological essence, TDA reveals patterns and structures that are otherwise hidden in the high-dimensional space.

TDA employs various techniques to analyze and visualize data. One fundamental concept in TDA is constructing a simplicial complex, a mathematical representation of the data as a topological space. This simplicial complex captures the relationships between data points and encodes the connectivity and proximity of the data in a geometric sense. By examining the simplicial complex, TDA can identify clusters, voids, loops, and higher-dimensional structures that characterize the data.

Visualization is not just a tool but a key to unlocking the understanding of complex data sets, and TDA offers several powerful visualization techniques. One popular method is the Mapper algorithm,

which provides a visual summary of the data by creating a graph representation of the simplicial complex [206]. Each node in the graph represents a subset of the data, and the edges indicate shared data points between subsets. This visualization identifies clusters, outliers, and transitions between different data regions, enabling a deeper understanding of its underlying structure.

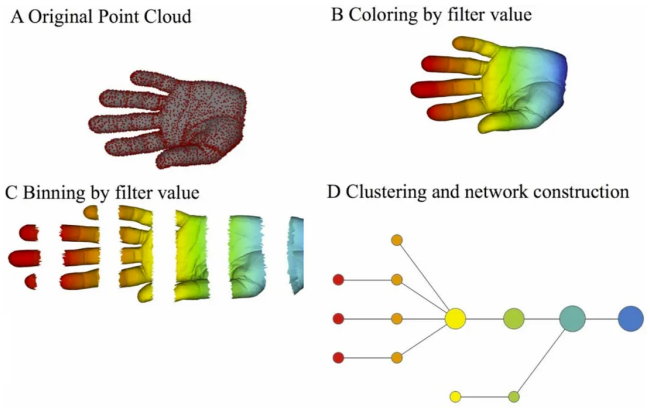


Figure 5.27: Mapper Algorithm [135].

Additionally, TDA offers a robust feature extraction and selection approach from complex data sets. TDA can identify informative features relevant for subsequent analysis or classification tasks by capturing topological structures. By considering the topological significance of each feature, TDA can evaluate its relevance and discard redundant or irrelevant features. This feature selection process reduces dimensionality, simplifies the data representation, and enhances interpretability and computational efficiency. With a lower number of features, subsequent analysis or classification tasks become more tractable and require fewer computational resources. This is particularly valuable when working with large-scale or high-dimensional data sets.

Clustering and classification are fundamental tasks in data analysis and machine learning. TDA offers a unique and powerful approach for classification [113] and clustering by leveraging the topological structures inherent in the data. Through analyzing these structures, TDA can effectively assign data points to specific classes or group similar data points into clusters, providing valuable insights and predictive capabilities.

The application of TDA for classification is relevant in various domains. In image analysis, TDA has been used to classify images based on their visual features, enabling tasks such as object recognition or image categorization [74][62][43]. In Natural Language Processing (NLP), TDA has assisted

in classifying text documents based on their linguistic features, enabling tasks such as sentiment analysis or topic classification. For instance, novel approaches like term frequency–term discrimination ability (TF-TDA) have been proposed to enhance feature representation in sentiment analysis tasks, showcasing significant improvements in classification performance [7]. Additionally, Alyasi et al. [63] introduced text classification and visualization for NLP using TDA, focusing on Persistent Homology and Mapper methods to classify Persian poems by "Ferdowsi" and "Hafez". These advancements highlight natural language processing technologies' continuous evolution and maturation for efficient text classification and analysis. Moreover, in the field of neuroscience, Ferrà et al. [67] proposed a TDA classifier in electroencephalogram (EEG) data analysis that used persistence diagrams to quantify topological features, focusing on shape over variance, providing insights into dataset structure and intrinsic dimensionality.

Clustering with TDA involves identifying data points with similar topological properties. TDA captures the connectivity and proximity of data points in a geometric sense, allowing for the identification of clusters based on their topological structures. By constructing a simplicial complex or a network representation of the data, TDA can reveal clusters as connected components or densely connected regions in the topological space. For example, Ohanuba et al. [167] proposed an automated method that utilizes TDA and machine learning clustering to identify floods (ARs) in big data. This approach has the potential to significantly reduce the impact of disasters caused by extreme events like floods and contribute to achieving Sustainable Development Goals (SDGs) related to flood management. Using spatial data, the researchers employed K-means clustering with k=2, achieving an 80% efficiency in identifying potential flood zones in Nigeria. They validated their results using the Silhouette coefficient, a cluster quality measure. These examples demonstrate the versatility of TDA in handling diverse types of data and improving classification and clustering performance.

TDA is also helpful in detecting anomalies [165][50] or outliers [101] in data sets. TDA can identify points that deviate significantly by comparing the topological structure of a given data point to the overall structure of the data set. This can be valuable in various domains, such as fraud detection [149][236] or network security [151][188].

TDA has been applied in shape recognition and image analysis tasks by representing shapes or images as topological spaces. TDA can capture their intrinsic topological properties by representing shapes or images as topological spaces. This enables the recognition of shapes based on their global

or local topological features and facilitates tasks such as object recognition [196][195] and image classification [43][74][207].

For the analysis of time series data, TDA plays a crucial role in capturing the temporal evolution of the underlying topological structure. This not only reveals patterns, trends, and relationships in time-dependent datasets but also aids in forecasting, anomaly detection, and most importantly, understanding the dynamics of complex systems. A review of TDA for time series has been done by Ravishanker et al. [182] with examples using R functions, finding features derived from TDA are useful in classification and clustering of time series and in detecting breaks in patterns. Additionally, Rivera-Castro et al. [187] focused on using TDA for business-to-business (B2B) customer relationship management, showcasing its potential in understanding customer loyalty and enhancing predictive models with commercial data sets.

In this study, in addition to the extensive analyses conducted in the time, frequency, and time-frequency domains, we have also leveraged the power of persistent homology to uncover the underlying topological structures within the dataset. Persistent homology is a powerful tool in topological data analysis that allows us to extract meaningful insights by studying the intrinsic shape and connectivity of the data.

By applying persistent homology, we could identify persistent topological features that remained stable across multiple scales of observation. These topological signatures gave us a complementary perspective on the complex, non-linear dynamics governing the system under study. Through the lens of persistent homology, we could detect and quantify the emergence of higher-dimensional structures and connectivity patterns that would have been difficult to discern using traditional statistical or signal-processing techniques alone.

5.11.2 Definitions

- **k-Simplex:**

  A k-simplex is the $k-$dimensional convex hull of $(k+1)$ vertices. The vertices are a set of geometrical independent points in $\mathbb{R}^k$, like $v_0, \ldots, v_k$. The $k-$simplex σ spanned by $v_0, \ldots, v_k$ is defined as the set of all points $x \in \mathbb{R}^k$ such that:

$$0 \leq i \leq k, \quad \alpha_i \in \mathbb{R}, \quad x \in \sum_{i=0}^{k} \alpha_i \mathbf{v}_i \tag{5.13}$$

where $\sum_{i=0}^{k} \alpha_i = 1$. The number k is called the dimension of σ, and any simplex spanned by a nonempty subset of the $(k+1)$ vertices is called a face of σ.

- **Simplicial Complex:**

  A simplicial complex C_* is a set of simplices such that:

  - every face of a simplex is also a simplex of C_*;
  - the intersection of any two simplices C_1 and C_2 in C_* is either a face of both C_1 and C_2, or the empty set.

- **Vietoris-Rips Complex:**

  Let $\chi = \{x_1, x_2, \ldots, x_n\}$ be a collection of points in $\mathbb{R}^d$ and let $\varepsilon > 0$. The Vietoris-Rips complex $R_\varepsilon(\chi)$ is constructed as follows:

  - The $0-$simplices (vertices) are the points in χ.
  - A $k-$simplex $[x_{i_0}, \ldots, x_{i_k}]$ is in $R_\varepsilon(\chi)$ if $\|x_{i_j} - x_{i_l}\| \leq \varepsilon$ for all $0 \leq j, l \leq k$

  Figure 5.28 represents the Vietoris-Rips complex $R_\varepsilon(\chi)$.

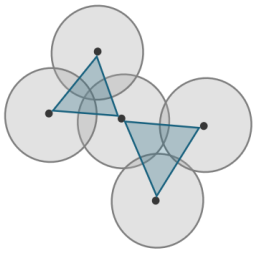


Figure 5.28: Vietoris-Rips complex $R_\varepsilon(\chi)$. There is a set of vertices and radius ε. Since all the pairwise intersections occur, the Vietoris-Rips complex does include the corresponding face.

- **Filtrations :**

  A filtration of a simplicial complex F is a nested family of subcomplexes F, where $T \subseteq \mathbb{R}$, such that for any $\varepsilon_1, \varepsilon_2 \in T$, if $\varepsilon_1 \leq \varepsilon_2$ then $F_{\varepsilon_1} \subseteq F_{\varepsilon_2}$ and $F = \bigcup_{\varepsilon \in T} F_\varepsilon$. The subset T may beeither finite or infinite (see Figure 5.29).

  Since we need to know which homology classes live for a small range of ε, and which homology classes persist longer, the "persistent" features of the homology of the data at different scales of should be investigated.

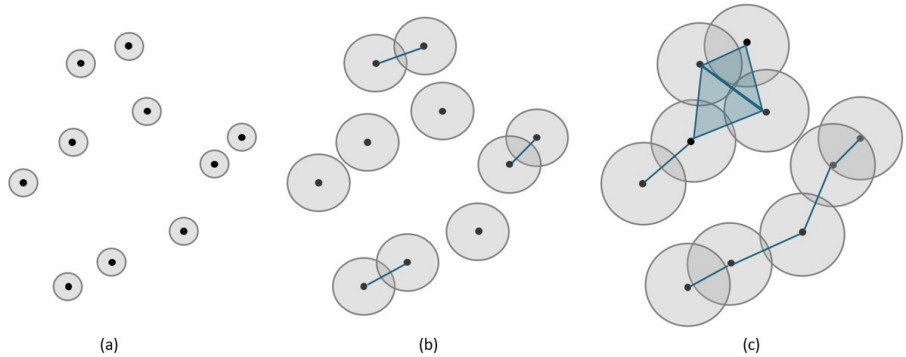


Figure 5.29: Example of filtration varying the filtration value ε which increased from (a) to (c). The black dot represents the point cloud data that are connected (blue line) when the $\varepsilon-$balls around them overlap. The top part of (c) is the union of two adjacent triangles.

- **Persistent Homology:**

  The $k-$th persistent homology of F is the sequence of $k-$th homology groups connected by homomorphisms obtained from a filtration F.

  Consider a finite filtration $\{F_\varepsilon|\varepsilon \in I\}$ of simplicial complexes. By definition, $\varepsilon_i \leq \varepsilon_j$ implies $F_{\varepsilon_i} \subseteq F_{\varepsilon_j}$. This relation yields a linear map between the homology groups $H_n(F_{\varepsilon_i}) \to H_n(F_{\varepsilon_j})$ for arbitrary k [40]. Moreover, there is an interval [b,d] where b is the minimum index where a $H_n(F_{\varepsilon_b})$ is non-zero, and d is the minimum index after which $H_k(F_{\varepsilon_d})$ are all 0. The numbers b and d are called the birth time and the death time of the corresponding homological features.

  Let $\mu_n^{(b,d)}$ denote the number of $n-$homology classes born at $H_n(F_b)$ and dying entering $H_n(F_d)$. Then, the $n-$th persistence diagram of a filtration F, denoted by $D_n(F)$, is the multiset of points (b, d) with multiplicity $\mu_n^{(b,d)}$ (together with the points of the diagonal with infinity multiplicity by convention). Figure 5.30 depicts a sample of persistence diagram.

5.11.3 Analysis of TDA Results in the Current Use Case

In our analysis, we have used the persistent homology value, or persistent value for short, as a metric to quantify the persistence or lifespan of topological features within the data. We then computed the average of this persistent value metric for each objective data set.

$$\text{persistent value} = d - b \tag{5.14}$$

The persistent homology values, represented in Table B.8, provide insights into the stability and

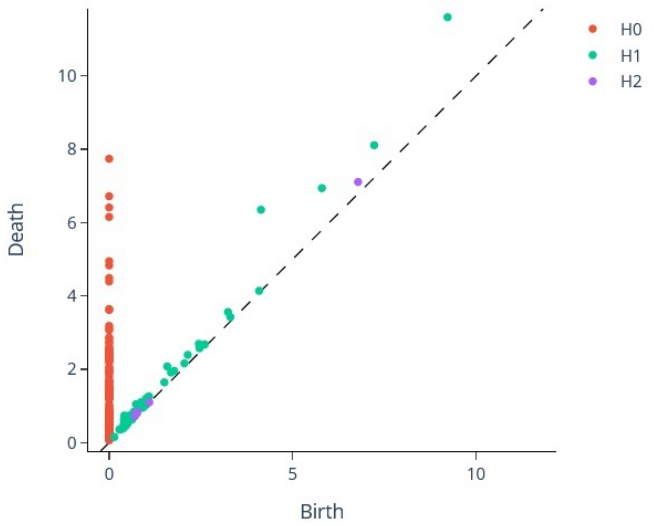


Figure 5.30: The corresponding persistence diagram with $H_0(x)$ in red and $H_1(x)$ in green, , and $H_2(x)$ in purple representing persistence of connected components.

variability of various physiological and behavioral signals measured in the non-intelligent and intelligent avatar studies. Due to the large volume of data, the table includes only the measurements that show significant differences ($p < 0.05$). Higher persistent values indicate more consistent and less variable signals, while lower values suggest greater instability and fluctuations over time.

The Intelligent avatar condition had significantly lower persistent values across the board for eye velocity and eye angular velocity in the x, y, and z dimensions. This suggests the eye movements were more stable and less erratic in the intelligent avatar scenario compared to the non-intelligent avatar. As mentioned before, such stabilization of eye movements could be an important factor in mitigating cybersickness, as disjointed or erratic eye movements are often associated with feelings of disorientation and nausea in virtual environments.

Similarly, the persistent values of the intelligent avatar condition for several eye positions and rotation components were significantly higher. This implies that eye positioning and orientation were more consistent, potentially contributing to a more coherent and comfortable visual experience for users.

The head movement signals, including head position and rotation, exhibited the same pattern, with the intelligent avatar condition showing higher persistent values. This indicates that the head movements were more stable and less variable. Turning to the ACC signals, the persistent values

along the x and y axes were significantly higher in the intelligent avatar condition. This suggests more consistent patterns of acceleration, which could be associated with smoother and more natural movements of the virtual avatar.

Lastly, the physiological signals, such as EDA, BVP, and TMP, also demonstrated significantly higher persistent values in the intelligent avatar condition. This points to more stable physiological responses, which may indicate a more comfortable and less stressful user experience.

The critical observation here is that the increase in persistent values is most pronounced for the second homology group (H_2), followed by a minor increase in H_1 and, finally, the slightest increase in H_0. This specific pattern of decreasing persistent values across the homology groups ($H_2 > H_1 > H_1$) has an important implication:

The fact that the increase is most significant for H_2, followed by H_1 and then H_0 suggests a hierarchical reduction in the complexity of the signal dynamics. H_2 captures the higher-order, more complex relationships and inter-dependencies in the data. The more significant increase in H_2 persistence indicates that the intelligent avatar is most effective at stabilizing and simplifying these complex, higher-dimensional patterns. The minor increase in H_1 and H_0 persistent values suggests that the intelligent avatar can also reduce the complexity at lower dimensional levels but to a lesser degree.

5.12 Discussion

The results of our study underscore the significant benefits of employing intelligent avatars in VR environments, particularly regarding user safety. These findings align with existing literature and highlight the crucial role of physiological feedback in creating engaging and adaptive user experiences. For instance, Nacke et al. [156] emphasize the necessity of integrating real-time physiological responses for improved interaction in VR. Our study supports this notion, demonstrating that users experience lower levels of discomfort when interacting with intelligent avatars.

Additionally, we emphasize that the analysis of time-series data should not be limited to raw time-dependent metrics, as this may restrict the depth of insights. Future research should prioritize more comprehensive analytical techniques to fully capture the dynamics of user interactions.

Moreover, the observed decrease in eye velocity, eye angular velocity, and electrodermal activity in the presence of intelligent avatars suggests that these physiological metrics can serve as valuable

indicators of cybersickness. This finding provides a foundation for future research into real-time monitoring systems aimed at enhancing user well-being by reducing cybersickness.

Our findings also highlight the potential of user-centric design in avatar development, as opposed to generic design, to significantly improve user satisfaction. By considering users' behavioral and physiological states, we have established a framework for creating more personalized and adaptive virtual companions. The data indicate that users feel more engaged and less anxious when interacting with avatars that respond to their individual needs, underscoring the potential for intelligent avatars to enhance overall user satisfaction in VR environments.

Looking ahead, several avenues for future research emerge. First, exploring the application of intelligent avatars across various VR contexts—such as educational or therapeutic settings—could provide insights into their versatility and effectiveness in diverse scenarios. Additionally, conducting longitudinal studies would allow researchers to assess the long-term effects of intelligent avatars on user experience and safety. Understanding how these avatars impact user engagement over time could inform the design of more effective VR applications.

Since SmartSimVR is a flexible modular framework, each module can be enhanced and expanded as needed. For instance, we can improve the measurement processes or the AI module to enhance functionality.

Expanding the range of physiological metrics analyzed in future studies could lead to a more comprehensive understanding of user states in VR. Incorporating emotional states measured through EEG, for example, could provide deeper insights into the complex interplay between user emotions and immersive experiences.

Further exploration of advanced machine learning techniques—such as reinforcement learning, deep learning, and integration of computer vision—could significantly enhance the adaptability of intelligent avatars. This would facilitate more sophisticated behavior modeling based on user interactions and preferences, leading to richer and more personalized user experiences.

Investigating diverse user populations will also be beneficial for future work. Understanding trustworthiness across varied demographics can help identify differing expectations and perceptions of avatars, which is crucial for designing universally effective solutions.

While this study offers valuable insights into the trustworthiness of intelligent avatars in VR,

several limitations should be acknowledged:

Firstly, the findings are based on a limited sample size, which may not adequately represent the broader population. Additionally, there was a lack of diversity regarding user age and experience with virtual games. Future research should aim for a larger and more varied sample to enhance the generalizability of the results.

Furthermore, the study was conducted in a controlled environment and focused on a single scenario, which may restrict the external validity of the findings. Real-world applications of intelligent avatars could present additional challenges that warrant further investigation.

5.13 Summary

This chapter presents a detailed overview of the experimental procedure employed in the study. The study began with a pre-exposure questionnaire, where participants completed a SSQ to establish their baseline health status and motion sickness tendencies. The necessary equipment, including the head-mounted display, was then set up and calibrated to ensure a comfortable and immersive virtual experience for the participants.

During the simulation, participants were instructed to navigate the simulated city streets, and their responses were closely monitored. After the virtual driving task, participants completed a post-exposure SSQ to assess any changes in their cybersickness symptoms.

The experimental protocol emphasized two key indicators: self-reported sickness levels and exposure time. Participants verbally reported their sickness level on a scale of 0 to 3 every minute during the simulation. The experiment duration was flexible, ranging from 10 to 12 minutes, and the system dynamically adjusted the protocol based on participant feedback. This adaptive approach was implemented to ensure the participant's safety and collect accurate data.

Following the experiment design, a section delved into assessing the distribution of the model's performance during training. The accuracy and F1 score metrics distribution across all participants demonstrated strong performance in predicting cybersickness.

We then meticulously examined and compared both objective and subjective data from sessions involving intelligent and non-intelligent avatars.

Time-domain analysis, a cornerstone of our methodology, focuses on key metrics such as mean, standard deviation, and root mean square (RMS). Complementing this conventional approach, frequency-domain analysis employed techniques like Power Spectral Density (PSD), spectral entropy, and mean frequency. Additionally, time-frequency domain analysis leveraged the capabilities of spectrograms and wavelet energy to investigate the evolution of frequency characteristics. This multifaceted approach yielded deeper insights into the behavior of the data.

We underscore the superiority of wavelet transforms over traditional methods, showcasing their efficacy in capturing transient features that are crucial for a nuanced understanding of the data. Throughout our analyses, we noted a significant decrease in eye velocity (both x and y axes), eye angular velocity (x and y), and electrodermal activity (EDA) in the intelligent avatar condition compared to the non-intelligent avatar. These findings suggest that these physiological signals may serve as sensitive indicators of the cybersickness experience.

The subjective data analysis further enriched our interpretation of the results, providing a complementary perspective to the quantitative measures. We explored the convergence of objective and subjective measures of intelligent avatar performance, revealing a consistency in findings that bolsters the overall conclusions drawn from our analysis.

In addition, topological analysis was conducted to investigate the structure of the data, uncovering underlying patterns that may influence the results and adding another layer of depth to our findings.

Ultimately, our hypothesis regarding the positive effects of adaptive intelligent avatars to improve user safety with cybersickness reduction orientation which was defined as safety indicator in our use case was substantiated.

Chapter 6

Conclusion and Perspective

6.1 General Conclusion

This research aimed to improve the advancement of avatar technology in VR, making it more closely aligned with the user's state. The central contribution of this thesis was introducing the concept of a trustworthy intelligent avatar for virtual immersion. In the context of VR avatars, "trustworthiness" refers to the perceived reliability, credibility, and confidence that users experience when interacting with the virtual environment. Trustworthiness in avatars arises from a range of factors, one of the most crucial being adaptability, which enables them to adjust their behavior, responses, and environmental interactions dynamically. By applying AI techniques, avatars can be given this crucial adaptable nature. The result is a trustworthy, intelligent avatar - a customized, self-adapting avatar equipped with robust human-centered AI capabilities.

To realize the avatar-based concept, an architectural framework called SmartSimVR was first designed to meet the requirements. This innovative architecture encompassed four key features: distributability, concurrency, shared virtual memory, and stream learning.

The distributability of SmartSimVR allows the system to be deployed in a decentralized manner, enabling greater scalability and resilience. This is a scalable architecture that supports any number of sensors needed for fusion and multimodal data processing, further enhancing the system's adaptability and responsiveness. The concurrent nature of the architecture not only facilitates the simultaneous processing of multiple user inputs and environmental data streams, but also ensures real-time responsiveness, making the system highly efficient. The shared virtual memory component enables seamless information exchange between the various subsystems, promoting coordination and coherence. Importantly, the stream learning capabilities of SmartSimVR empower the avatar to continuously adapt and refine its behavior based on the user's evolving preferences and the dynamic simulation environment. This allows for the implementation of highly personalized and closed-loop systems that can dynamically adjust to the user's needs and the changing conditions of the simulated world. As a result, the virtual application enables a more tailored and responsive experience for the user.

The flexible and extensible nature of the SmartSimVR architecture means it can be applied to a wide range of auto-adaptive smart avatar approaches. This versatility allows researchers and developers to study and implement a variety of immersive virtual experiences. For example, the SmartSimVR framework could be leveraged to create avatars that dynamically adjust their visual fidelity, anima-

tion, and interactive capabilities based on the user's hardware capabilities and network conditions. Similarly, the architecture could support the development of avatars that learn to anticipate and accommodate the user's unique behavioral patterns and cognitive biases, further enhancing the sense of embodiment and natural interaction.

To validate the proposed concept, a driving simulator was chosen as a use case. In this simulator, the user assumes the role of the driver and has full control of the car (non-humanoid avatar). In other words, the car was considered an avatar in a functional sense that effectively fulfills this role by representing the user, facilitating interaction, and enhancing immersion. This driving simulator was implemented using SmartSimVR.

This model was designed to consider the user's current state and enhance their safety during the driving and navigation tasks. A key focus was reducing the incidence of cybersickness, a common issue in virtual reality experiences. The implemented driving simulator features an Adaptive Virtual Reality System (AVRS) that powers an auto-adapted intelligent avatar. This AVRS is designed around three core components:

- **Performance Measures:** The system collects and analyzes physiological, behavioral, and user feedback data to assess the user's state and experience within the simulation.
- **Adaptive Logic:** An advanced AI-based classification algorithm processes the performance data in real-time, providing insights that inform the system's adaptation.
- **Adaptive Variables:** Based on the adaptive logic, the system dynamically adjusts the car's acceleration and handling to mitigate cybersickness.

By seamlessly integrating these three components, the system enables the intelligent avatar to continuously evolve and adapt to each user's individual state, delivering a highly customized virtual reality experience and creating a more comfortable and sustainable driving experience within the simulated environment by proactively identifying and mitigating cybersickness.

The analysis of subjective and objective data has provided compelling evidence for the effectiveness of the intelligent avatar approach in our use case. Our in-depth analysis, including time, frequency, and time-frequency domain analysis, eye velocity, eye angular velocity, and EDA, revealed a significant reduction in the intelligent avatar condition. These physiological data, which we considered as the

cybersickness signature, provided a strong validation of our approach. The correlation between SSQ and the wavelet energy of these signature objective data further confirmed our hypothesis.

Additionally, the analysis of objective data in terms of topology and data connectivity pointed to more stable physiological responses using intelligent avatars, which may indicate a more comfortable and less stressful user experience.

Furthermore, the performance analysis has uncovered valuable insights into the scalability and optimization opportunities for the VR application. By addressing the identified areas of improvement, particularly the ASR component, we can significantly enhance the system's ability to deliver timely and accurate results. This enhancement will lead to a more robust and capable solution, further improving the user experience.

6.2 Research Perspective

Further enhancements of the capabilities of auto-adaptive avatars could create an even more immersive and realistic virtual experience across various applications. A significant area for future development involves incorporating advanced physiological monitoring techniques to assess users' states more accurately, moving beyond reliance on self-reported feedback. Current systems depend heavily on participant self-reporting as a critical input for the adaptive logic that informs avatar adjustments to the simulation environment. However, this method has inherent limitations, as users may struggle to accurately perceive or articulate symptoms such as nausea, disorientation, or fatigue. Such discrepancies between perceived and actual physiological states can compromise the effectiveness of the adaptive system.

To address these challenges, future development could explore the use of EEG signals as direct classification labels rather than solely relying on user feedback. EEG has been extensively studied for detecting neurophysiological indicators associated with cybersickness, including changes in alpha and theta band power, increased frontal lobe activity, and altered brain connectivity patterns. Shifting from self-reported to physiologically grounded labels can significantly enhance the reliability and responsiveness of the adaptive system, allowing avatars to make more informed adjustments in real-time.

Moreover, the concept of auto-adaptive avatars can significantly enhance user experiences by dynamically adjusting their behaviors based on users' skill levels, emotional states, and attention spans.

For instance, an auto-adaptive avatar might modify its level of assistance, providing more guidance and support to novice users while offering greater autonomy to experienced users. This adaptability ensures that users feel both challenged and supported, fostering a more engaging and immersive environment.

Integrating semi-autonomous capabilities into avatars can further enhance user safety and comfort. In scenarios where users may become overwhelmed or distracted, the avatar could take proactive measures to assist. For example, it might provide timely alerts or suggestions to help users navigate complex situations, ensuring a smoother experience. Additionally, the avatar could automatically manage certain tasks, such as adjusting the environment or offering contextual information, without undermining the user's sense of agency.

By balancing user autonomy with intelligent system interventions, auto-adaptive avatars can serve as effective companions across diverse virtual contexts, from education to therapy. This approach not only mitigates risks associated with user errors but also enhances overall engagement, making virtual experiences more responsive to individual needs and preferences. Furthermore, the integration of advanced AI-powered decision-making algorithms could empower avatars to anticipate potential scenarios, evaluate risks, and make proactive choices to optimize the user's experience, ultimately leading to a safer and more personalized virtual environment.

6.3 Scientific Publications

All the research results have been presented in a series of scientific publications, including peer-reviewed journal article and presentations at international conferences, and will be further extended in future work.

1. Journal

    - Hadadi, A., Guillet, C., Chardonnet, J. R., Langovoy, M., Wang, Y., & Ovtcharova, J. (2022). Prediction of cybersickness in virtual environments using topological data analysis and machine learning. Frontiers in Virtual Reality, 3, 973236.

2. Conferences

    - Hadadi, A.; Chardonnet, J.-R.; Guillet, C. & Ovtcharova, J., Machine Learning Application

for Real-Time Simulator, Proceedings of the 2024 9th International Conference on Machine Learning Technologies, Association for Computing Machinery, 2024, 1–5

- Hadadi, A., Chardonnet, J. R., Guillet, C., & Ovtcharova, J. (2024, February). Smart-SimVR: An Architecture Integrating Machine Learning and Virtual Environment for Real-Time Simulation Adaptation. In 2024 10th International Conference on Automation, Robotics and Applications (ICARA) (pp. 531-535). IEEE.
- Hadadi, A., Chardonnet, J. R., Guillet, C., & Ovtcharova, J. (2024, January). Intelligent Virtual Platform for Real-time Cybersickness Detection and Adaptation. In 2024 IEEE International Conference on Artificial Intelligence and eXtended and Virtual Reality (AIxVR) (pp. 231-235). IEEE.
- Hadadi, Azadeh, Jean-Rémy Chardonnet, Mikhail Langovoy, Christophe Guillet, Jivka Ovtacharova. "Intelligent VR in Driving Simulator." Poster presented at the DSC 2023.

Bibliography

[1] R. Abbasi-Asl, M. Keshavarzi, and D. Y. Chan. Brain-computer interface in virtual reality. In *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 1220–1224. IEEE, 2019.

[2] S. Akbas, A. Evren Yantac, T. Eskenazi, K. Kuscu, S. Semsioglu, O. Topal Sumer, and A. Ozturk. Virtual dance mirror: A functional approach to avatar representation through movement in immersive vr. In *Proceedings of the 8th International Conference on Movement and Computing*, pages 1–4, 2022.

[3] M. N. Aladwan, F. M. Awaysheh, S. Alawadi, M. Alazab, T. F. Pena, and J. C. Cabaleiro. Truste-vc: Trustworthy evaluation framework for industrial connected vehicles in the cloud. *IEEE transactions on industrial informatics*, 16(9):6203–6213, 2020.

[4] T. Alam, S. Qamar, A. Dixit, and M. Benaida. Genetic algorithm: Reviews, implementations, and applications. *arXiv preprint arXiv:2007.12673*, 2020.

[5] B. Albert, G. Ricard, H. Geoff, and P. Bernhard. Machine learning for data streams with practical examples in moa, 2018.

[6] A. Almeida, F. Rebelo, P. Noriega, and E. Vilar. Virtual reality self induced cybersickness: an exploratory study. In *Advances in Ergonomics in Design: Proceedings of the AHFE 2017 International Conference on Ergonomics in Design, July 17- 21, 2017, The Westin Bonaventure Hotel, Los Angeles, California, USA 8*, pages 26–33. Springer, 2018.

[7] A. Alshehri and A. Algarni. Tf-tda: A novel supervised term weighting scheme for sentiment analysis. *Electronics*, 12(7):1632, 2023.

[8] E. J. Amézquita, M. Y. Quigley, T. Ophelders, E. Munch, and D. H. Chitwood. The shape of things to come: Topological data analysis and biology, from molecules to organisms. *Developmental Dynamics*, 249(7):816–833, 2020.

[9] S. R. Anderson, M. Gianola, N. A. Medina, J. M. Perry, T. D. Wager, and E. A. R. Losin. Doctor trustworthiness influences pain and its neural correlates in virtual medical interactions. *Cerebral Cortex*, 33(7):3421–3436, 2023.

[10] B. Arcioni, S. Palmisano, D. Apthorp, and J. Kim. Postural stability predicts the likelihood of cybersickness in active hmd-based virtual reality. *Displays*, 58:3–11, 2019.

[11] F. Argelaguet and C. Andujar. A survey of 3d object selection techniques for virtual environments. *Computers & Graphics*, 37(3):121–136, 2013.

[12] F. Argelaguet, L. Hoyet, M. Trico, and A. Lécuyer. The role of interaction in virtual embodiment: Effects of the virtual hand representation. In *2016 IEEE virtual reality (VR)*, pages 3–10. IEEE, 2016.

[13] B. Arnaldi, P. Fuchs, and J. Tisseau. Chapitre 1 du volume 1 du traité de la réalité virtuelle. *Les Presses de l'Ecole des Mines de Paris*, 1:131, 2003.

[14] L. L. Arns and M. M. Cerney. The relationship between age and incidence of cybersickness among immersive environment users. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.*, pages 267–268. IEEE, 2005.

[15] A. Asaad and S. Jassim. Topological data analysis for image tampering detection. In *Digital Forensics and Watermarking: 16th International Workshop, IWDW 2017, Magdeburg, Germany, August 23-25, 2017, Proceedings 16*, pages 136–146. Springer, 2017.

[16] C. Axenie, A. Becher, D. Kurz, and T. Grauschopf. Meta-learning for avatar kinematics reconstruction in virtual reality rehabilitation. In *2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE)*, pages 617–624. IEEE, 2019.

[17] J. N. Bailenson, N. Yee, J. Blascovich, A. C. Beall, N. Lundblad, and M. Jin. The use of immersive virtual reality in the learning sciences: Digital transformations of teachers, students, and social context. *The journal of the learning sciences*, 17(1):102–141, 2008.

[18] C. Baker and S. H. Fairclough. Adaptive virtual reality. In *Current research in neuroadaptive technology*, pages 159–176. Elsevier, 2022.

[19] O. Bălan, A. Moldoveanu, and M. Leordeanu. A machine learning approach to automatic phobia therapy with virtual reality. *Modern Approaches to Augmentation of Brain Function*, pages 607–636, 2021.

[20] S. Benford, C. Greenhalgh, T. Rodden, and J. Pycock. Collaborative virtual environments. *Communications of the ACM*, 44(7):79–85, 2001.

[21] B. Bergström. Morphology of the vestibular nerve: Ii. the number of myelinated vestibular nerve fibers in man at various ages. *Acta oto-laryngologica*, 76(1-6):173–179, 1973.

[22] A. Bernstein, E. Burnaev, M. Sharaev, E. Kondrateva, and O. Kachan. Topological data analysis in computer vision. In *Twelfth International Conference on Machine Vision (ICMV 2019)*, volume 11433, pages 673–679. SPIE, 2020.

[23] A. Bifet and R. Gavalda. Learning from time-changing data with adaptive windowing. In *Proceedings of the 2007 SIAM international conference on data mining*, pages 443–448. SIAM, 2007.

[24] P. Bimberg, T. Weissker, and A. Kulik. On the usage of the simulator sickness questionnaire for virtual reality research. In *2020 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW)*, pages 464–467. IEEE, 2020.

[25] E. B. Blackford and J. R. Estepp. Measurements of pulse rate using long-range imaging photoplethysmography and sunlight illumination outdoors. In *Optical Diagnostics and Sensing XVII: Toward Point-of-Care Diagnostics*, volume 10072, pages 122–134. SPIE, 2017.

[26] E. B. Blackford, J. R. Estepp, and D. J. McDuff. Remote spectral measurements of the blood volume pulse with applications for imaging photoplethysmography. In *Optical diagnostics and sensing XVIII: toward point-of-care diagnostics*, volume 10501, pages 192–199. SPIE, 2018.

[27] O. Blanke and T. Metzinger. Full-body illusions and minimal phenomenal selfhood. *Trends in cognitive sciences*, 13(1):7–13, 2009.

[28] C. Böffel and J. Müsseler. Action effect consistency and body ownership in the avatar-simon task. *PLoS One*, 14(8):e0220817, 2019.

[29] B. Bowins. Motion sickness: A negative reinforcement model. *Brain research bulletin*, 81(1): 7–11, 2010.

[30] D. A. Bowman. Interaction techniques for common tasks in immersive virtual environments. *Georgia Institute of Technology*, 1999.

[31] D. A. Bowman and R. P. McMahan. Virtual reality: how much immersion is enough? *Computer*, 40(7):36–43, 2007.

[32] J. O. Brooks, R. R. Goodenough, M. C. Crisler, N. D. Klein, R. L. Alley, B. L. Koon, W. C. Logan Jr, J. H. Ogle, R. A. Tyrrell, and R. F. Wills. Simulator sickness during driving simulation studies. *Accident analysis & prevention*, 42(3):788–796, 2010.

[33] L. Bryant, B. Hemsley, B. Bailey, A. Bluff, V. Nguyen, P. Stubbs, D. Barnett, C. Jacobs, C. Lucas, and E. Power. Opportunities for immersive virtual reality in rehabilitation: focus on communication disability. In *Proceedings of the 53rd Hawaii International Conference of System Sciences 2020*, pages 3567–3576, 2020.

[34] J. Campbell and M. Fraser. Switching it up: Designing adaptive interfaces for virtual reality exergames. In *Proceedings of the 31st European conference on cognitive ergonomics*, pages 177–184, 2019.

[35] J. Campbell and M. Fraser. Switching it up: Designing adaptive interfaces for virtual reality exergames. In *Proceedings of the 31st European Conference on Cognitive Ergonomics*, pages 177–184, 2019.

[36] G. Carlsson. Topology and data. *Bulletin of the American Mathematical Society*, 46(2):255–308, 2009.

[37] M. Chancel and H. H. Ehrsson. Which hand is mine? discriminating body ownership perception in a two-alternative forced-choice task. *Attention, Perception, & Psychophysics*, 82(8):4058–4083, 2020.

[38] J.-R. Chardonnet, M. A. Mirzaei, and F. Mérienne. Features of the postural sway signal as indicators to estimate and predict visually induced motion sickness in virtual reality. *International Journal of Human–Computer Interaction*, 33(10):771–785, 2017.

[39] F. Chazal. High-dimensional topological data analysis. In *Handbook of Discrete and Computational Geometry*, pages 663–683. Chapman and Hall/CRC, 2017.

[40] F. Chazal, V. De Silva, M. Glisse, and S. Oudot. *The structure and stability of persistence modules*, volume 10. Springer, 2016.

[41] L. Chen, C. Cao, F. De la Torre, J. Saragih, C. Xu, and Y. Sheikh. High-fidelity face tracking for ar/vr via deep lighting adaptation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 13059–13069, 2021.

[42] R. Chengoden, N. Victor, T. Huynh-The, G. Yenduri, R. H. Jhaveri, M. Alazab, S. Bhattacharya, P. Hegde, P. K. R. Maddikunta, and T. R. Gadekallu. Metaverse for healthcare: a survey on potential applications, challenges and future directions. *IEEE Access*, 11:12765–12795, 2023.

[43] S. Choe and S. Ramanna. Cubical homology-based machine learning: An application in image classification. *Axioms*, 11(3):112, 2022.

[44] J. Y. Chun, H.-J. Kim, J.-W. Hur, D. Jung, H.-J. Lee, S. P. Pack, S. Lee, G. Kim, C.-Y. Cho, S.-M. Lee, et al. Prediction of specific anxiety symptoms and virtual reality sickness using in situ autonomic physiological signals during virtual reality treatment in patients with social anxiety disorder: Mixed methods study. *JMIR Serious Games*, 10(3):e38284, 2022.

[45] W. Chung and M. Barnett-Cowan. Sensory reweighting: a common mechanism for subjective visual vertical and cybersickness susceptibility. *Virtual Reality*, 27(3):2029–2041, 2023.

[46] T. Collingwoode-Williams, Z. O'Shea, M. Gillies, and X. Pan. The impact of self-representation and consistency in collaborative virtual environments. *Frontiers in Virtual Reality*, 2:648601, 2021.

[47] C. Cruz-Neira, D. J. Sandin, T. A. DeFanti, R. V. Kenyon, and J. C. Hart. The cave: Audio visual experience automatic virtual environment. *Communications of the ACM*, 35(6):64–73, 1992.

[48] C. Cruz-Neira, D. J. Sandin, and T. A. DeFanti. Surround-screen projection-based virtual reality: the design and implementation of the cave. In *Seminal Graphics Papers: Pushing the Boundaries, Volume 2*, pages 51–58. 2023.

[49] A. Dave, J. C. Vaz, J. Kim, N. Kosanovic, N. Kassai, and P. Y. Oh. Avatar-darwin a social humanoid with telepresence abilities aimed at embodied avatar systems. In *2022 IEEE-RAS 21st International Conference on Humanoid Robots (Humanoids)*, pages 47–52. IEEE, 2022.

[50] T. Davies. Topological data analysis for anomaly detection in host-based logs. *arXiv preprint arXiv:2204.12919*, 2022.

[51] M. S. de Aquino and F. d. F. de Souza. Adaptive virtual environments: the role of intelligent agents. In *Practical Applications of Agent-Based Technology*, pages 87–110. INTECH Open Science, 2012.

[52] G. De Haan and A. Van Leest. Improved motion robustness of remote-ppg by using the blood volume pulse signature. *Physiological measurement*, 35(9):1913, 2014.

[53] E. J. De Visser, S. S. Monfort, R. McKendrick, M. A. Smith, P. E. McKnight, F. Krueger, and R. Parasuraman. Almost human: Anthropomorphism increases trust resilience in cognitive agents. *Journal of Experimental Psychology: Applied*, 22(3):331, 2016.

[54] H. G. Debarba, E. Molla, B. Herbelin, and R. Boulic. Characterizing embodied interaction in first and third person perspective viewpoints. In *2015 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 67–72. IEEE, 2015.

[55] H. G. Debarba, S. Chagué, and C. Charbonnier. On the plausibility of virtual body animation features in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 28(4): 1880–1893, 2020.

[56] L. Deng and D. O'Shaughnessy. *Speech processing: a dynamic and optimization-oriented approach*. CRC Press, 2003.

[57] N. Devane, N. Behn, J. Marshall, A. Ramachandran, S. Wilson, and K. Hilari. The use of virtual reality in the rehabilitation of aphasia: a systematic review. *Disability and Rehabilitation*, 45 (23):3803–3822, 2023.

[58] D. Dewez, L. Hoyet, A. Lécuyer, and F. Argelaguet. Do you need another hand? investigating dual body representations during anisomorphic 3d manipulation. *IEEE Transactions on Visualization and Computer Graphics*, 28(5):2047–2057, 2022.

[59] Edelsbrunner, Letscher, and Zomorodian. Topological persistence and simplification. *Discrete & computational geometry*, 28:511–533, 2002.

[60] H. Edelsbrunner. Persistent homology: theory and practice. 2013.

[61] H. H. Ehrsson, K. Wiech, N. Weiskopf, R. J. Dolan, and R. E. Passingham. Threatening a rubber hand that you feel is yours elicits a cortical anxiety response. *Proceedings of the National Academy of Sciences*, 104(23):9828–9833, 2007.

[62] N. Elyasi and M. Hosseini Moghadam. Classification of skin lesions by tda alongside xception neural network. *Journal of AI and Data Mining*, 10(3):333–344, 2022.

[63] N. Elyasi and M. H. Moghadam. An introduction to a new text classification and visualization for natural language processing using topological data analysis. *arXiv preprint arXiv:1906.01726*, 2019.

[64] B. Englebert, L. Marsman, and J. Crijnen. Evaluating the effectiveness of mixed reality as a cybersickness mitigation strategy in helicopter flight simulation. *Human Factors and Simulation*, 2023. URL https://api.semanticscholar.org/CorpusID:259731925.

[65] B. T. Fasy, J. Kim, F. Lecci, and C. Maria. Introduction to the r package tda. *arXiv preprint arXiv:1411.1830*, 2014.

[66] D. J. Ferguson, B. Hailemeskel, and F. Fullas. Anti-motion sickness properties of ginger capsules, peppermint oil inhalation, acupressure wristbands and vitamin b6: A survey of pharmacy students. *International Journal of Scholarly Research in Biology and Pharmacy*, 2024. URL https://api.semanticscholar.org/CorpusID:270541038.

[67] A. Ferrà, G. Cecchini, F.-P. Nobbe Fisas, C. Casacuberta, and I. Cos. A topological classifier to characterize brain states: When shape matters more than variance. *Plos one*, 18(10):e0292049, 2023.

[68] L. Frank, R. S. Kennedy, R. S. Kellogg, M. E. McCauley, and E. C. O. FL. Simulator sickness: A reaction to a transformed perceptual world. 1. scope of the problem. In *Proceedings of the Second Symposium of Aviation Psychology, Ohio State University, Columbus OH*, pages 25–28, 1983.

[69] J. P. Freiwald, J. Schenke, N. Lehmann-Willenbrock, and F. Steinicke. Effects of avatar appearance and locomotion on co-presence in virtual reality collaborations. In *Proceedings of Mensch Und Computer 2021*, pages 393–401. 2021.

[70] R. Fribourg, N. Ogawa, L. Hoyet, F. Argelaguet, T. Narumi, M. Hirose, and A. Lécuyer. Virtual co-embodiment: evaluation of the sense of agency while sharing the control of a virtual body among two individuals. *IEEE Transactions on Visualization and Computer Graphics*, 27(10): 4023–4038, 2020.

[71] S. Füllbrunn, K. Richwien, and A. Sadrieh. Trust and trustworthiness in anonymous virtual worlds. *Journal of Media Economics*, 24(1):48–63, 2011.

[72] I. Gabriel. Artificial intelligence, values, and alignment. *Minds and machines*, 30(3):411–437, 2020.

[73] 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*, 8(6):439–444, 2019.

[74] A. Garin and G. Tauzin. A topological" reading" lesson: Classification of mnist using tda. In *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*, pages 1551–1556. IEEE, 2019.

[75] 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*, 26(4):1347–1371, 2022.

[76] C. Geniesse, O. Sporns, G. Petri, and M. Saggar. Generating dynamical neuroimaging spatiotemporal representations (dyneusr) using topological data analysis. *Network neuroscience*, 3 (3):763–778, 2019.

[77] A. P. Gilakjani et al. Visual, auditory, kinaesthetic learning styles and their impacts on english language teaching. *Journal of studies in education*, 2(1):104–113, 2012.

[78] M. Gillies, X. Pan, and M. Slater. Piavca: a framework for heterogeneous interactions with virtual characters. *Virtual reality*, 14:221–228, 2010.

[79] M. Gongora and D. Irvine. Adaptive intelligent agents based on efficient behaviour differentiation models. In *2010 IEEE ANDESCON*, pages 1–6. IEEE, 2010.

[80] G. Gorisse, O. Christmann, E. A. Amato, and S. Richir. First-and third-person perspectives in immersive virtual environments: presence and performance analysis of embodied users. *Frontiers in Robotics and AI*, 4:33, 2017.

[81] Z. Gracia-Tabuenca, J. C. Díaz-Patiño, I. Arelio-Ríos, M. B. Moreno-García, F. A. Barrios, and S. Alcauter. Development of the functional connectome topology in adolescence: Evidence from topological data analysis. *Eneuro*, 10(2), 2023.

[82] C. Grosan, A. Abraham, C. Grosan, and A. Abraham. Rule-based expert systems. *Intelligent systems: A modern approach*, pages 149–185, 2011.

[83] L. F. Guerrero-Vásquez, D. X. Landy-Rivera, J. F. Bravo-Torres, M. López-Nores, R. Castro-Serrano, and P. E. Vintimilla-Tapia. Avatar: Contribution to human-computer interaction processes through the adaptation of semi-personalized virtual agents. In *2018 IEEE biennial congress of Argentina (Argencon)*, pages 1–4. IEEE, 2018.

[84] K. Gupta, R. Hajika, Y. S. Pai, A. Duenser, M. Lochner, and M. Billinghurst. In ai we trust: Investigating the relationship between biosignals, trust and cognitive load in vr. In *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology*, pages 1–10, 2019.

[85] M. Guy, C. Jeunet-Kelway, G. Moreau, and J.-M. Normand. Manipulating the sense of embodiment in virtual reality: a study of the interactions between the senses of agency, self-location and ownership. In *ICAT-EGVE2022, the joint international conference of the 32nd International Conference on Artificial Reality and Telexistence & the 27th Eurographics Symposium on Virtual Environments (2022)*, pages 1–11, 2022.

[86] D.-I. D. Han, Y. Bergs, and N. Moorhouse. Virtual reality consumer experience escapes: preparing for the metaverse. *Virtual Reality*, 26(4):1443–1458, 2022.

[87] J. Hartfill, J. Gabel, L. Kruse, S. Schmidt, K. Riebandt, S. Kühn, and F. Steinicke. Analysis of detection thresholds for hand redirection during mid-air interactions in virtual reality. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology*, pages 1–10, 2021.

[88] S. Z. Hassan, P. Salehi, M. A. Riegler, M. S. Johnson, G. A. Baugerud, P. Halvorsen, and S. S. Sabet. A virtual reality talking avatar for investigative interviews of maltreat children. In *Proceedings of the 19th international conference on content-based multimedia indexing*, pages 201–204, 2022.

[89] A. Hatcher. *Algebraic topology*. 2005.

[90] J. A. Heathers, E. Fink, R.-L. Kuhnert, and M. de Rosnay. Blood volume pulse (bvp) derived vagal tone (vt) between 5 and 7 years of age: A methodological investigation of measurement and longitudinal stability. *Developmental Psychobiology*, 56(1):23–35, 2014.

[91] M. L. Heilig. Sensorama simulator. *US PAT. 3,050,870*, 1962.

[92] J. Hoey, P. Poupart, A. von Bertoldi, T. Craig, C. Boutilier, and A. Mihailidis. Automated handwashing assistance for persons with dementia using video and a partially observable markov decision process. *Computer Vision and Image Understanding*, 114(5):503–519, 2010.

[93] J. Huang and Y. Jung. Perceived authenticity of virtual characters makes the difference. *Frontiers in Virtual Reality*, 3:1033709, 2022.

[94] X. Huang, A. Acero, H.-W. Hon, and R. Reddy. *Spoken language processing: A guide to theory, algorithm, and system development*. Prentice hall PTR, 2001.

[95] 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*, pages 561–572, 2017.

[96] L. Jayaraj, J. Wood, and M. Gibson. Improving the immersion in virtual reality with real-time avatar and haptic feedback in a cricket simulation. In *2017 IEEE international symposium on mixed and augmented reality (ISMAR-adjunct)*, pages 310–314. IEEE, 2017.

[97] 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)*, pages 827–835. IEEE, 2019.

[98] C. Jeunet, L. Albert, F. Argelaguet, and A. Lécuyer. "do you feel in control?": towards novel approaches to characterise, manipulate and measure the sense of agency in virtual environments. *IEEE transactions on visualization and computer graphics*, 24(4):1486–1495, 2018.

[99] Y. Jiang, Z. Li, M. He, D. Lindlbauer, and Y. Yan. Handavatar: Embodying non-humanoid virtual avatars through hands. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pages 1–17, 2023.

[100] K. Jung, V. T. Nguyen, D. Piscarac, and S.-C. Yoo. Meet the virtual jeju dol harubang—the mixed vr/ar application for cultural immersion in korea's main heritage. *ISPRS International Journal of Geo-Information*, 9(6):367, 2020.

[101] S. Kandanaarachchi and R. J. Hyndman. Leave-one-out kernel density estimates for outlier detection. *Journal of Computational and Graphical Statistics*, 31(2):586–599, 2022.

[102] I. Kastanis and M. Slater. Reinforcement learning utilizes proxemics: An avatar learns to manipulate the position of people in immersive virtual reality. *ACM Transactions on Applied Perception (TAP)*, 9(1):1–15, 2012.

[103] 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*, 65(1):119–127, 2018.

[104] A. Kemeny, J.-R. Chardonnet, and F. Colombet. Getting rid of cybersickness. *Virtual Reality, Augmented Reality, and Simulators*, 2020.

[105] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator sickness question-

naire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220, 1993.

[106] B. P. Kerfoot, H. E. Baker, M. O. Koch, D. Connelly, D. B. Joseph, and M. L. Ritchey. Randomized, controlled trial of spaced education to urology residents in the united states and canada. *The Journal of urology*, 177(4):1481–1487, 2007.

[107] B. Keshavarz and H. Hecht. Validating an efficient method to quantify motion sickness. *Human factors*, 53(4):415–426, 2011.

[108] B. Keshavarz, L. Hettinger, D. Vena, and J. Campos. Combined effects of auditory and visual cues on the perception of vection. *Experimental brain research. Experimentelle Hirnforschung. Experimentation cerebrale*, 232, 12 2013. doi:10.1007/s00221-013-3793-9.

[109] K. Kilteni, R. Groten, and M. Slater. The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, 21(4):373–387, 2012.

[110] 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*, 69:66–73, 2018.

[111] 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*, pages 10580–10589, 2019.

[112] S.-Y. Kim, H. Park, M. Jung, and K. Kim. Impact of body size match to an avatar on the body ownership illusion and user's subjective experience. *Cyberpsychology, Behavior, and Social Networking*, 23(4):234–241, 2020.

[113] R. Kindelan, J. Frías, M. Cerda, and N. Hitschfeld. Classification based on topological data analysis. *arXiv preprint arXiv:2102.03709*, 2021.

[114] A. Klimek, I. Mannheim, G. Schouten, E. J. Wouters, and M. W. Peeters. Wearables measuring electrodermal activity to assess perceived stress in care: a scoping review. *Acta Neuropsychiatrica*, pages 1–11, 2023.

[115] P. Knierim, V. Schwind, A. M. Feit, F. Nieuwenhuizen, and N. Henze. Physical keyboards in virtual reality: Analysis of typing performance and effects of avatar hands. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, pages 1–9, 2018.

[116] G. Kong, K. He, and K. Wei. Sensorimotor experience in virtual reality enhances sense of agency associated with an avatar. *Consciousness and cognition*, 52:115–124, 2017.

[117] P. Kourtesis, R. Amir, J. Linnell, F. Argelaguet, and S. E. MacPherson. Cybersickness, cognition, & motor skills: The effects of music, gender, and gaming experience. *IEEE Transactions on Visualization and Computer Graphics*, 29(5):2326–2336, 2023.

[118] M. Kramár, A. Goullet, L. Kondic, and K. Mischaikow. Persistence of force networks in compressed granular media. *Physical Review E*, 87(4):042207, 2013.

[119] A. Krekhov, S. Cmentowski, K. Emmerich, and J. Krüger. Beyond human: Animals as an escape from stereotype avatars in virtual reality games. In *Proceedings of the annual symposium on computer-human interaction in play*, pages 439–451, 2019.

[120] S. Krishnakumar, V. Manoharan, N. Saritakumar, et al. Digital signal processing on 3-axis accelerometer. In *2023 International Conference on Intelligent Systems for Communication, IoT and Security (ICISCoIS)*, pages 317–321. IEEE, 2023.

[121] J. Kritikos, G. Alevizopoulos, and D. Koutsouris. Personalized virtual reality human-computer interaction for psychiatric and neurological illnesses: a dynamically adaptive virtual reality environment that changes according to real-time feedback from electrophysiological signal responses. *Frontiers in Human Neuroscience*, 15:596980, 2021.

[122] P. Kruszewski and T. J. Mahamad. The ai powered magic mirror: building immersive ar/vr experiences with only webcams and deep learning. In *ACM SIGGRAPH 2018 Virtual, Augmented, and Mixed Reality*, pages 1–1. 2018.

[123] 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)*, pages 91–92. IEEE, 2018.

[124] E. Lallart, X. Lallart, and R. Jouvent. Agency, the sense of presence, and schizophrenia. *Cyberpsychology & behavior*, 12(2):139–145, 2009.

[125] V. R. Lee. Youth engagement during making: using electrodermal activity data and first-person video to generate evidence-based conjectures. 2021. URL https://api.semanticscholar.org/CorpusID:236309560.

[126] 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*, 8:126784–126796, 2020.

[127] 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*, 24(10):1689–1700, 2013.

[128] J. Lin and N. Li. Towards a framework to model intelligent avatars in immersive virtual environments for studying human behavior in building fire emergencies. In J. Y. Chen and G. Fragomeni, editors, *Virtual, Augmented and Mixed Reality. Multimodal Interaction*, pages 349–360, Cham, 2019. Springer International Publishing. ISBN 978-3-030-21607-8.

[129] Y. Ling, W.-P. Brinkman, H. T. Nefs, C. Qu, and I. Heynderickx. Cybersickness and anxiety in virtual environments. In *2011 Joint Virtual Reality Conference. Nottingham UK: VTT Symposium*, pages 80–82, 2011.

[130] Z. Liu, T. D. Nguyen, A. Ene, and H. Nguyen. Adaptive accelerated (extra-) gradient methods with variance reduction. In *International Conference on Machine Learning*, pages 13947–13994. PMLR, 2022.

[131] C. K. Lo. What is the impact of chatgpt on education? a rapid review of the literature. *Education Sciences*, 13(4):410, 2023.

[132] M. Lombard and T. Ditton. At the heart of it all: The concept of presence. *Journal of computer-mediated communication*, 3(2):JCMC321, 1997.

[133] R. Lou, F. Mérienne, and D. Bechmann. General framework of geometric simplification for mitigating cybersickness. In *2022 International Conference on Future Trends in Smart Communities (ICFTSC)*, pages 1–5. IEEE, 2022.

[134] J.-L. Lugrin, J. Latt, and M. E. Latoschik. Anthropomorphism and illusion of virtual body ownership. *ICAT-EGVE*, 15:1–8, 2015.

[135] P. Y. Lum, G. Singh, A. Lehman, T. Ishkanov, M. Vejdemo-Johansson, M. Alagappan, J. Carlsson, and G. Carlsson. Extracting insights from the shape of complex data using topology. *Scientific reports*, 3(1):1236, 2013.

[136] M. Machneva, A. M. Evans, and O. Stavrova. Consensus and (lack of) accuracy in perceptions of avatar trustworthiness. *Computers in Human Behavior*, 126:107017, 2022.

[137] P. Majaranta and A. Bulling. Eye tracking and eye-based human–computer interaction. In *Advances in physiological computing*, pages 39–65. Springer, 2014.

[138] Z. A. A. Manjiyani, R. T. Jacob, R. Keerthan Kumar, and B. Varghese. Development of mems based 3-axis accelerometer for hand movement monitoring. *International Journal of Scientific and Research Publications*, 4(2):1, 2014.

[139] C. Maria, J.-D. Boissonnat, M. Glisse, and M. Yvinec. The gudhi library: Simplicial complexes and persistent homology. In *Mathematical Software–ICMS 2014: 4th International Congress, Seoul, South Korea, August 5-9, 2014. Proceedings 4*, pages 167–174. Springer, 2014.

[140] J. Marino, P. Lin, N. Karlova, and M. Eisenberg. Avatar transparency and the establishment of trust in virtual information eco-systems. *Proceedings of the American Society for Information Science and Technology*, 47(1):1–2, 2010.

[141] G. Markkula, R. Romano, R. Waldram, O. Giles, C. Mole, and R. Wilkie. Modelling visual-vestibular integration and behavioural adaptation in the driving simulator. *Transportation research part F: traffic psychology and behaviour*, 66:310–323, 2019.

[142] A. Maselli and M. Slater. The building blocks of the full body ownership illusion. *Frontiers in human neuroscience*, 7:83, 2013.

[143] T. Matahari. Webxr asset management in developing virtual reality learning media. *Indonesian Journal of Computing, Engineering, and Design (IJoCED)*, 4(1):38–46, 2022.

[144] D. McCarthy et al. Distributed vr rendering using nvidia optix. *Electronic Imaging*, 29:36–41, 2017.

[145] J. McIntosh, H. D. Zajac, A. N. Stefan, J. Bergström, and K. Hornbæk. Iteratively adapting avatars using task-integrated optimisation. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, pages 709–721, 2020.

[146] F. Miao, I. V. Kozlenkova, H. Wang, T. Xie, and R. W. Palmatier. An emerging theory of avatar marketing. *Journal of Marketing*, 86(1):67–90, 2022.

[147] J. N. Mikeska and H. Howell. Authenticity perceptions in virtual environments. *Information and Learning Sciences*, 122(7/8):480–502, 2021.

[148] P. Milgram and F. Kishino. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, 77(12):1321–1329, 1994.

[149] S. Mitra and K. R. JV. Experiments on fraud detection use case with qml and tda mapper. In *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)*, pages 471–472. IEEE, 2021.

[150] N. Mohajer, H. Abdi, K. Nelson, and S. Nahavandi. Vehicle motion simulators, a key step towards road vehicle dynamics improvement. *Vehicle System Dynamics*, 53(8):1204–1226, 2015.

[151] G. F. Monkam, M. J. De Lucia, and N. D. Bastian. Preprocessing network traffic using topological data analysis for data poisoning detection. In *2023 IEEE Conference on Dependable and Secure Computing (DSC)*, pages 1–8. IEEE, 2023.

[152] J. Montiel, M. Halford, S. M. Mastelini, G. Bolmier, R. Sourty, R. Vaysse, A. Zouitine, H. M. Gomes, J. Read, T. Abdessalem, et al. River: machine learning for streaming data in python. *Journal of Machine Learning Research*, 22(110):1–8, 2021.

[153] R. Moradinezhad and E. T. Solovey. Investigating trust in interaction with inconsistent embodied virtual agents. *International Journal of Social Robotics*, 13(8):2103–2118, 2021.

[154] S. Morélot, A. Garrigou, J. Dedieu, and B. N'Kaoua. Virtual reality for fire safety training: Influence of immersion and sense of presence on conceptual and procedural acquisition. *Computers & Education*, 166:104145, 2021.

[155] B. J. Mount, J. Scott, R. L. Hastings, D. Bennett, S. G. Latta, D. J. McCulloch, K. A. Geisner, J. T. Steed, and M. J. Scavezze. Augmented reality playspaces with adaptive game rules, Sept. 27 2016. US Patent 9,454,849.

[156] 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*, pages 103–112, 2011.

[157] T. Nakamura, Y. Hiraoka, A. Hirata, E. G. Escolar, and Y. Nishiura. Persistent homology and many-body atomic structure for medium-range order in the glass. *Nanotechnology*, 26(30): 304001, 2015.

[158] S. Narang, A. Best, and D. Manocha. Inferring user intent using bayesian theory of mind in shared avatar-agent virtual environments. *IEEE transactions on visualization and computer graphics*, 25(5):2113–2122, 2019.

[159] S. Neyret, A. I. Bellido Rivas, X. Navarro, and M. Slater. Which body would you like to have? the impact of embodied perspective on body perception and body evaluation in immersive virtual reality. *Frontiers in Robotics and AI*, 7:31, 2020.

[160] R. Niiyama, M. Ikeda, and Y. A. Seong. Inflatable humanoid cybernetic avatar for physical human–robot interaction. *International Journal of Automation Technology*, 17(3):277–283, 2023.

[161] Y. Niu, D. Wang, Z. Wang, F. Sun, K. Yue, and N. Zheng. User experience evaluation in virtual reality based on subjective feelings and physiological signals. *Journal of Imaging Science and Technology*, 63(6):60413–1, 2019.

[162] S. A. Nooij, P. Pretto, D. Oberfeld, H. Hecht, and H. H. Bülthoff. Vection is the main contributor to motion sickness induced by visual yaw rotation: Implications for conflict and eye movement theories. *PloS one*, 12(4):e0175305, 2017.

[163] K. L. Nowak and J. Fox. Avatars and computer-mediated communication: a review of the definitions, uses, and effects of digital representations. *Review of Communication Research*, 6: 30–53, 2018.

[164] M. Offroy and L. Duponchel. Topological data analysis: A promising big data exploration tool in biology, analytical chemistry and physical chemistry. *Analytica chimica acta*, 910:1–11, 2016.

[165] D. Ofori-Boateng, I. S. Dominguez, C. Akcora, M. Kantarcioglu, and Y. R. Gel. Topological anomaly detection in dynamic multilayer blockchain networks. In *Machine Learning and Knowledge Discovery in Databases. Research Track: European Conference, ECML PKDD 2021, Bilbao, Spain, September 13–17, 2021, Proceedings, Part I 21*, pages 788–804. Springer, 2021.

[166] N. Ogawa, T. Narumi, and M. Hirose. Virtual hand realism affects object size perception in body-based scaling. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 519–528. IEEE, 2019.

[167] F. O. Ohanuba, M. T. Ismail, M. K. M. Ali, E. Alih, and P. N. Ezra. On the cluster validity test (s) in unsupervised machine learning tda approach for atmospheric river patterns on flood detection in nigeria. 2021.

[168] A. Oyanagi and R. Ohmura. Transformation to a bird: overcoming the height of fear by inducing the proteus effect of the bird avatar. In *Proceedings of the 2nd International Conference on Image and Graphics Processing*, ICIGP '19, page 145–149, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450360920. doi:10.1145/3313950.3313976. URL https://doi.org/10.1145/3313950.3313976.

[169] E. Pacherie. The sense of control and the sense of agency. *Psyche*, 13(1):1–30, 2007.

[170] B. Pätzold, A. Rochow, M. Schreiber, R. Memmesheimer, C. Lenz, M. Schwarz, and S. Behnke. Audio-based roughness sensing and tactile feedback for haptic perception in telepresence. In *2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 1387–1392. IEEE, 2023.

[171] E. F. Pavone, G. Tieri, G. Rizza, E. Tidoni, L. Grisoni, and S. M. Aglioti. Embodying others in immersive virtual reality: electro-cortical signatures of monitoring the errors in the actions of an avatar seen from a first-person perspective. *Journal of Neuroscience*, 36(2):268–279, 2016.

[172] T. Piumsomboon, G. A. Lee, J. D. Hart, B. Ens, R. W. Lindeman, B. H. Thomas, and M. Billinghurst. Mini-me: An adaptive avatar for mixed reality remote collaboration. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, pages 1–13, 2018.

[173] J. Plouzeau, J.-R. Chardonnet, and F. Merienne. Using cybersickness indicators to adapt navigation in virtual reality: a pre-study. In *2018 IEEE conference on virtual reality and 3D user interfaces (VR)*, pages 661–662. IEEE, 2018.

[174] B. Porras-Garcia, E. Serrano-Troncoso, M. Carulla-Roig, P. Soto-Usera, M. Ferrer-Garcia, N. Figueras-Puigderrajols, L. Yilmaz, Y. Onur Sen, N. Shojaeian, and J. Gutiérrez-Maldonado. Virtual reality body exposure therapy for anorexia nervosa. a case report with follow-up results. *Frontiers in Psychology*, 11:956, 2020.

[175] T. Porssut, Y. Hou, O. Blanke, B. Herbelin, and R. Boulic. Adapting virtual embodiment through reinforcement learning. *IEEE Transactions on Visualization and Computer Graphics*, 28(9):3193–3205, 2021.

[176] A. S. Praetorius and D. Görlich. The proteus effect: how avatars influence their users' self-perception and behaviour. In *Augmented Reality and Virtual Reality: New Trends in Immersive Technology*, pages 109–122. Springer, 2021.

[177] U. Proske and S. C. Gandevia. The proprioceptive senses: their roles in signaling body shape, body position and movement, and muscle force. *Physiological reviews*, 2012.

[178] X. Qian, F. He, X. Hu, T. Wang, A. Ipsita, and K. Ramani. Scalar: Authoring semantically adaptive augmented reality experiences in virtual reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pages 1–18, 2022.

[179] R. Rabadán and A. J. Blumberg. *Topological data analysis for genomics and evolution: topology in biology*. Cambridge University Press, 2019.

[180] S. Rangelova, D. Motus, and E. André. Cybersickness among gamers: an online survey. In *Advances in Human Factors in Wearable Technologies and Game Design: Proceedings of the AHFE 2019 International Conference on Human Factors and Wearable Technologies, and the AHFE International Conference on Game Design and Virtual Environments, July 24-28, 2019, Washington DC, USA 10*, pages 192–201. Springer, 2020.

[181] R. Ratan. Cars and contemporary communication| when automobiles are avacars: A self-other-utility approach to cars and avatars. *International Journal of Communication*, 13:19, 2019.

[182] N. Ravishanker and R. Chen. Topological data analysis (tda) for time series. *arXiv preprint arXiv:1909.10604*, 2019.

[183] J. T. Reason. Motion sickness adaptation: a neural mismatch model. *Journal of the royal society of medicine*, 71(11):819–829, 1978.

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

[185] E. Richter. Quantitative study of human scarpa's ganglion and vestibular sensory epithelia. *Acta oto-laryngologica*, 90(1-6):199–208, 1980.

[186] J. A. Rincon, E. Garcia, V. Julian, and C. Carrascosa. Developing adaptive agents situated in intelligent virtual environments. In *Hybrid Artificial Intelligence Systems: 9th International Conference, HAIS 2014, Salamanca, Spain, June 11-13, 2014. Proceedings 9*, pages 98–109. Springer, 2014.

[187] R. Rivera-Castro, P. Pilyugina, A. Pletnev, I. Maksimov, W. Wyz, and E. Burnaev. Topological data analysis of time series data for b2b customer relationship management. *arXiv preprint arXiv:1906.03956*, 2019.

[188] N. Rizzo. *Topological Data Analysis for Evaluation of Network Security Data.* State University of New York at Albany, 2020.

[189] A. Robb, A. Kleinsmith, A. Cordar, C. White, A. Wendling, S. Lampotang, and B. Lok. Training together: how another human trainee's presence affects behavior during virtual human-based team training. *Frontiers in ICT*, 3:17, 2016.

[190] T. B. Rodrigues, C. Ó. Catháin, N. E. O'Connor, and N. Murray. A qoe evaluation of haptic and augmented reality gait applications via time and frequency-domain electrodermal activity (eda) analysis. *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 297–302, 2022. URL https://api.semanticscholar.org/CorpusID:254736786.

[191] B. Rossen and B. Lok. A crowdsourcing method to develop virtual human conversational agents. *International Journal of Human-Computer Studies*, 70(4):301–319, 2012.

[192] D. Roth, J.-L. Lugrin, D. Galakhov, A. Hofmann, G. Bente, M. E. Latoschik, and A. Fuhrmann. Avatar realism and social interaction quality in virtual reality. In *2016 IEEE virtual reality (VR)*, pages 277–278. IEEE, 2016.

[193] M. Salehi, N. Javadpour, B. Beisner, M. Sanaei, and S. Gilbert. Cybersickness detection through head movement patterns: A promising approach. 2024.

[194] D. P. Salgado, R. Flynn, E. L. M. Naves, and N. Murray. The impact of jerk on quality of experience and cybersickness in an immersive wheelchair application. *2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)*, pages 1–6, 2020. URL https://api.semanticscholar.org/CorpusID:220078921.

[195] E. U. Samani and A. G. Banerjee. Persistent homology meets object unity: Object recognition in clutter. *IEEE Transactions on Robotics*, 2023.

[196] E. U. Samani, X. Yang, and A. G. Banerjee. Visual object recognition in indoor environments using topologically persistent features. *IEEE Robotics and Automation Letters*, 6(4):7509–7516, 2021.

[197] J. R. Schoenherr, R. Abbas, K. Michael, P. Rivas, and T. D. Anderson. Designing ai using a human-centered approach: Explainability and accuracy toward trustworthiness. *IEEE Transactions on Technology and Society*, 4(1):9–23, 2023.

[198] V. Schwind, P. Knierim, C. Tasci, P. Franczak, N. Haas, and N. Henze. " these are not my hands!" effect of gender on the perception of avatar hands in virtual reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 1577–1582, 2017.

[199] V. Schwind, D. Halbhuber, J. Fehle, J. Sasse, A. Pfaffelhuber, C. Tögel, J. Dietz, and N. Henze. The effects of full-body avatar movement predictions in virtual reality using neural networks. In *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology*, pages 1–11, 2020.

[200] T. Selvi, S. N. Sri, B. Devi, M. Thomas, Rajkumar, and S. P. Ramdoss. Mental immersion in virtual reality avatar (mivra)—social communication rehabilitation assistive tool for autism children. In *Proceedings of International Conference on Innovations in Information and Communication Technologies: ICI2CT 2020*, pages 13–19. Springer, 2021.

[201] J.-C. Servotte, M. Goosse, S. H. Campbell, N. Dardenne, B. Pilote, I. L. Simoneau, M. Guillaume, I. Bragard, and A. Ghuysen. Virtual reality experience: Immersion, sense of presence, and cybersickness. *Clinical Simulation in Nursing*, 38:35–43, 2020.

[202] B. Shneiderman. Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human–Computer Interaction*, 36(6):495–504, 2020.

[203] B. Shneiderman. *Human-centered AI*. Oxford University Press, 2022.

[204] S. Siehl, K. Kammler-Sücker, S. Guldner, Y. Janvier, R. Zohair, and F. Nees. To trust or not to trust? face and voice modulation of virtual avatars. *Frontiers in Virtual Reality*, 5:1301322, 2024.

[205] R. M. Silva and C. Brasil. A adaptabilidade de npcs em jogos com estilo presa-predador usando algoritmo genético. *Anais do Computer on the Beach*, 11(1):056–058, 2020.

[206] G. Singh, F. Mémoli, G. E. Carlsson, et al. Topological methods for the analysis of high dimensional data sets and 3d object recognition. *PBG@ Eurographics*, 2:091–100, 2007.

[207] Y. Singh, W. Jons, J. E. Eaton, J. D. Sobek, J. Jagtap, G. M. Conte, E. G. Fuemmeler, K. Zhang, Y. Wei, D. V. V. Garcia, et al. A new tda-based machine learning classifier framework for predicting hepatic decompensation from mr images. In *Medical Imaging 2022: Imaging Informatics for Healthcare, Research, and Applications*, volume 12037, pages 117–121. SPIE, 2022.

[208] Y. Singh, C. M. Farrelly, Q. A. Hathaway, T. Leiner, J. Jagtap, G. E. Carlsson, and B. J. Erickson. Topological data analysis in medical imaging: current state of the art. *Insights into Imaging*, 14(1):58, 2023.

[209] A. E. Sizemore, J. E. Phillips-Cremins, R. Ghrist, and D. S. Bassett. The importance of the whole: topological data analysis for the network neuroscientist. *Network Neuroscience*, 3(3): 656–673, 2019.

[210] R. Skarbez, M. Smith, and M. C. Whitton. Revisiting milgram and kishino's reality-virtuality continuum. *Frontiers in Virtual Reality*, 2:647997, 2021.

[211] F. Škola, S. Rizvić, M. Cozza, L. Barbieri, F. Bruno, D. Skarlatos, and F. Liarokapis. Virtual reality with 360-video storytelling in cultural heritage: Study of presence, engagement, and immersion. *Sensors*, 20(20):5851, 2020.

[212] M. Slater. A note on presence terminology. *Presence connect*, 3(3):1–5, 2003.

[213] M. Slater. Immersion and the illusion of presence in virtual reality. *British journal of psychology*, 109(3):431–433, 2018.

[214] M. Slater, B. Spanlang, M. V. Sanchez-Vives, and O. Blanke. First person experience of body transfer in virtual reality. *PloS one*, 5(5):e10564, 2010.

[215] J. J. Slob. State-of-the-art driving simulators, a literature survey. *DCT report*, 107, 2008.

[216] A. D. Smith, P. Dłotko, and V. M. Zavala. Topological data analysis: concepts, computation, and applications in chemical engineering. *Computers & Chemical Engineering*, 146:107202, 2021.

[217] A. G. Solimini, A. Mannocci, D. Di Thiene, and G. La Torre. A survey of visually induced symptoms and associated factors in spectators of three dimensional stereoscopic movies. *BMC public health*, 12:1–11, 2012.

[218] A. Somrak, I. Humar, M. S. Hossain, M. F. Alhamid, M. A. Hossain, and J. Guna. Estimating vr sickness and user experience using different hmd technologies: An evaluation study. *Future Generation Computer Systems*, 94:302–316, 2019.

[219] M. Sra. Enhancing the sense of presence in virtual reality. *IEEE Computer Graphics and Applications*, 43(4):90–96, 2023.

[220] 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)*, pages 121–127. IEEE, 2018.

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

[222] J. Steuer, F. Biocca, M. R. Levy, et al. Defining virtual reality: Dimensions determining telepresence. *Communication in the age of virtual reality*, 33:37–39, 1995.

[223] T. A. Stoffregen and G. E. Riccio. An ecological critique of the sensory conflict theory of motion sickness. *Ecological Psychology*, 3(3):159–194, 1991.

[224] T. A. Stoffregen and L. J. Smart Jr. Postural instability precedes motion sickness. *Brain research bulletin*, 47(5):437–448, 1998.

[225] G. Strang. *Linear algebra and its applications.* Belmont, CA: Thomson, Brooks/Cole, 2006.

[226] S. Subramanian, P. L. Purdon, R. Barbieri, and E. N. Brown. Elementary integrate-and-fire process underlies pulse amplitudes in electrodermal activity. *PLoS Computational Biology*, 17 (7):e1009099, 2021.

[227] N. Sun and J. Botev. Intelligent autonomous agents and trust in virtual reality. *Computers in Human Behavior Reports*, 4:100146, 2021.

[228] A. Surely, S. Taherzadeh, V. Misal, and A. Kleinsmith. Exploring affective dimension perception from bodily expressions and electrodermal activity in paramedic simulation training. *2022 10th International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 1–8, 2022. URL https://api.semanticscholar.org/CorpusID:253881939.

[229] I. E. Sutherland. A head-mounted three dimensional display. In *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*, pages 757–764, 1968.

[230] G. Tauzin, U. Lupo, L. Tunstall, J. B. Pérez, M. Caorsi, A. M. Medina-Mardones, A. Dassatti, and K. Hess. giotto-tda:: A topological data analysis toolkit for machine learning and data exploration. *Journal of Machine Learning Research*, 22(39):1–6, 2021.

[231] J. Teixeira and S. Palmisano. Effects of dynamic field-of-view restriction on cybersickness and presence in hmd-based virtual reality. *Virtual Reality*, 25(2):433–445, 2021.

[232] N. Tian, P. Lopes, and R. Boulic. A review of cybersickness in head-mounted displays: raising attention to individual susceptibility. *Virtual Reality*, 26(4):1409–1441, 2022.

[233] J. Tierny, G. Favelier, J. A. Levine, C. Gueunet, and M. Michaux. The topology toolkit. *IEEE transactions on visualization and computer graphics*, 24(1):832–842, 2017.

[234] I. Y. Torshin and K. Rudakov. Topological data analysis in materials science: the case of high-temperature cuprate superconductors. *Pattern recognition and image analysis*, 30(2):264–276, 2020.

[235] M. Treisman. Motion sickness: an evolutionary hypothesis. *Science*, 197(4302):493–495, 1977.

[236] S. Tymochko, J. Chaput, T. Doster, E. Purvine, J. Warley, and T. Emerson. Con connections: Detecting fraud from abstracts using topological data analysis. In *2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 403–408. IEEE, 2021.

[237] S. Valluripally, V. Akashe, M. Fisher, D. Falana, K. A. Hoque, and P. Calyam. Rule-based adaptations to control cybersickness in social virtual reality learning environments. In *2021 8th International Conference on Future Internet of Things and Cloud (FiCloud)*, pages 350–358. IEEE, 2021.

[238] M. Van Der Boon, L. Fermoselle, F. Ter Haar, S. Dijkstra-Soudarissanane, and O. Niamut. Deep learning augmented realistic avatars for social vr human representation. In *Proceedings of the 2022 ACM International Conference on Interactive Media Experiences*, pages 311–318, 2022.

[239] N. E. van der Waal, J. A. van Bokhorst, and L. N. van der Laan. Identifying emotions toward an overweight avatar in virtual reality: The moderating effects of visuotactile stimulation and drive for thinness. *Frontiers in Virtual Reality*, 3:989676, 2022.

[240] A. Vasalou and A. N. Joinson. Me, myself and i: The role of interactional context on self-presentation through avatars. *Computers in human behavior*, 25(2):510–520, 2009.

[241] N. Wagener, A. Ackermann, G.-L. Savino, B. Dänekas, J. Niess, and J. Schöning. Influence of passive haptic and auditory feedback on presence and mindfulness in virtual reality environments. In *Proceedings of the 2022 International Conference on Multimodal Interaction*, pages 558–569, 2022.

[242] T. Waltemate, D. Gall, D. Roth, M. Botsch, and M. E. Latoschik. The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response. *IEEE transactions on visualization and computer graphics*, 24(4):1643–1652, 2018.

[243] T. Waltemate, D. Gall, D. Roth, M. Botsch, and M. E. Latoschik. The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response. *IEEE transactions on visualization and computer graphics*, 24(4):1643–1652, 2018.

[244] J. Wang, Y. Qiu, K. Chen, Y. Ding, and Y. Pan. Fully automatic blendshape generation for stylized characters. In *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pages 347–355. IEEE, 2023.

[245] R. Warp, M. Zhu, I. Kiprijanovska, J. Wiesler, S. Stafford, and I. Mavridou. Validating the effects of immersion and spatial audio using novel continuous biometric sensor measures for virtual reality. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 262–265. IEEE, 2022.

[246] N. A. Webb and M. J. Griffin. Optokinetic stimuli: Motion sickness, visual acuity and eye movements. *Aviation, Space and Environmental Medicine*, 73(4):351–358, 2002.

[247] D. B. West et al. *Introduction to graph theory*, volume 2. Prentice hall Upper Saddle River, 2001.

[248] E. Widmaier, H. Raff, K. Strang, and A. Vander. *Vander's Human Physiology: The Mechanisms of Body Function*. McGraw-Hill Higher Education, Boston, MA, 2008.

[249] R. Wilkie and J. Wann. Controlling steering and judging heading: retinal flow, visual direction, and extraretinal information. *Journal of Experimental Psychology: Human Perception and Performance*, 29(2):363, 2003.

[250] C. A. Wingrave, D. A. Bowman, and N. Ramakrishnan. Towards preferences in virtual environment interfaces. In *EGVE*, volume 2, pages 63–72, 2002.

[251] B. Xie, H. Liu, R. Alghofaili, Y. Zhang, Y. Jiang, F. D. Lobo, C. Li, W. Li, H. Huang, M. Akdere, et al. A review on virtual reality skill training applications. *Frontiers in Virtual Reality*, 2:645153, 2021.

[252] E. Yarnell. Herbs for motion sickness. *Alternative and Complementary Therapies*, 22(2):74–78, 2016.

[253] N. Yee, J. N. Bailenson, and N. Ducheneaut. The proteus effect: Implications of transformed digital self-representation on online and offline behavior. *Communication Research*, 36(2):285–312, 2009.

[254] C. Yildirim. An immersive model of user trust in conversational agents in virtual reality. *2021 Third International Conference on Transdisciplinary AI (TransAI)*, pages 17–18, 2021. URL https://api.semanticscholar.org/CorpusID:239041250.

[255] K. Yin, E.-L. Hsiang, J. Zou, Y. Li, Z. Yang, Q. Yang, P.-C. Lai, C.-L. Lin, and S.-T. Wu. Advanced liquid crystal devices for augmented reality and virtual reality displays: principles and applications. *Light: Science & Applications*, 11(1):161, 2022.

[256] W. Yu, Y. Fan, Y. Zhang, X. Wang, F. Yin, Y. Bai, Y.-P. Cao, Y. Shan, Y. Wu, Z. Sun, et al. Nofa: Nerf-based one-shot facial avatar reconstruction. In *ACM SIGGRAPH 2023 Conference Proceedings*, pages 1–12, 2023.

[257] P. Zaal, F. Nieuwenhuizen, M. Mulder, and M. M. Van Paassen. Perception of visual and motion cues during control of self-motion in optic flow environments. volume 2, 08 2006. doi:10.2514/6.2006-6627.

[258] A. Zaman, M. R. Abir, and S. Mursalin. Extended reality in education and training: Enhancing trustworthiness. *International Journal of Science and Research Archive*, 2024. URL https://api.semanticscholar.org/CorpusID:267718595.

[259] F. Zhang, T.-Y. Wu, J.-S. Pan, G. Ding, and Z. Li. Human motion recognition based on svm in vr art media interaction environment. *Human-centric Computing and Information Sciences*, 9(1):40, 2019.

[260] K. Zhang, E. Deldari, Y. Yao, and Y. Zhao. A diary study in social virtual reality: Impact of avatars with disability signifiers on the social experiences of people with disabilities. In *Proceedings of the 25th International ACM SIGACCESS Conference on Computers and Accessibility*, pages 1–17, 2023.

[261] L. Zhang, B. Verma, D. Tjondronegoro, and V. Chandran. Facial expression analysis under partial occlusion: A survey. *ACM Computing Surveys (CSUR)*, 51(2):1–49, 2018.

[262] S.-H. Zhang, C.-H. Chen, F. Zheng, Y.-L. Yang, and S.-M. Hu. Adaptive optimization algorithm for resetting techniques in obstacle-ridden environments. *IEEE Transactions on Visualization and Computer Graphics*, 29(4):2080–2092, 2022.

[263] Y. Zhang and S.-B. Tsai. Application of adaptive virtual reality with ai-enabled techniques in modern sports training. *Mobile Information Systems*, 2021:1–10, 2021.

[264] J. Zhao, Z. Wang, and Y. Wang. Effects of different levels of self-representation on spatial awareness, self-presence and spatial presence during virtual locomotion. In *2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 4128–4133. IEEE, 2023.

[265] X. Zhu. Persistent homology: An introduction and a new text representation for natural language processing. In *IJCAI*, number 2013, pages 1953–1959, 2013.

[266] A. Zomorodian and G. Carlsson. Computing persistent homology. In *Proceedings of the twentieth annual symposium on Computational geometry*, pages 347–356, 2004.

Appendix A

Experiment Material, General Information Form, and SSQ

Figure A.1: Recruitment Flyer: This flyer was designed and distributed on social media platforms as well as within the university community to invite participants for the main study.

Figure A.2: First pictogram: This visual aid was designed to help participants understand and identify the various sickness indicators.

Status Reporting

While driving, a question will be asked verbally every minute: **"What is your score?"**

Upon hearing the question, promptly express your current status loud and clearly using the following pattern: ***"Number x."*** The value of ***'x'*** should correspond to a score ranging from 0 to 3, based on the following symptom assessment:

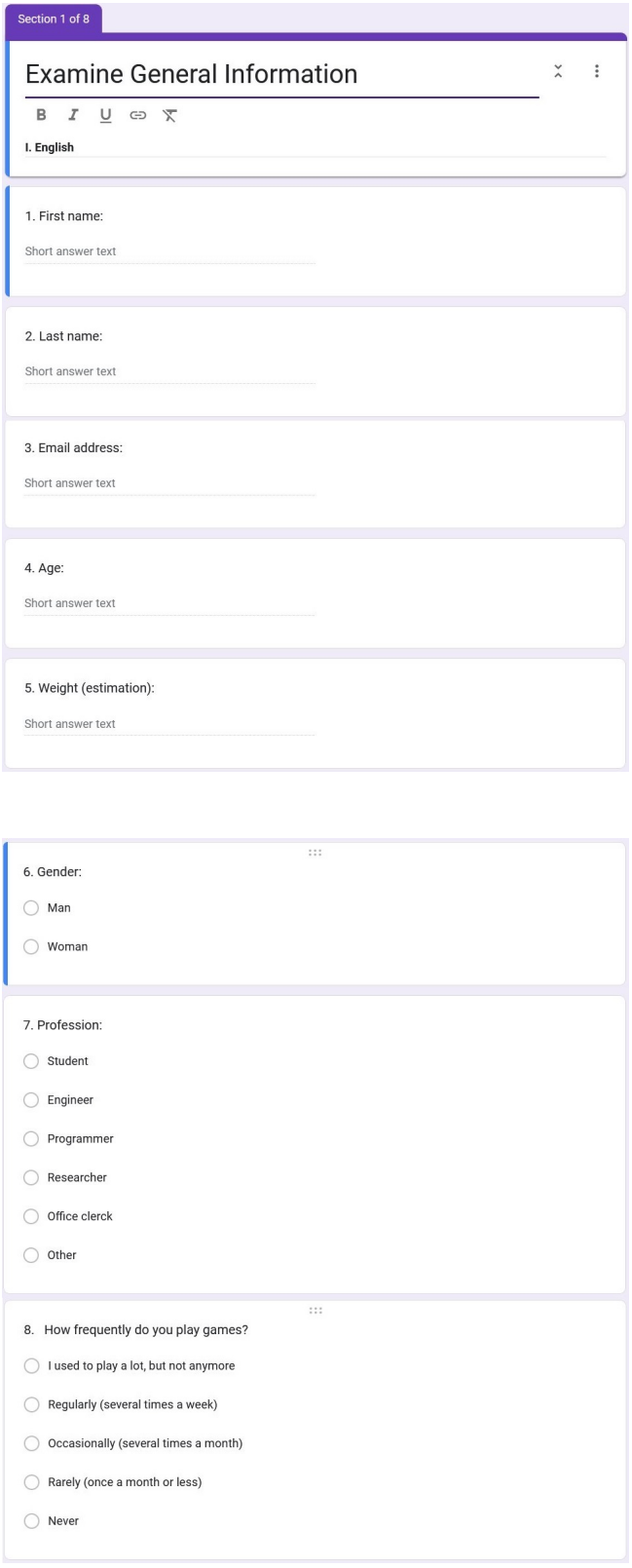
State	Description (symptoms and detail)	Score
Stage 0 (OK)	No symptoms	0
Stage 1 (Initial) (Not OK)	Fatigue Headache Eye strain	1
Stage 2 (Moderate) (Not OK)	General discomfort Difficulty concentrating Stomach awareness Increase of salivation Sweating Burping	2
Stage 3 (Severe) (Not OK)	Difficulty focusing Fullness of head Blurred vision Vertigo Dizziness (eye open) Dizziness (eye closed) Nausea	3

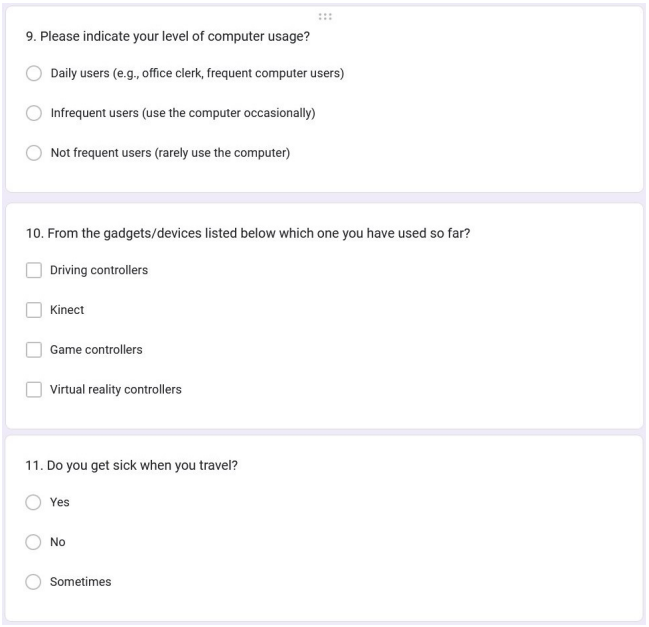
Note:

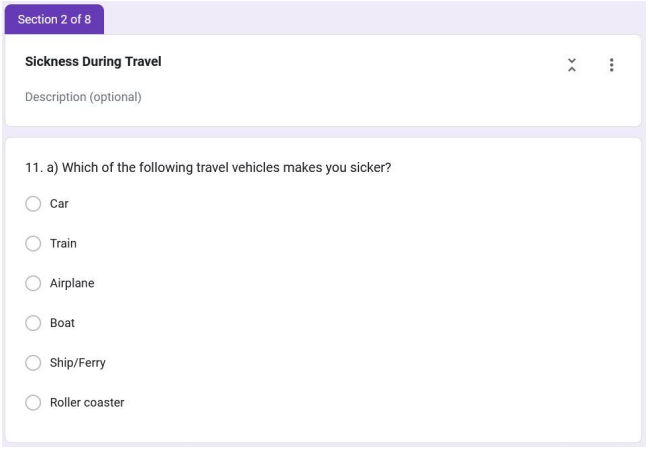
- Please do not speak other than **your current status** when the question is been asked.

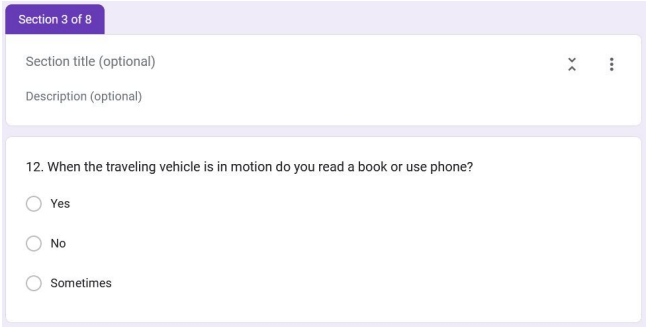
Figure A.3: Second pictogram: Self-reporting score and correlated sickness symptoms.

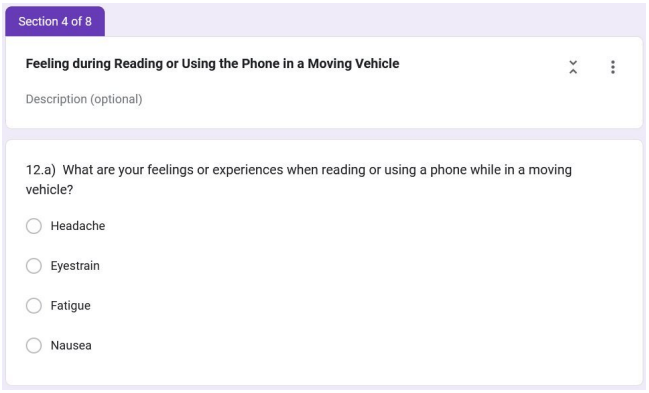
Figure A.4: Third pictogram: User interface.

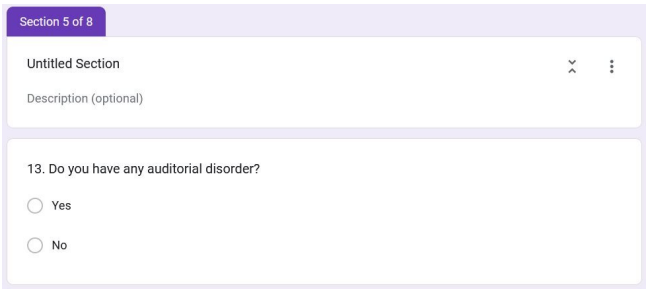

Section 1 of 8
Examine General Information
I. English
1. First name:
Short answer text
2. Last name:
Short answer text
3. Email address:
Short answer text
4. Age:
Short answer text
5. Weight (estimation):
Short answer text
6. Gender:
Man
Woman
7. Profession:
Student
Engineer
Programmer
Researcher
Office clerck
Other
8. How frequently do you play games?
I used to play a lot, but not anymore
Regularly (several times a week)
Occasionally (several times a month)
Rarely (once a month or less)
Never

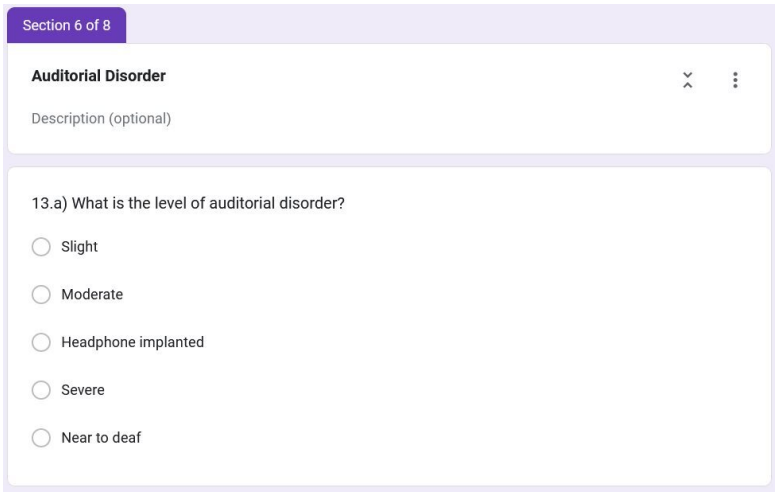

9. Please indicate your level of computer usage?
Daily users (e.g., office clerk, frequent computer users)
Infrequent users (use the computer occasionally)
Not frequent users (rarely use the computer)
10. From the gadgets/devices listed below which one you have used so far?
Driving controllers
Kinect
Game controllers
Virtual reality controllers
11. Do you get sick when you travel?
Yes
No
Sometimes

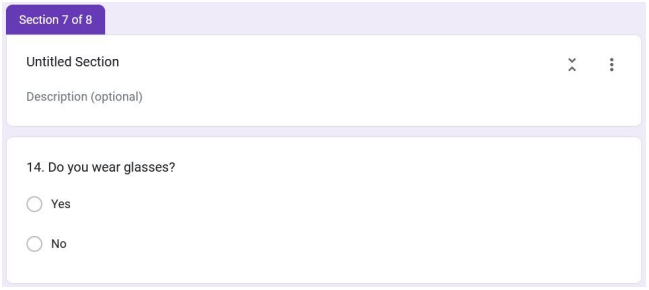

Section 2 of 8
Sickness During Travel
Description (optional)
11. a) Which of the following travel vehicles makes you sicker?
Car
Train
Airplane
Boat
Ship/Ferry
Roller coaster

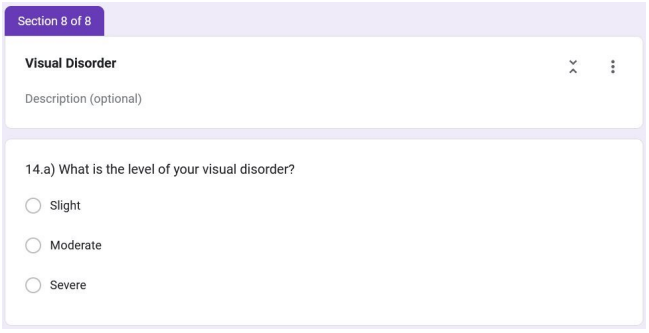

Section 3 of 8
Section title (optional)
Description (optional)
12. When the traveling vehicle is in motion do you read a book or use phone?
Yes
No
Sometimes

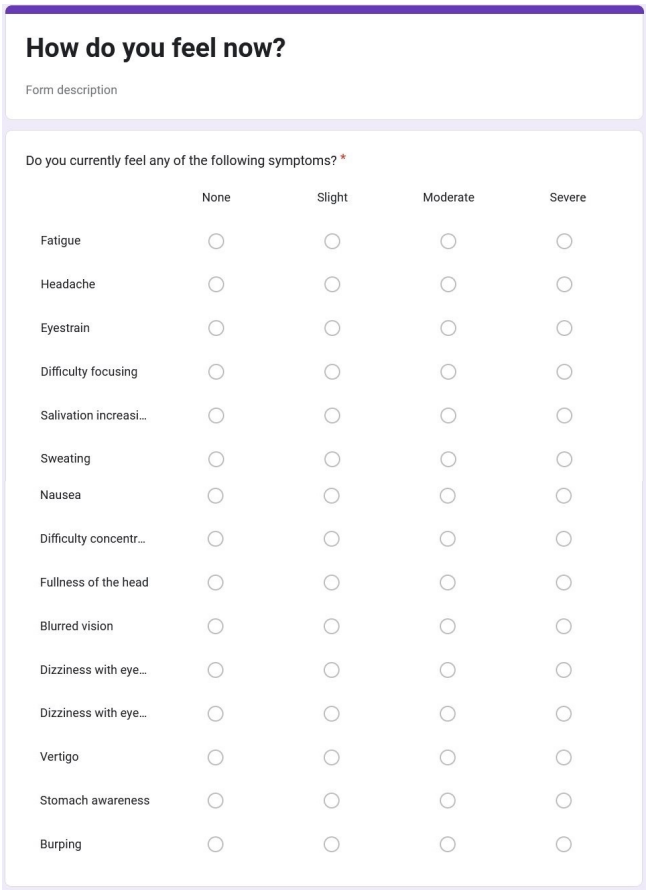

Section 4 of 8
Feeling during Reading or Using the Phone in a Moving Vehicle
Description (optional)
12.a) What are your feelings or experiences when reading or using a phone while in a moving vehicle?
Headache
Eyestrain
Fatigue
Nausea


Section 5 of 8
Untitled Section
Description (optional)
13. Do you have any auditorial disorder?
Yes
No


Section 6 of 8
Auditorial Disorder
Description (optional)
13.a) What is the level of auditorial disorder?
Slight
Moderate
Headphone implanted
Severe
Near to deaf


Section 7 of 8
Untitled Section
Description (optional)
14. Do you wear glasses?
Yes
No


Section 8 of 8
Visual Disorder
Description (optional)
14.a) What is the level of your visual disorder?
Slight
Moderate
Severe


How do you feel now?
Form description
Do you currently feel any of the following symptoms? *
None
Slight
Moderate
Severe
Fatigue
Headache
Eyestrain
Difficulty focusing
Salivation increasi...
Sweating
Nausea
Difficulty concentr...
Fullness of the head
Blurred vision
Dizziness with eye...
Dizziness with eye...
Vertigo
Stomach awareness
Burping

Appendix B

Data Analysis and Result Tables

Table B.1: Comparison of the mean and standard deviation (SD) between physiological and behavioral signals in Non-Intelligent and Intelligent avatar studies, along with results of an independent t-test. * denotes significant with p-value < 0.05.

Measurement	Mean	SD	t-value	$p < 0.05$
Eye Velocity (x) (Non-Int. Avatar)	6.8e-5	0.010	1.7	*
Eye Velocity (x) (Int. Avatar)	-1.2e-5	0.008		
Eye Velocity (y) (Non-Int. Avatar)	6.9e-5	0.004	1.5	*
Eye Velocity (y) (Int. Avatar)	-5.7e-5	0.003		
Eye Velocity (z) (Non-Int. Avatar)	-1.5e-5	0.006	-0.501	
Eye Velocity (z) (Int. Avatar)	-4.1e-5	0.005		
Eye Ang. Velocity (x) (Non-Int. Avatar)	5.1e-4	0.038	0.824	*
Eye Ang. Velocity (x) (Int. Avatar)	-1.8e-4	0.024		
Eye Ang. Velocity (y) (Non-Int. Avatar)	5.3e-4	0.072	0.797	*
Eye Ang. Velocity (y) (Int. Avatar)	-3.4e-4	0.049		
Eye Ang. Velocity (z) (Non-Int. Avatar)	-1.8e-4	0.025	0.385	
Eye Ang. Velocity (z) (Int. Avatar)	-3.1e-4	0.023		
Eye Position (x) (Non-Int. Avatar)	-0.055	0.013	1.36	
Eye Position (x) (Int. Avatar)	-0.075	0.013		
Eye Position (y) (Non-Int. Avatar)	1.147	0.004	0.481	
Eye Position (y) (Int. Avatar)	1.141	0.004		
Eye Position (z) (Non-Int. Avatar)	-0.172	0.011	1.143	
Eye Position (z) (Int. Avatar)	-0.210	0.009		
Eye Rotation (x) (Non-Int. Avatar	0.019	0.2	0.744	
Eye Rotation (x) (Int. Avatar)	0.006	0.2		
Eye Rotation (y) (Non-Int. Avatar)	0.005	0.045	-0.071	
Eye Rotation (y) (Int. Avatar)	0.005	0.043		
Eye Rotation (z) (Non-Int. Avatar)	0.023	0.018	0.949	
Eye Rotation (z) (Int. Avatar)	0.003	0.018		
Head Position (x) (Non-Int. Avatar)	-0.051	0.013	1.964	*
Head Position (x) (Int. Avatar)	-0.082	0.013		
Head Position (y) (Non-Int. Avatar)	1.147	0.005	0.481	
Head Position (y) (Int. Avatar)	1.141	0.005		
Head Position (z) (Non-Int. Avatar)	-0.172	0.011	1.143	
Head Position (z) (Int. Avatar)	-0.21	0.009		
Head Rotation (x) (Non-Int. Avatar)	0.019	0.020	0.743	
Head Rotation (x) (Int. Avatar)	0.007	0.020		
Head Rotation (y) (Non-Int. Avatar)	0.005	0.045	-0.071	
Head Rotation (y) (Int. Avatar)	0.005	0.043		
Head Rotation (z) (Non-Int. Avatar)	0.026	0.018	3.073	*
Head Rotation (z) (Int. Avatar)	0.004	0.018		
Head Rotation (w) (Non-Int. Avatar)	-0.677	0.004	1.413	
Head Rotation (w) (Int. Avatar)	-0.916	0.004		
ACC (x) (Non-Int. Avatar)	-49.093	9.834	0.212	
ACC (x) (Int. Avatar)	-49.742	9.035		
ACC (y) (Non-Int. Avatar)	-17.643	6.627	0.148	
ACC (y) (nt. Avatar)	-18.098	6.625		
ACC (z) (Non-Int. Avatar)	29.893	13.555	2.029	*
ACC (z) (Int. Avatar)	21.511	10.431		
EDA (Non-Int. Avatar)	2.276	0.480	2.119	*
EDA (Int. Avatar)	1.204	0.198		
BVP (Non-Int. Avatar)	0.1	50.068	0.93	
BVP (Int. Avatar)	-0.921	40.444		
TMP (Non-Int. Avatar)	33.234	0.213	0.626	
TMP (Int. Avatar)	32.895	0.113		

Table B.2: Comparison of the mean and standard deviation (SD) of RMS between Non-Intelligent and Intelligent avatar studies, along with results of an independent t-test. * denotes significant with p-value < 0.05.

Measurement	RMS (Mean)	RMS (SD)	t-value	$p < 0.05$
Eye Velocity (x) (Non-Int. Avatar)	0.010	0.005	1.975	*
Eye Velocity (x) (Int. Avatar)	0.007	0.003		
Eye Velocity (y) (Non-Int. Avatar)	0.004	0.002	2.054	*
Eye Velocity (y) (Int. Avatar)	0.003	0.001		
Eye Velocity (z) (Non-Int. Avatar)	0.007	0.004	0.012	
Eye Velocity (z) (Int. Avatar)	0.007	0.002		
Eye Ang. Velocity (x) (Non-Int. Avatar)	0.039	0.022	2.177	*
Eye Ang. Velocity (x) (Int. Avatar)	0.026	0.014		
Eye Ang. Velocity (y) (Non-Int. Avatar)	0.072	0.047	2.040	*
Eye Ang. Velocity (y) (Int. Avatar)	0.049	0.018		
Eye Ang. Velocity (z) (Non-Int. Avatar)	0.030	0.014	2.057	*
Eye Ang. Velocity (z) (Int. Avatar)	0.022	0.009		
Eye Position (x) (Non-Int. Avatar)	0.078	0.035	0.016	
Eye Position (x) (Int. Avatar)	0.077	0.034		
Eye Position (y) (Non-Int. Avatar)	1.157	0.040	2.272	*
Eye Position (y) (Int. Avatar)	1.130	0.037		
Eye Position (z) (Non-Int. Avatar)	0.203	0.092	0.049	
Eye Position (z) (Int. Avatar)	0.201	0.080		
Eye Rotation (x) (Non-Int. Avatar	0.065	0.033	2.986	*
Eye Rotation (x) (Int. Avatar)	0.040	0.020		
Eye Rotation (y) (Non-Int. Avatar)	0.054	0.026	0.18	
Eye Rotation (y) (Int. Avatar)	0.053	0.017		
Eye Rotation (z) (Non-Int. Avatar)	0.033	0.019	0.163	
Eye Rotation (z) (Int. Avatar)	0.032	0.013		
Head Position (x) (Non-Int. Avatar)	0.078	0.036	0.016	
Head Position (x) (Int. Avatar)	0.077	0.034		
Head Position (y) (Non-Int. Avatar)	1.153	0.043	2.007	*
Head Position (y) (Int. Avatar)	1.129	0.033		
Head Position (z) (Non-Int. Avatar)	0.207	0.093	0.334	
Head Position (z) (Int. Avatar)	0.198	0.080		
Head Rotation (x) (Non-Int. Avatar)	0.064	0.032	2.663	*
Head Rotation (x) (Int. Avatar)	0.042	0.022		
Head Rotation (y) (Non-Int. Avatar)	0.053	0.025	0.028	
Head Rotation (y) (Int. Avatar)	0.052	0.017		
Head Rotation (z) (Non-Int. Avatar)	0.033	0.019	0.202	
Head Rotation (z) (Int. Avatar)	0.032	0.016		
Head Rotation (w) (Non-Int. Avatar)	0.995	0.040	-0.506	
Head Rotation (w) (Int. Avatar)	0.995	0.030		
ACC (x) (Non-Int. Avatar)	50.591	10.547	0.238	
ACC (x) (Int. Avatar)	49.947	8.455		
ACC (y) (Non-Int. Avatar)	21.563	8.491	0.783	
ACC (y) (nt. Avatar)	19.799	6.684		
ACC (z) (Non-Int. Avatar)	30.847	12.754	0.859	
ACC (z) (Int. Avatar)	27.924	9.585		
EDA (Non-Int. Avatar)	2.319	1.883	1.229	*
EDA (Int. Avatar)	1.230	1.558		
BVP (Non-Int. Avatar)	50.624	30.570	0.960	
BVP (Int. Avatar)	42.284	26.939		
TMP (Non-Int. Avatar)	33.234	1.822	0.626	
TMP (Int. Avatar)	32.896	1.764		

Table B.3: Comparison of the mean and standard deviation (SD) of PSD between Non-Intelligent and Intelligent avatar studies, along with results of an independent t-test. * denotes significant with p-value < 0.05.

Measurement	PSD (Mean)	PSD (SD)	t-value	$p < 0.05$
Eye Velocity (x) (Non-Int. Avatar)	5.47e-5	5.109e-5	2.076	*
Eye Velocity (x) (Int. Avatar)	2.97e-5	1.740e-5		
Eye Velocity (y) (Non-Int. Avatar)	8.45e-6	8.282e-6	2.064	*
Eye Velocity (y) (Int. Avatar)	4.29e-6	3.591e-6		
Eye Velocity (z) (Non-Int. Avatar)	2.040e-5	2.685e-5	0.113	
Eye Velocity (z) (Int. Avatar)	1.955e-5	2.593e-5		
Eye Ang. Velocity (x) (Non-Int. Avatar)	6.3e-4	6.373e-4	2.069	*
Eye Ang. Velocity (x) (Int. Avatar)	3e-4	2.271e-4		
Eye Ang. Velocity (y) (Non-Int. Avatar)	0.0032	0.0041	2.136	*
Eye Ang. Velocity (y) (Int. Avatar)	0.0012	8.122e-4		
Eye Ang. Velocity (z) (Non-Int. Avatar)	3.85e-4	4.920e-4	0.769	
Eye Ang. Velocity (z) (Int. Avatar)	3.03e-4	2.149e-4		
Eye Position (x) (Non-Int. Avatar)	0.004	0.003	0.577	
Eye Position (x) (Int. Avatar)	0.003	0.003		
Eye Position (y) (Non-Int. Avatar)	0.662	0.049	2.065	*
Eye Position (y) (Int. Avatar)	0.634	0.037		
Eye Position (z) (Non-Int. Avatar)	0.025	0.018	0.561	
Eye Position (z) (Int. Avatar)	0.023	0.015		
Eye Rotation (x) (Non-Int. Avatar	0.003	0.002	3.061	*
Eye Rotation (x) (Int. Avatar)	0.001	9.051e-4		
Eye Rotation (y) (Non-Int. Avatar)	0.002	0.002	0.470	
Eye Rotation (y) (Int. Avatar)	0.001	8.695e-4		
Eye Rotation (z) (Non-Int. Avatar)	0.007	8.503e-4	0.103	
Eye Rotation (z) (Int. Avatar)	0.006	5.043e-4		
Head Position (x) (Non-Int. Avatar)	0.004	2.6e-3	0.908	
Head Position (x) (Int. Avatar)	0.003	2.5e-3		
Head Position (y) (Non-Int. Avatar)	0.66	0.051	2.271	*
Head Position (y) (Int. Avatar)	0.63	0.035		
Head Position (z) (Non-Int. Avatar)	0.026	0.019	0.771	
Head Position (z) (Int. Avatar)	0.022	0.015		
Head Rotation (x) (Non-Int. Avatar)	0.003	0.002	2.746	*
Head Rotation (x) (Int. Avatar)	0.001	0.001		
Head Rotation (y) (Non-Int. Avatar)	0.002	0.002	0.467	
Head Rotation (y) (Int. Avatar)	0.001	8.693e-4		
Head Rotation (z) (Non-Int. Avatar)	0.001	8.546e-4	0.195	
Head Rotation (z) (Int. Avatar)	0.001	4.935e-4		
Head Rotation (w) (Non-Int. Avatar)	0.493	3.5e-3	0.153	
Head Rotation (w) (Int. Avatar)	0.493	2.4e-3		
ACC (x) (Non-Int. Avatar)	1304.796	475.577	0.048	
ACC (x) (Int. Avatar)	1298.726	414.542		
ACC (y) (Non-Int. Avatar)	254.413	220.465	0.17	
ACC (y) (nt. Avatar)	245.203	157.565		
ACC (z) (Non-Int. Avatar)	553.870	399.411	1.288	
ACC (z) (Int. Avatar)	422.718	261.978		
EDA (Non-Int. Avatar)	4.615	6.707	2.294	*
EDA (Int. Avatar)	1.033	2.490		
BVP (Non-Int. Avatar)	1598.395	2.090e+3	0.245	
BVP (Int. Avatar)	1454.023	1.895e+3		
TMP (Non-Int. Avatar)	555.504	58.195	1.072	
TMP (Int. Avatar)	536.553	56.381		

Table B.4: Comparison of the mean and standard deviation (SD) of spectral entropy between physiological and behavioral signals in Non-Intelligent and Intelligent avatar studies, along with results of an independent t-test. * denotes significant with p-value < 0.05.

Measurement	Spectral Entropy (Mean)	Spectral Entropy (SD)	t-value	$p < 0.05$
Eye Velocity (x) (Non-Int. Avatar)	8.958	0.897	2.084	*
Eye Velocity (x) (Int. Avatar)	8.445	0.844		
Eye Velocity (y) (Non-Int. Avatar)	9.237	0.775	2.451	*
Eye Velocity (y) (Int. Avatar)	8.671	0.755		
Eye Velocity (z) (Non-Int. Avatar)	9.096	0.793	2.139	*
Eye Velocity (z) (Int. Avatar)	8.617	0.789		
Eye Ang. Velocity (x) (Non-Int. Avatar)	9.281	0.783	2.753	*
Eye Ang. Velocity (x) (Int. Avatar)	8.705	0.589		
Eye Ang. Velocity (y) (Non-Int. Avatar)	8.932	0.950	2.299	*
Eye Ang. Velocity (y) (Int. Avatar)	8.311	0.838		
Eye Ang. Velocity (z) (Non-Int. Avatar)	9.080	0.867	2.132	*
Eye Ang. Velocity (z) (Int. Avatar)	8.563	0.850		
Eye Position (x) (Non-Int. Avatar)	0.984	1.390	0.487	
Eye Position (x) (Int. Avatar)	0.809	1.148		
Eye Position (y) (Non-Int. Avatar)	5.11e-4	9.374e-4	0.245	
Eye Position (y) (Int. Avatar)	4.59e-4	3.926e-4		
Eye Position (z) (Non-Int. Avatar)	0.397	1.340	0.137	
Eye Position (z) (Int. Avatar)	0.351	1.017		
Eye Rotation (x) (Non-Int. Avatar	2.417	2.346	0.130	
Eye Rotation (x) (Int. Avatar)	2.339	1.420		
Eye Rotation (y) (Non-Int. Avatar)	5.669	1.359	2.301	*
Eye Rotation (y) (Int. Avatar)	4.81	1.273		
Eye Rotation (z) (Non-Int. Avatar)	3.068	1.979	0.632	
Eye Rotation (z) (Int. Avatar)	2.734	1.748		
Head Position (x) (Non-Int. Avatar)	0.982	1.388	0.485	
Head Position (x) (Int. Avatar)	0.808	1.147		
Head Position (y) (Non-Int. Avatar)	5.11e-4	9.369e-4	0.246	
Head Position (y) (Int. Avatar)	4.58e-4	3.923e-4		
Head Position (z) (Non-Int. Avatar)	0.396	1.338	0.138	
Head Position (z) (Int. Avatar)	0.350	1.015		
Head Rotation (x) (Non-Int. Avatar)	2.416	2.346	0.134	
Head Rotation (x) (Int. Avatar)	2.335	1.418		
Head Rotation (y) (Non-Int. Avatar)	5.663	1.357	2.3	*
Head Rotation (y) (Int. Avatar)	4.807	1.272		
Head Rotation (z) (Non-Int. Avatar)	3.3	2.099	2.038	*
Head Rotation (z) (Int. Avatar)	2.149	1.405		
Head Rotation (w) (Non-Int. Avatar)	6.52e-4	9.126e-4	0.426	
Head Rotation (w) (Int. Avatar)	5.51e-4	7.573e-4		
ACC (x) (Non-Int. Avatar)	0.640	0.793	0.252	
ACC (x) (Int. Avatar)	0.588	0.683		
ACC (y) (Non-Int. Avatar)	1.391	1.304	2.020	*
ACC (y) (nt. Avatar)	0.756	0.608		
ACC (z) (Non-Int. Avatar)	2.696	1.993	2.575	*
ACC (z) (Int. Avatar)	1.445	0.991		
EDA (Non-Int. Avatar)	0.397	0.285	2.363	*
EDA (Int. Avatar)	0.216	0.205		
BVP (Non-Int. Avatar)	9.41	1.186	2.183	*
BVP (Int. Avatar)	8.77	0.861		
TMP (Non-Int. Avatar)	0.001	0.003	1.013	
TMP (Int. Avatar)	7.21e-4	0.002		

Table B.5: Comparison of the mean and standard deviation (SD) of mean frequency between physiological and behavioral signals in Non-Intelligent and Intelligent avatar studies, along with results of an independent t-test. * denotes significant with p-value < 0.05.

Measurement	Mean Frequency (Mean)	Mean Frequency (SD)	t-value	$p < 0.05$
Eye Velocity (x) (Non-Int. Avatar)	0.748	0.130	2.023	*
Eye Velocity (x) (Int. Avatar)	0.681	0.092		
Eye Velocity (y) (Non-Int. Avatar)	0.879	0.118	2.620	*
Eye Velocity (y) (Int. Avatar)	0.756	0.111		
Eye Velocity (z) (Non-Int. Avatar)	0.778	0.133	0.330	
Eye Velocity (z) (Int. Avatar)	0.766	0.120		
Eye Ang. Velocity (x) (Non-Int. Avatar)	0.856	0.155	2.134	*
Eye Ang. Velocity (x) (Int. Avatar)	0.765	0.105		
Eye Ang. Velocity (y) (Non-Int. Avatar)	0.750	0.141	2.321	*
Eye Ang. Velocity (y) (Int. Avatar)	0.653	0.122		
Eye Ang. Velocity (z) (Non-Int. Avatar)	0.837	0.141	2.285	*
Eye Ang. Velocity (z) (Int. Avatar)	0.741	0.136		
Eye Position (x) (Non-Int. Avatar)	0.016	0.028	0.754	
Eye Position (x) (Int. Avatar)	0.011	0.016		
Eye Position (y) (Non-Int. Avatar)	2.51e-6	2.829e-6	0.158	
Eye Position (y) (Int. Avatar)	2.39e-6	2.258e-6		
Eye Position (z) (Non-Int. Avatar)	0.004	0.011	1.571	
Eye Position (z) (Int. Avatar)	4.54e-4	9.852e-4		
Eye Rotation (x) (Non-Int. Avatar	0.050	0.065	0.892	
Eye Rotation (x) (Int. Avatar)	0.036	0.034		
Eye Rotation (y) (Non-Int. Avatar)	0.131	0.055	2.150	*
Eye Rotation (y) (Int. Avatar)	0.095	0.051		
Eye Rotation (z) (Non-Int. Avatar)	0.072	0.074	2.150	*
Eye Rotation (z) (Int. Avatar)	0.034	0.025		
Head Position (x) (Non-Int. Avatar)	0.015	0.028	0.750	
Head Position (x) (Int. Avatar)	0.011	0.016		
Head Position (y) (Non-Int. Avatar)	2.40e-6	2.839e-6	0.578	
Head Position (y) (Int. Avatar)	2.00e-6	1.439e-6		
Head Position (z) (Non-Int. Avatar)	0.004	0.011	1.568	
Head Position (z) (Int. Avatar)	4.51e-4	9.792e-4		
Head Rotation (x) (Non-Int. Avatar)	0.048	0.064	0.547	
Head Rotation (x) (Int. Avatar)	0.039	0.037		
Head Rotation (y) (Non-Int. Avatar)	0.125	0.055	2.087	*
Head Rotation (y) (Int. Avatar)	0.090	0.051		
Head Rotation (z) (Non-Int. Avatar)	0.071	0.072	2.151	*
Head Rotation (z) (Int. Avatar)	0.034	0.025		
Head Rotation (w) (Non-Int. Avatar)	7.72e-6	1.153e-5	2.032	*
Head Rotation (w) (Int. Avatar)	2.38e-6	2.277e-6		
ACC (x) (Non-Int. Avatar)	0.008	0.008	0.038	
ACC (x) (Int. Avatar)	0.008	0.007		
ACC (y) (Non-Int. Avatar)	0.019	0.016	3.025	*
ACC (y) (nt. Avatar)	0.008	0.005		
ACC (z) (Non-Int. Avatar)	0.049	0.059	2.077	*
ACC (z) (Int. Avatar)	0.021	0.020		
EDA (Non-Int. Avatar)	0.001	8.485e-4	3.076	*
EDA (Int. Avatar)	6.42e-4	4.508e-4		
BVP (Non-Int. Avatar)	1.294	0.305	0.596	
BVP (Int. Avatar)	1.252	0.171		
TMP (Non-Int. Avatar)	1.26e-6	2.31e-6	0.369	
TMP (Int. Avatar)	1.04e-6	1.851e-6		

Table B.6: Comparison of the mean and standard deviation (SD) of power of spectrogram between physiological and behavioral signals in Non-Intelligent and Intelligent avatar studies, along with results of an independent t-test. * denotes significant with p-value < 0.05.

Measurement	Spectrogram Power (Mean)	Spectrogram Power (SD)	t-value	$p < 0.05$
Eye Velocity (x) (Non-Int. Avatar)	1.63e-5	1.7e-5	2.042	*
Eye Velocity (x) (Int. Avatar)	9.02e-6	8.6e-6		
Eye Velocity (y) (Non-Int. Avatar)	3.70e-6	3.8e-6	2.050	*
Eye Velocity (y) (Int. Avatar)	1.40e-6	1.4e-6		
Eye Velocity (z) (Non-Int. Avatar)	3.85e-6	4.0e-6	0.649	
Eye Velocity (z) (Int. Avatar)	3.38e-6	3.4e-6		
Eye Ang. Velocity (x) (Non-Int. Avatar)	2.40e-4	2.6e-4	2.006	*
Eye Ang. Velocity (x) (Int. Avatar)	1.08e-4	1.0e-4		
Eye Ang. Velocity (y) (Non-Int. Avatar)	9.36e-4	9.8e-4	2.111	*
Eye Ang. Velocity (y) (Int. Avatar)	3.52e-4	3.5e-4		
Eye Ang. Velocity (z) (Non-Int. Avatar)	1.41e-4	1.4e-4	0.559	
Eye Ang. Velocity (z) (Int. Avatar)	1.15e-4	1.1e-4		
Eye Position (x) (Non-Int. Avatar)	5.72e-6	4.4e-6	0.485	
Eye Position (x) (Int. Avatar)	5.21e-6	3.9e-6		
Eye Position (y) (Non-Int. Avatar)	6.02e-4	1.3e-4	0.510	
Eye Position (y) (Int. Avatar)	5.39e-4	1.3e-4		
Eye Position (z) (Non-Int. Avatar)	2.52e-5	7.4e-6	0.251	
Eye Position (z) (Int. Avatar)	2.26e-5	7.0e-6		
Eye Rotation (x) (Non-Int. Avatar	7.57e-6	7.3e-6	0.250	
Eye Rotation (x) (Int. Avatar)	7.06e-6	6.8e-6		
Eye Rotation (y) (Non-Int. Avatar)	3.30e-5	3.3e-5	0.673	
Eye Rotation (y) (Int. Avatar)	2.65e-5	2.9e-5		
Eye Rotation (z) (Non-Int. Avatar)	6.98e-6	6.4e-6	0.304	
Eye Rotation (z) (Int. Avatar)	6.31e-6	6.2e-6		
Head Position (x) (Non-Int. Avatar)	5.68e-6	4.3e-6	0.490	
Head Position (x) (Int. Avatar)	5.16e-6	4.2e-6		
Head Position (y) (Non-Int. Avatar)	5.91e-4	1.3e-4	0.344	
Head Position (y) (Int. Avatar)	5.50e-4	1.3e-4		
Head Position (z) (Non-Int. Avatar)	2.52e-5	7.7e-6	0.249	
Head Position (z) (Int. Avatar)	2.26e-5	7.2e-6		
Head Rotation (x) (Non-Int. Avatar)	7.37e-6	7.1e-6	0.241	
Head Rotation (x) (Int. Avatar)	6.89e-6	6.6e-6		
Head Rotation (y) (Non-Int. Avatar)	3.25e-5	3.3e-5	0.649	
Head Rotation (y) (Int. Avatar)	2.63e-5	2.9e-5		
Head Rotation (z) (Non-Int. Avatar)	6.88e-6	6.3e-6	0.287	
Head Rotation (z) (Int. Avatar)	6.25e-6	6.1e-6		
Head Rotation (w) (Non-Int. Avatar)	4.56e-4	1.1e-4	0.429	
Head Rotation (w) (Int. Avatar)	4.15e-4	1.0e-4		
ACC (x) (Non-Int. Avatar)	2.01	2.013	0.522	
ACC (x) (Int. Avatar)	1.73	1.734		
ACC (y) (Non-Int. Avatar)	0.476	0.444	1.672	
ACC (y) (nt. Avatar)	0.260	0.353		
ACC (z) (Non-Int. Avatar)	1.504	1.421	1.369	
ACC (z) (Int. Avatar)	0.877	1.205		
EDA (Non-Int. Avatar)	3.35e-3	1.1e-3	2.129	*
EDA (Int. Avatar)	1.02e-3	3.0e-4		
BVP (Non-Int. Avatar)	885.014	928.9	1.112	
BVP (Int. Avatar)	579.431	813.2		
TMP (Non-Int. Avatar)	0.448	0.100	0.057	
TMP (Int. Avatar)	0.443	0.098		

Table B.7: Comparison of the mean and standard deviation (SD) wavelet energy between physiological and behavioral signals in Non-Intelligent and Intelligent avatar studies, along with results of an independent t-test. * denotes significant with p-value < 0.05.

Measurement	Wavelet Energy (Mean)	Wavelet Energy (SD)	t-value	$p < 0.05$
Eye Velocity (x) (Non-Int. Avatar)	0.439	0.208	2.016	*
Eye Velocity (x) (Int. Avatar)	0.210	0.089		
Eye Velocity (y) (Non-Int. Avatar)	0.047	0.019	2.080	*
Eye Velocity (y) (Int. Avatar)	0.019	0.008		
Eye Velocity (z) (Non-Int. Avatar)	0.138	0.049	2.113	*
Eye Velocity (z) (Int. Avatar)	0.052	0.023		
Eye Ang. Velocity (x) (Non-Int. Avatar)	4.185	1.736	2.193	*
Eye Ang. Velocity (x) (Int. Avatar)	1.543	0.650		
Eye Ang. Velocity (y) (Non-Int. Avatar)	25.056	11.947	2.240	*
Eye Ang. Velocity (y) (Int. Avatar)	10.412	4.794		
Eye Ang. Velocity (z) (Non-Int. Avatar)	2.898	1.222	1.017	
Eye Ang. Velocity (z) (Int. Avatar)	1.986	0.868		
Eye Position (x) (Non-Int. Avatar)	1.852	1.448	2.236	*
Eye Position (x) (Int. Avatar)	1.110	0.960		
Eye Position (y) (Non-Int. Avatar)	70.323	62.188	2.049	*
Eye Position (y) (Int. Avatar)	67.402	59.686		
Eye Position (z) (Non-Int. Avatar)	3.004	2.652	0.640	
Eye Position (z) (Int. Avatar)	2.663	2.663		
Eye Rotation (x) (Non-Int. Avatar	2.726	2.136	0.876	
Eye Rotation (x) (Int. Avatar)	1.977	1.567		
Eye Rotation (y) (Non-Int. Avatar)	19.680	15.488	2.272	*
Eye Rotation (y) (Int. Avatar)	9.844	8.419		
Eye Rotation (z) (Non-Int. Avatar)	2.594	2.183	0.758	
Eye Rotation (z) (Int. Avatar)	1.844	1.597		
Head Position (x) (Non-Int. Avatar)	1.832	1.417	2.256	*
Head Position (x) (Int. Avatar)	1.050	0.906		
Head Position (y) (Non-Int. Avatar)	70.323	62.188	2.048	*
Head Position (y) (Int. Avatar)	67.403	59.686		
Head Position (z) (Non-Int. Avatar)	3.003	2.652	0.639	
Head Position (z) (Int. Avatar)	2.662	2.371		
Head Rotation (x) (Non-Int. Avatar)	2.630	2.063	2.187	*
Head Rotation (x) (Int. Avatar)	2.089	1.788		
Head Rotation (y) (Non-Int. Avatar)	18.869	14.852	2.019	*
Head Rotation (y) (Int. Avatar)	10.322	8.872		
Head Rotation (z) (Non-Int. Avatar)	2.592	2.184	0.758	
Head Rotation (z) (Int. Avatar)	1.844	1.598		
Head Rotation (w) (Non-Int. Avatar)	52.476	46.462	2.392	*
Head Rotation (w) (Int. Avatar)	52.083	46.053		
ACC (x) (Non-Int. Avatar)	645714.375	471496	2.672	*
ACC (x) (Int. Avatar)	332569.295	246672		
ACC (y) (Non-Int. Avatar)	287560.378	215842	2.140	*
ACC (y) (nt. Avatar)	126280.097	92058		
ACC (z) (Non-Int. Avatar)	1315045.537	975146	2.367	*
ACC (z) (Int. Avatar)	593301.805	455958		
EDA (Non-Int. Avatar)	757.456	616.653	2.561	*
EDA (Int. Avatar)	203.696	173.553		
BVP (Non-Int. Avatar)	3444041	11918320	2.156	*
BVP (Int. Avatar)	1050880	1786581		
TMP (Non-Int. Avatar)	58686.768	51972	1.049	
TMP (Int. Avatar)	56736.686	50213		

Table B.8: Comparison of mean of persistent value between physiological and behavioral signals in Non-Intelligent and Intelligent avatar studies, along with results of an independent t-test. * denotes significant with p-value < 0.05.

Measurement	Persistent Value Int. Avatar			Persistent Value Non-Int. Avatar			$p < 0.05$		
	H_0	H_1	H_2	H_0	H_1	H_2	H_0	H_1	H_2
Eye Velocity (x)	0.076	0.0008	0.0003	0.0053	0.0005	0.0001	*	*	*
Eye Velocity (y)			1.1e-4			5.7e-5			*
Eye Velocity (z)	0.0043	0.0004	1.4e-4	0.0027	0.0003	9.8e-5	*	*	*
Eye Angular Velocity (x)	0.0221	0.0021	0.001	0.0148	0.0014	0.0005	*	*	*
Eye Angular Velocity (y)		0.0041			0.0028			*	
Eye Angular Velocity (z)		0.0021	0.0008		0.0014	0.0004		*	*
Eye Position (x)			0.0003			0.0001			*
Eye Position (y)									
Eye Position (z)			1.1e-4			6.1e-5			*
Eye Rotation (x)			0.0004			0.0002			*
Eye Rotation (y)			9.8e-4			4.1e-4			*
Eye Rotation (z)									
Head Position (x)			3.8e-4			1.5e-4			*
Head Position (y)			8.2e-5			4.4e-5			*
Head Position (z)									
Head Rotation (x)									
Head Rotation (y)									
Head Rotation (z)									
Head Rotation (w)			6.9e-5			3.3e-5			*
ACC (x)			0.1936			0.0567			*
ACC (y)			0.185			0.054			*
ACC (z)		0.8396	0.320		0.5575	0.1699		*	*
EDA	0.0776	0.0160	0.0057	0.0281	0.0055	0.0022	*	*	*
BVP			1.4840			0.8823			*
TMP			0.0024			0.0017			*

Appendix C

Résumé Substantiel en Langue Française

C.1 Introduction

La réalité virtuelle (RV) est une technologie transformative qui offre des expériences immersives, brouillant les frontières entre le physique et le numérique. Au cœur de l'expérience RV se trouvent les avatars, qui facilitent l'engagement des utilisateurs et l'interaction entre les mondes réel et virtuel. L'efficacité des avatars à bien représenter les utilisateurs reste un sujet de débat.

Les avatars possèdent deux aspects principaux : l'apparence et la fonctionnalité. L'apparence cherche à créer des représentations fidèles des utilisateurs, favorisant la familiarité dans l'espace virtuel. Les avancées technologiques ont permis de réaliser des avatars réalistes.

La fonctionnalité concerne la capacité des avatars à reproduire les actions des utilisateurs en temps réel, renforçant ainsi le sentiment d'immersion. Cependant, il est crucial d'évaluer la sécurité et la prise en compte des états des utilisateurs. Les environnements immersifs impactent le bien-être psychologique et physiologique, et les avatars doivent réagir aux conditions émotionnelles des utilisateurs pour offrir une expérience RV enrichissante.

C.1.1 Contribution

Motivé par le besoin de combler le fossé entre les avatars et l'état et la sécurité des utilisateurs, cette thèse vise à examiner de manière critique l'état actuel des avatars en RV, en mettant l'accent sur leurs aspects fonctionnels et en se concentrant explicitement sur le développement d'avatars de première personne. En explorant les technologies émergentes et les directions de recherche, cette recherche cherche à contribuer à l'élaboration de systèmes d'avatars plus sécurisés, adaptatifs et centrés sur

l'utilisateur, qui correspondent aux besoins et préferences des utilisateurs humains dans le domaine virtuel.

Ces avatars seraient équipés d'intelligence artificielle (IA) centrée sur l'humain pour s'aligner avec l'état de l'utilisateur. Pour atteindre cet objectif, une attention particulière a été accordée à l'état physiologique et comportemental des utilisateurs. Ce modèle virtuel facilite l'auto-adaptation et renforce le sentiment d'engagement de l'utilisateur lors des expériences virtuelles. En d'autres termes, il a servi de représentation innovante et dynamique de l'utilisateur, capturant efficacement son état actuel et répondant de manière réactive à ses besoins spécifiques dans l'environnement virtuel.

La thèse apporte plusieurs contributions significatives en termes de développements scientifiques et techniques, notamment :

- Réalisation d'une revue de literature complète sur les avatars intelligents et fourniture d'une définition claire et bien définie du concept d'avatar intelligent de confiance.
- Développement d'un cadre modulaire et scalable novateur pour créer un avatar intelligent centré sur l'humain, personnalisable en temps réel et qui s'adapte dynamiquement à l'état de l'utilisateur. Le cadre utilise des techniques de fusion de capteurs pour intégrer les données provenant de plusieurs capteurs de manière transparente.
- Réalisation d'analyses de données complètes et d'études dans divers domaines, y compris des analyses statistiques et topologiques, pour enquêter sur l'impact d'un avatar intelligent fiable sur l'expérience utilisateur.

C.1.2 Structure de la Thése

La figure C.1 illustre la structure de la thèse et les interrelations entre les chapitres, centrés principalement sur les avatars.

Le chapitre 1 est une revue de la littérature sur la RV et l'immersion, explorant les avatars en RV, leur apparence, interactivité, effet Proteus, capacités fonctionnelles et fiabilité. Il aborde également les avatars intelligents et leurs caractéristiques.

Le chapitre 2 détaille l'approche de recherche, les objectifs spécifiques de la thèse et les questions de recherche.

Le chapitre 3 se concentre sur l'"Avatar Intelligent Fiable", discutant du comportement adaptatif, du modèle hybride avatar/agent, des avatars centrés sur l'humain et des défis de l'apprentissage en temps réel.

Le chapitre 4 présente l'architecture logicielle d'implémentation d'un avatar intelligent adaptatif, incluant l'application de l'IA et l'intégration des modules développés. Un cas d'usage sur la simulation de conduite est décrit ensuite pour réduire le cybermalaise. Une présentation des techniques utilisées, ainsi que des méthodes multimodales, est faite. Le chapitre est conclu par une évaluation des performances de notre architecture.

Le chapitre 5 couvre la conception expérimentale associée au cas d'usage, la collecte des données et les résultats, incluant l'analyse statistique et l'homologie persistante pour extraire des informations des données.

Le chapitre de conclusion (chapitre 6) résume les conclusions de l'étude, souligne les contributions de la recherche et propose des directions futures.

C.2 État de l'art

Avec l'avenue des plateformes numériques et d'internet, la société a commencé à chercher des moyens de s'exprimer aux autres dans des environnements en ligne. Comme la communication par texte s'est révélée limitée, un besoin de représentations visuelles est né pour transmettre son identité aux autres utilisateurs. Cela a conduit à lémergence des photos de profil, communément appelées avatars, initialement popularisées dans les jeux vidéo et les romans. Au fil du temps, le terme "avatar" a été élargi pour englober toutes les formes de représentation des utilisateurs dans les environnements virtuels. Des jeux vidéo aux plateformes de discussion en passant par la RV, les avatars servent d'identités virtuelles des utilisateurs [198]. Ils peuvent être personnalisés et agissent comme des procurations virtuelles des utilisateurs dans le monde numérique. La définition d'un "avatar" peut varier selon les communautés et les chercheurs. Nowak et Fox [163] fournissent une définition complète, affirmant qu'un avatar est une représentation numérique d'un utilisateur humain qui facilite l'interaction avec d'autres utilisateurs, entités ou l'environnement. De plus, Miao et al. [146] définissent les avatars comme des entités capables d'incarner des agents artificiels dans des environnements immersifs de RV.

Les avatars se sont rapidement développés dans le domaine de la RV, trouvant des applications

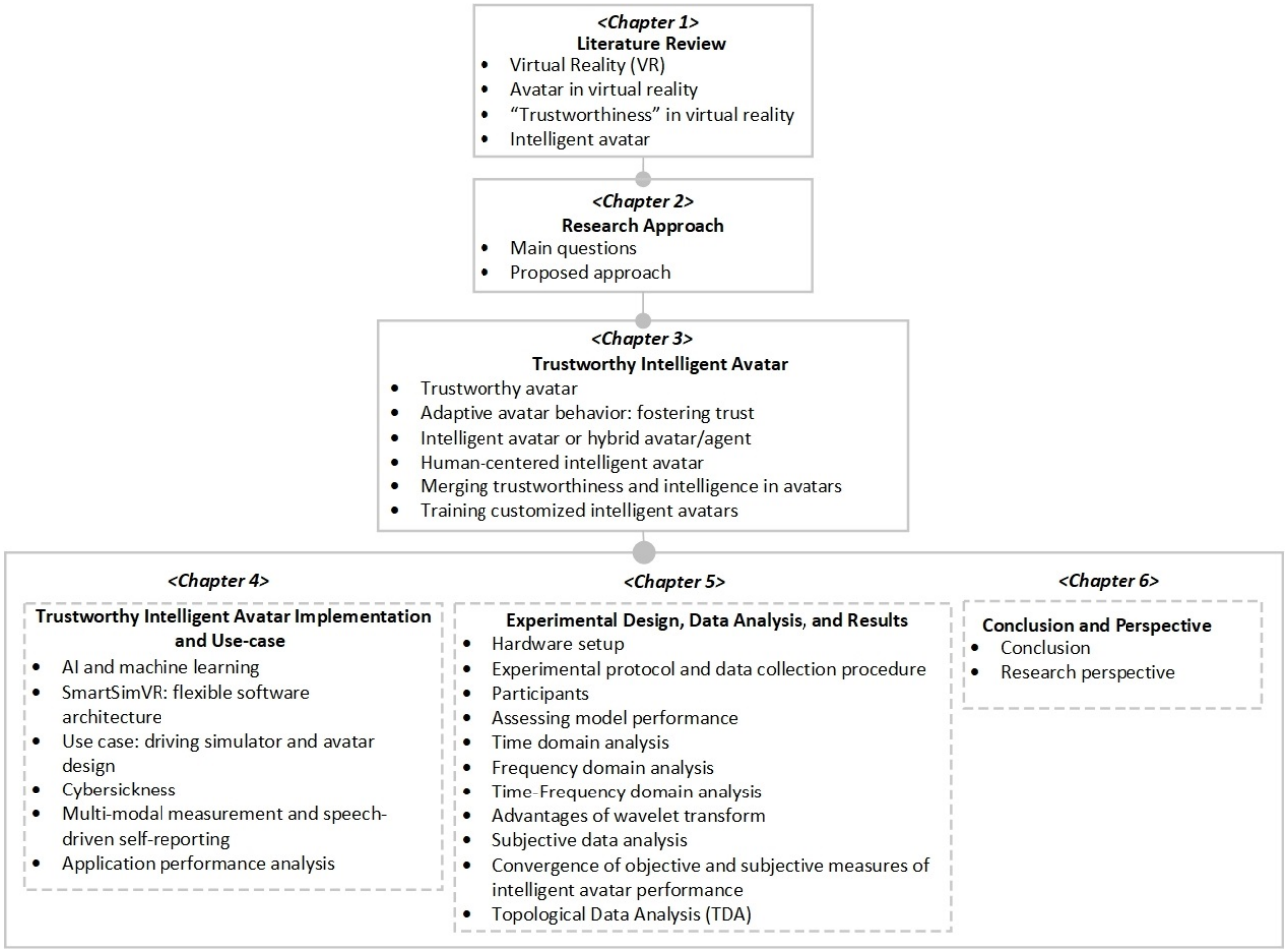


Figure C.1: Aperçu de la structure de la thèse

diverses allant de la reconstruction d'avatars faciaux [256] aux interventions concernant l'image corporelle [174]. Un domaine crucial dans lequel les avatars sont largement utilisés est celui des systèmes collaboratifs. Le rôle des avatars dans la facilitation de la communication et de l'interaction au sein des environnements virtuels est vital pour la collaboration virtuelle [20]. Les avatars permettent aux utilisateurs de se transmettre par des gestes et le langage corporel, qui sont fondamentaux pour les interactions sociales [189] [191]. De plus, ils représentent visuellement les activités des autres utilisateurs, améliorant le sentiment de présence et de conscientisation dans l'environnement virtuel. Les avatars influencent également les capacités des utilisateurs et fournissent des informations sur leur position dans le monde virtuel. L'apparence d'un avatar reflète comment les utilisateurs souhaitent se présenter aux autres et leur auto-perception [240]. Les interactions des utilisateurs et les réactions émotionnelles avec les avatars dans les environnements virtuels ont été largement étudiées. Les travaux

de recherche de Majaranta et Bulling [137], Zhang et al. [261], et Waltemate et al. [242] contribuent collectivement à l'expansion des connaissances concernant la relation entre les utilisateurs et leurs avatars dans les environnements virtuels. Ces études explorent l'attention visuelle, l'expression émotionnelle, la personnalisation et l'immersion, toutes visant à approfondir notre compréhension de cet aspect significatif de l'interaction virtuelle.

L'étude de la relation utilisateur-avatar est une exploration interdisciplinaire qui met en lumière les dynamiques complexes et l'influence profonde que les avatars ont sur la formation des expériences et des interactions des utilisateurs dans les environnements virtuels. Cette exploration englobe diverses dimensions, y compris l'interaction incarnée et la possession du corps, contribuant ainsi à la définition et à la fonctionnalité multifacettes des avatars en RV. La complexité de cette étude interdisciplinaire est vraiment fascinante.

C.2.1 Interactivité des Avatars : Immersion dans les Environnements Virtuels

Dans la vie réelle, nos corps permettent d'effectuer des actions et de recevoir des retours sensoriels. En RV, l'avatar sert d'intermédiaire entre le corps physique et l'environnement virtuel, facilitant l'interaction et la manipulation active du monde virtuel. Il reflète les actions des utilisateurs et fournit des retours visuels et sensoriels.

Les utilisateurs peuvent recevoir des retours auditifs, haptiques et proprioceptifs, renforçant le sentiment de présence [241], [96], [177]. L'interactivité des avatars influence l'immersion et l'expérience utilisateur, augmentant les sentiments d'agence et de possession [213]. Les interactions sociales avec d'autres avatars enrichissent également l'engagement [253].

Cette interactivité a des applications dans des domaines tels que l'éducation, la formation et la thérapie, facilitant l'apprentissage pratique et la réhabilitation. Par exemple, les avatars interactifs dans les simulations médicales offrent un environnement sécurisé pour le développement des compétences [106].

En conclusion, l'interactivité des avatars est cruciale pour l'engagement dans les environnements virtuels, renforçant l'agence et la présence, et menant à des expériences immersives.

C.2.2 Apparence : Avatar Humanoïde ou Non-Humanoïde

L'apparence des avatars en RV est essentielle. Les avatars humanoïdes, qui imitent l'apparence humaine, ont longtemps été privilégiés pour favoriser la familiarité et l'interaction sociale, notamment dans des applications comme les jumeaux numériques et les systèmes de téléprésence [160], [49].

Cependant, les avatars non-humanoïdes gagnent en popularité, permettant aux utilisateurs d'explorer des formes créatives comme des animaux ou des entités fantastiques. Cela favorise l'expression de soi et améliore l'immersion [99], [168]. Ces avatars peuvent également s'adapter aux contextes spécifiques des applications, comme des avatars robotiques dans des jeux de science-fiction [119].

En outre, les avatars non-humanoïdes améliorent l'accessibilité pour les utilisateurs avec des limitations physiques ou cognitives. Le choix entre humanoïdes et non-humanoïdes dépend des objectifs et préférences de l'expérience RV, chaque type offrant des avantages uniques pour l'interaction et l'immersion.

C.2.3 Capacités Fonctionnelles des Avatars dans les Environnements Virtuels

Les avatars en RV possèdent plusieurs capacités qui améliorent l'expérience utilisateur et l'interaction. Ces capacités incluent la représentation visuelle, les animations, la locomotion, la manipulation d'objets et la communication.

Les animations donnent vie aux avatars, permettant des mouvements et des expressions émotionnelles, ce qui améliore le réalisme et la communication non verbale [55]. Par exemple, Akbas et al. [2] ont créé une expérience de danse en RV où les mouvements d'un danseur sont reflétés sur un modèle 3D, facilitant la collaboration.

Les capacités de locomotion permettent aux utilisateurs de naviguer à l'aide de joysticks, de gestes ou d'interfaces cerveau-ordinateur [1]. La manipulation d'objets est essentielle pour interagir avec des éléments virtuels de manière intuitive [12]. Knierim et al. [115] ont montré que la conception des avatars affecte l'interaction utilisateur.

Les avatars permettent également la communication par la voix, le texte et les gestes, favorisant l'interaction sociale [78], [69]. Une étude de Lo [131] a suggéré que l'intégration d'avatars avec des agents conversationnels pourrait enrichir l'éducation en RV.

De plus, les avatars sont bénéfiques dans la réhabilitation sociale des enfants autistes, facilitant

l'engagement et le développement des compétences sociales [200].

En conclusion, les avatars en RV offrent des capacités qui contribuent à une expérience immersive et interactive, favorisant l'engagement et les interactions sociales.

C.2.4 La "Fiabilité" en Réalité Virtuelle

La "fiabilité" en RV désigne le degré auquel les utilisateurs perçoivent les systèmes, applications ou agents de RV comme fiables, crédibles et sécurisés. Cette notion est essentielle, surtout dans des applications impliquant des informations sensibles ou des tâches critiques. La fiabilité comprend plusieurs dimensions : transparence, confidentialité, utilisation éthique, formation des utilisateurs et expérience globale.

Des recherches montrent que la confiance des patients dans leur médecin virtuel peut influencer leur perception de la douleur [9]. Dans le cadre de communautés virtuelles comme Second Life, la fiabilité influence les comportements et interactions [71]. Une étude utilisant une approche multisensorielle a mesuré la confiance envers un agent virtuel, révélant que la puissance alpha des EEG était un indicateur de fiabilité [84].

Plusieurs facteurs contribuent à la fiabilité :

- **Authenticité :** Les utilisateurs font davantage confiance aux avatars qui affichent des comportements authentiques. Mikeska et al. [147] ont montré que la perception de l'authenticité influençait la confiance des enseignants en formation.

- **Fiabilité :** Les avatars fiables répondent de manière cohérente aux entrées des utilisateurs et fournissent des informations précises. Visser et al. [53] ont constaté que la fiabilité des avatars augmente la confiance des utilisateurs.

- **Transparence :** Fournir des informations claires sur le fonctionnement du système RV renforce la confiance des utilisateurs. Zaman et al. [258] soulignent l'importance de la transparence pour établir la confiance dans la technologie XR.

- **Confidentialité :** Protéger les données des utilisateurs contre tout accès non autorisé est crucial pour bâtir la confiance, en particulier dans les applications sensibles.

- **Utilisation Éthique :** Respecter des principes moraux dans la conception des systèmes RV est essentiel pour éviter des pratiques manipulatrices [258].
- **Formation des Utilisateurs :** Offrir une formation adéquate aide les utilisateurs à comprendre et à utiliser efficacement les systèmes RV, augmentant leur confiance en la technologie.
- **Expérience Utilisateur :** Une expérience utilisateur positive, caractérisée par des interfaces intuitives et des interactions réactives, renforce la fiabilité perçue des systèmes RV [254].

C.2.5 Avatar Intelligent

Le terme "Avatar Intelligent (AI)" désigne un avatar doté de capacités avancées d'intelligence artificielle (IA) [72]. Ces capacités permettent à l'avatar d'exhiber un comportement intelligent, de prendre des décisions et d'interagir de manière sophistiquée avec les utilisateurs [128]. Les avatars intelligents jouent un rôle clé à l'intersection de l'IA et de la RV, améliorant l'expérience utilisateur dans les environnements virtuels.

Les systèmes basés sur des règles sont les formes les plus simples d'IA, permettant des ajustements statiques en fonction d'entrées spécifiques [82]. Des cadres d'adaptation comme le 3QS permettent aux environnements virtuels de s'ajuster dynamiquement aux besoins des utilisateurs, réduisant des problèmes comme le cybersickness [237].

Des techniques d'IA plus avancées, telles que les algorithmes d'optimisation et l'apprentissage par renforcement, offrent une adaptabilité accrue. Par exemple, des algorithmes d'optimisation peuvent ajuster les caractéristiques des avatars en fonction des performances des utilisateurs [262]. L'apprentissage par renforcement a été utilisé pour former des personnages virtuels à interagir de manière réaliste avec les participants [102].

L'apprentissage automatique peut également personnaliser les environnements virtuels en fonction des états émotionnels des utilisateurs, comme le montre le travail de Van et al. [238]. De plus, des avatars peuvent être utilisés pour la réhabilitation sensorimotrice, comme indiqué par Bălan et al. [19].

Des recherches récentes, comme celles de Guerrero Vásquez et al. [83], proposent d'animer des modèles 3D pour les transformer en agents virtuels, améliorant ainsi les interactions homme-machine. Chen et al. [41] ont développé un modèle de transfert de mouvement facial à un avatar 3D photoréaliste,

tandis qu'Aquino et al. [51] ont montré comment les avatars peuvent analyser l'environnement et interpréter les intentions des utilisateurs.

Zhang et Tsai [263] ont exploré l'application des techniques d'IA dans l'entraînement sportif, tandis que Hassan et al. [88] ont développé un programme de formation pour les agents de protection de l'enfance utilisant des avatars intelligents. D'autres études se concentrent sur la reconnaissance des mouvements humains et l'amélioration de la réactivité des avatars en utilisant des algorithmes avancés [259].

Enfin, des cadres comme JACALIVE facilitent l'adaptation des environnements virtuels intelligents, tandis que Wang et al. [244] ont proposé une méthode de retargeting facial en temps réel, utilisant l'apprentissage profond pour générer des formes de mélange réalistes.

C.3 Approche de Recherche

La recherche sur l'incarnation, comme le montre le travail de Kilteni et al. [109], révèle que la conception des avatars en RV présente plusieurs lacunes. Parmi celles-ci :

1. **Représentation Inadéquate des Utilisateurs :** Les avatars ne tiennent souvent pas compte des états physiologiques des utilisateurs, comme le rythme cardiaque ou les réponses émotionnelles, limitant ainsi les expériences personnalisées. Nacke et al. [156] soulignent l'importance d'intégrer un retour physiologique pour améliorer l'engagement des utilisateurs. Bailenson et al. [17] insistent sur la nécessité d'incarner des avatars avec des caractéristiques réalistes pour influencer positivement les interactions sociales.

2. **Préoccupations de Sécurité :** La sécurité des utilisateurs est souvent négligée dans la conception des avatars, ce qui peut entraîner des effets indésirables comme le cybermalaise ou des blessures physiques. Bowman et al. [31] recommandent d'intégrer des facteurs de sécurité dans la conception, tels que des mécanismes d'évitement de collision.

Ces lacunes soulignent les défis associés à la conception d'avatars, ouvrant la voie à des expériences virtuelles plus engageantes et sûres.

C.3.1 Questions de Recherche

La recherche sur les avatars s'est souvent concentrée sur leur apparence. Cette thèse vise à répondre aux questions suivantes :

Q_1: *Quels caractéristiques et comportements doivent être intégrés aux avatars virtuels?*

Q_2: *Comment développer un avatar personnalisé pour améliorer l'expérience d'immersion virtuelle?*

C.3.2 Approche Proposée

Notre approche vise à comprendre les états psychologiques et comportementaux des utilisateurs pour construire un modèle personnalisé de compagnon virtuel, intégrant des paramètres objectifs comme l'activité électrodermale et le mouvement oculaire.

Nous avons choisi un simulateur de conduite comme cas d'utilisation, car ces applications offrent des environnements immersifs et réalistes. Cela permet de recueillir des données précieuses sur les réponses physiologiques et psychologiques des utilisateurs, fournissant une base solide pour le développement d'agents intelligents.

Ces simulateurs permettent également une interaction en temps réel, facilitant l'intervention personnalisée pour améliorer le confort des utilisateurs et atténuer le cybersickness. En utilisant des outils d'IA, nous pouvons analyser les données en temps réel pour adapter continuellement notre modèle.

Notre objectif est de créer des jumeaux numériques, des répliques virtuelles qui reflètent les caractéristiques uniques des utilisateurs. En développant ces avatars personnalisés, nous visons à améliorer l'expérience utilisateur en adaptant les interactions et les interventions à leurs besoins spécifiques.

C.4 Avatar Intelligent Fiable

Dans le domaine des avatars, être "fiable" désigne la perception d'un avatar qui inspire confiance et crédibilité chez les utilisateurs. Un avatar fiable est perçu comme dépendable, honnête et capable de remplir sa fonction [136]. Plusieurs facteurs influencent cette fiabilité.

Le premier chapitre (section C.2.4) introduit des facteurs généraux tels que l'authenticité, la fiabilité et la transparence, qui contribuent à la confiance dans divers contextes. Dans les environnements virtuels, la fiabilité peut également dépendre d'autres facteurs, tels que le réalisme visuel, un com-

portement naturel et réactif, des actions prévisibles, des signaux sociaux appropriés, une communication efficace et l'adaptabilité. L'importance de ces facteurs peut varier selon l'application spécifique, l'objectif de recherche ou le domaine d'étude.

C.4.1 Comportement Adaptatif des Avatars : Favoriser la Confiance

L'adaptabilité des avatars est essentielle pour renforcer leur fiabilité, car elle leur permet d'ajuster leur comportement et leurs réponses. Les avatars adaptatifs améliorent considérablement l'expérience utilisateur en RV grâce à des interactions personnalisées et des environnements sur mesure [178][18]. Ils ont diverses applications, notamment l'induction d'états émotionnels, la prédiction de traits individuels et la réhabilitation sensorimotrice [35][16]. De plus, ils jouent un rôle crucial dans la représentation d'utilisateurs distants en réalité mixte (MR), renforçant la présence sociale [172].

Les interfaces adaptatives en RV peuvent également motiver et intensifier l'exercice physique en réagissant aux niveaux d'effort et aux données physiologiques des utilisateurs, améliorant ainsi les performances d'entraînement et de jeu [34]. Les utilisateurs apprécient particulièrement les avatars capables de s'ajuster dynamiquement à leurs besoins, soulignant leur nature centrée sur l'utilisateur [243].

Sun et Botev [227] ont exploré les divers facteurs établissant la confiance entre les utilisateurs et les agents adaptatifs intelligents (IAA), tandis que Kritikos et al. [121] ont présenté un système VR capable de reconnaître les différences individuelles et d'ajuster les scénarios de simulation. Piumsomboon et al. [172] ont introduit Mini-Me, un avatar adaptatif pour la collaboration MR. McIntosh et al. [145] ont proposé une approche systématique pour adapter l'avatar en fonction des performances des utilisateurs.

La question intrigante demeure : comment doter un avatar de cette nature adaptable ?

C.4.2 Avatar Intelligent ou Avatar/Agent Hybride

Dans les études, les termes "avatar" et "agent" sont parfois utilisés de manière interchangeable, mais il est essentiel de les différencier pour définir précisément un avatar intelligent.

Les avatars sont des représentations virtuelles contrôlées par des humains, permettant l'expression de soi et l'interaction dans les environnements virtuels. Ils servent d'interface centrée sur l'utilisateur,

personnalisable selon les préférences. À l'inverse, les agents sont contrôlés par des algorithmes informatiques, exhibant un comportement autonome et des capacités décisionnelles, fonctionnant sans intervention directe de l'utilisateur.

L'avatar intelligent (IA) combine ces deux concepts, intégrant des capacités avancées d'intelligence artificielle, comme l'apprentissage automatique et le traitement du langage naturel. Cela lui permet de démontrer des comportements intelligents et de s'adapter aux interactions avec les utilisateurs, améliorant ainsi l'expérience immersive.

L'intégration de l'intelligence artificielle dans un avatar passe par deux phases : l'apprentissage et le déploiement. L'apprentissage peut être en ligne (durant les interactions) ou hors ligne (après les interactions). Pendant la phase d'apprentissage, l'avatar agit comme une entité virtuelle contrôlée par l'utilisateur. En revanche, lors du déploiement, l'avatar intelligent adopte une nature hybride, combinant des traits d'avatar et d'agent (Figure C.2).

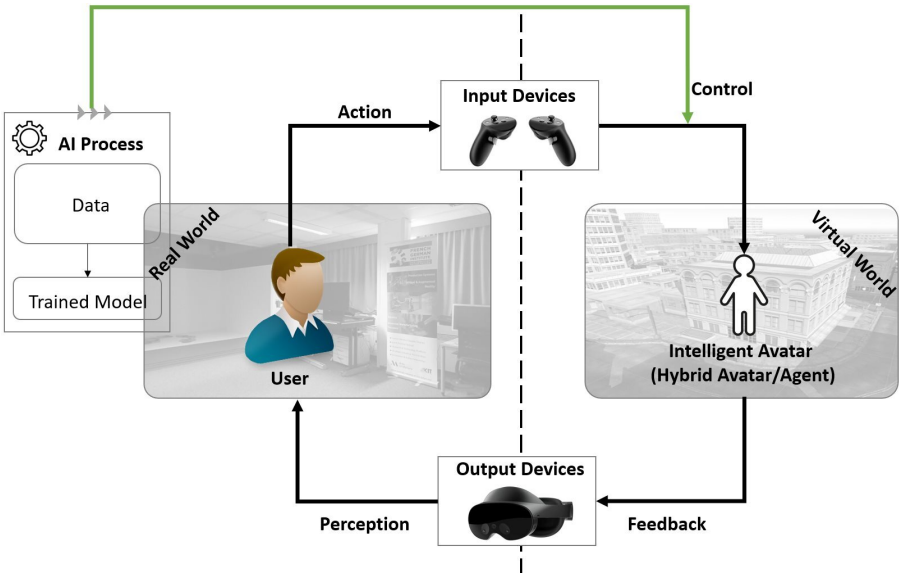


Figure C.2: Caractéristiques hybrides de l'avatar et de l'agent durant la phase de déploiement.

Cette entité hybride représente une présence numérique et utilise des capacités d'intelligence artificielle pour analyser des données et adapter son comportement, offrant ainsi une expérience plus immersive et interactive. Cette fusion vise à combler le fossé entre les utilisateurs humains et la technologie d'intelligence artificielle.

C.4.3 Avatar Intelligent Centré sur l'Humain

En plus de l'"Avatar Intelligent", un autre concept est l'avatar équipé d'une "IA centrée sur l'humain". Ce concept désigne une approche où les systèmes et technologies d'IA utilisés dans l'avatar mettent l'accent sur les besoins, préférences et bien-être des utilisateurs [203]. L'objectif est de créer une expérience personnalisée et centrée sur l'utilisateur.

Bien que les deux concepts utilisent l'IA, l'"IA centrée sur l'humain dans l'Avatar" met en avant cette approche centrée sur l'utilisateur. En adaptant l'avatar aux besoins de l'utilisateur, cette IA vise à établir une confiance, rendant les utilisateurs à l'aise et sécurisés lors des interactions [197]. Shneiderman [202] souligne l'importance de l'IA centrée sur l'humain, en insistant sur la nécessité de systèmes d'IA fiables et dignes de confiance, révélant ainsi l'impact de la fiabilité de l'IA sur l'expérience utilisateur.

C.4.4 Fusion de la Fiabilité et de l'Intelligence dans les Avatars

La personnalisation et l'adaptabilité sont essentielles pour établir la confiance dans les avatars en RV. Les avatars fiables reconnaissent et s'adaptent aux caractéristiques, préférences et besoins des utilisateurs, renforçant ainsi la connexion et la confiance. Grâce à des techniques avancées d'IA, ces avatars peuvent analyser des données utilisateur, telles que les interactions passées et les préférences, pour adapter leurs réponses [18]. En ajustant dynamiquement leur style de communication, leurs gestes, leur interaction avec l'environnement virtuel et même leur apparence, ils créent une expérience plus personnalisée, renforçant ainsi la confiance au fil du temps [92].

En général, la capacité des avatars à reconnaître et s'adapter aux caractéristiques individuelles et aux indices contextuels contribue à une expérience utilisateur immersive et personnalisée, favorisant une connexion plus profonde dans l'environnement virtuel.

C.4.5 Formation d'Avatars Intelligents Personnalisés

Cette section examine la formation d'avatars intelligents personnalisés, en mettant en lumière comment ces avatars s'adaptent en recueillant et en analysant des données utilisateur lors des interactions. Cette approche d'apprentissage en temps réel contraste avec les limitations de l'apprentissage hors ligne, où des ensembles de données statiques ne permettent pas de saisir les nuances individuelles,

entraînant des expériences moins personnalisées. Cependant, des obstacles techniques et pratiques se posent lors de la mise en œuvre d'un entraînement en temps réel efficace. Ces sous-sections soulignent l'importance de l'apprentissage en temps réel pour développer des avatars intelligents véritablement adaptatifs tout en reconnaissant les défis des méthodes traditionnelles.

C.4.5.1 Apprentissage des Avatars Intelligents Personnalisés en Temps Réel

Pour créer un véritable jumeau virtuel personnalisé de l'utilisateur dans un environnement virtuel, il est crucial que l'avatar apprenne de l'utilisateur lors de leurs interactions. En exploitant des techniques d'apprentissage en temps réel, l'avatar peut continuellement recueillir et analyser les caractéristiques physiologiques et comportementales de l'utilisateur ainsi que ses réponses dans l'environnement virtuel. Ce processus d'apprentissage continu permet à l'avatar de s'adapter et de personnaliser son comportement, créant ainsi une représentation virtuelle qui s'aligne étroitement avec les caractéristiques et les désirs de l'utilisateur, entraînant une expérience utilisateur plus immersive et personnalisée.

Avec l'intelligence artificielle et les algorithmes d'apprentissage automatique, l'avatar peut apprendre à interpréter ces signaux et à personnaliser ses interactions en conséquence. Par exemple, si l'utilisateur montre des signes d'anxiété ou de stress, l'avatar peut ajuster son comportement pour fournir des réponses apaisantes et de soutien. En tirant parti de l'entraînement en temps réel, l'avatar peut continuellement affiner sa compréhension des préférences, des états émotionnels et des schémas de communication de l'utilisateur, conduisant à une représentation plus précise et personnalisée.

C.4.5.2 Limitations de l'Apprentissage Hors Ligne des Avatars

L'apprentissage hors ligne des avatars, utilisant des données préenregistrées ou un ensemble de données statique, présente des inconvénients majeurs par rapport à l'apprentissage adaptatif en temps réel. L'une des principales limitations est la difficulté à saisir les caractéristiques nuancées des utilisateurs individuels.

L'avatar est formé à partir de données d'un panel d'utilisateurs, ce qui peut ne pas modéliser adéquatement les préférences spécifiques de chacun. Cela peut entraîner un comportement qui ne s'aligne pas sur les besoins contextuels de l'utilisateur, rendant l'expérience moins engageante.

De plus, l'apprentissage hors ligne manque de réactivité face à l'évolution des comportements et des préférences humaines. Même avec des données préenregistrées d'un utilisateur, l'avatar peut devenir

obsolète, diminuant l'immersion et provoquant de la frustration.

Enfin, l'adaptabilité limitée aux circonstances changeantes est un problème, notamment dans des applications dynamiques comme les simulateurs de conduite, où les modèles hors ligne ne peuvent pas s'auto-adapter aux préférences fluctuantes de l'utilisateur.

C.4.5.3 Défis de l'Apprentissage en Temps Réel

L'apprentissage en temps réel pour les avatars intelligents centrés sur l'humain présente plusieurs défis à relever pour une mise en œuvre efficace. Tout d'abord, l'acquisition et le traitement des données physiologiques et comportementales en temps réel nécessitent des technologies de détection robustes. La capture précise de signaux comme le rythme cardiaque et les expressions faciales est cruciale. Cependant, la synchronisation des capteurs à fréquences différentes en temps réel peut être un défi.

Un autre défi majeur est la complexité computationnelle. Les avatars personnalisés doivent analyser les données utilisateur avec des algorithmes d'apprentissage automatique, ce qui exige des ressources computationnelles importantes pour traiter rapidement les flux de données.

De plus, la protection de la vie privée et la sécurité des données sensibles sont des préoccupations critiques. Il est essentiel de garantir la confidentialité des informations utilisateur par le biais de mesures de sécurité robustes et de conformité aux réglementations.

Enfin, définir des critères d'arrêt pour l'apprentissage de l'avatar est crucial. Sans un point d'arrêt clair, l'avatar peut continuer à apprendre indéfiniment, entraînant un surajustement. Établir des critères appropriés pour déterminer quand l'avatar a atteint un état fiable est essentiel.

Surmonter ces défis nécessite des efforts interdisciplinaires, notamment des avancées dans les technologies de capteurs, les capacités computationnelles et le développement d'algorithmes. Cela permettra des techniques d'apprentissage en temps réel plus sophistiquées, offrant des expériences utilisateur plus riches et immersives.

C.5 Implémentation et Cas d'Utilisation

Nous avons développé un simulateur de conduite pour démontrer l'implémentation pratique de notre approche, servant de cas d'étude convaincant. Dans ce simulateur, l'utilisateur joue le rôle de conducteur, contrôlant entièrement la voiture, qui devient l'avatar, désigné comme un *"avacar"* dans

la littérature [181]. Nous prévoyons d'équiper la voiture d'une intelligence centrée sur l'humain pour interagir de manière réactive avec l'utilisateur.

Une question clé est : comment cet avatar intelligent peut-il améliorer l'expérience utilisateur dans les environnements virtuels ? Nous visons à atténuer le cybermalaise, un défi commun dans les applications de RV, afin de créer une expérience plus confortable. La Figure 4.1 présente un aperçu de l'application.

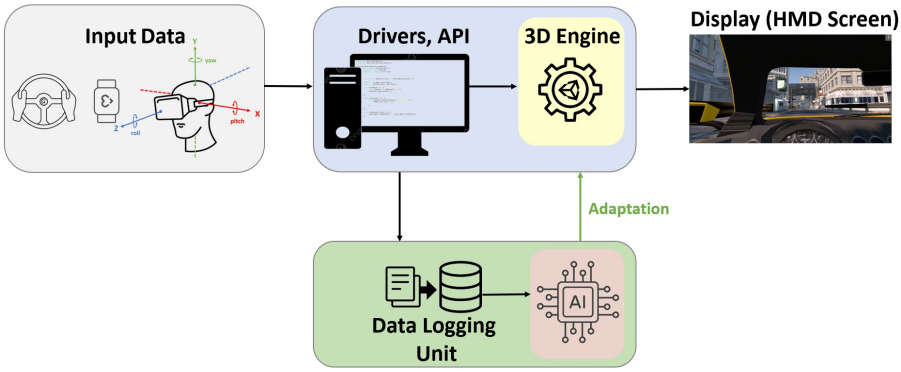


Figure C.3: Aperçu de l'Application.

Cependant, le traitement des données en temps réel et l'auto-adaptation posent des défis, notamment pour les simulateurs de conduite, qui nécessitent des environnements détaillés et des dynamiques de véhicule précises. Bien que des avancées matérielles aient amélioré les performances, l'optimisation des logiciels est cruciale pour une utilisation efficace des ressources disponibles.

Chaque Système de Réalité Virtuelle Adaptatif (AVRS) repose sur trois caractéristiques clés :

- **Mesures de performance:** Évaluent l'efficacité de l'utilisateur à accomplir des tâches, ainsi que des données physiologiques et cinématiques.
- **Logique adaptative:** Utilise des systèmes basés sur des règles et des algorithmes d'IA pour analyser les entrées de l'utilisateur, permettant des interactions dynamiques.
- **Variables adaptatives:** Ajustent les éléments de l'environnement virtuel selon les préférences de l'utilisateur.

Ce chapitre explorera les fondements théoriques du cybermalaise, le développement de cet avatar intelligent, et comment il peut relever les défis liés aux applications de RV, notamment l'atténuation

du cybermalaise. Nous visons ainsi à contribuer à l'avancement de la technologie des avatars et à enrichir l'expérience de RV pour les utilisateurs.

C.5.1 SmartSimVR : Architecture Logicielle Flexible

L'implémentation de l'avatar intelligent auto-adaptatif repose sur l'architecture *"SmartSimVR"*, intégrant des applications de RV avec adaptation en temps réel.

Cette architecture comprend quatre composants fondamentaux :

1. **Distribution :** Modules répartis sur plusieurs systèmes, avec des méthodes de transfert de données comme TCP/IP.
2. **Processus Concurrentiels :** Système multi-threadé pour un traitement efficace des données en temps réel.
3. **Mémoire Virtuelle Partagée :** Synchronisation des données entre processus pour améliorer les performances.
4. **Apprentissage en Flux :** Approche permettant une adaptation continue.

La Figure C.4 montre l'architecture logicielle, flexible et modulaire, apte à intégrer divers capteurs et à s'appliquer à différents scénarios de RV.

C.5.1.1 Détails de l'Implémentation Logicielle

La Figure C.5 présente le flux de données de SmartSimVR, utilisant une connexion TCP/IP pour l'échange entre client et serveur.

Le côté client comprend l'**Application RV**, qui offre une expérience immersive et capture des données comportementales. Le côté serveur englobe plusieurs processus clés :

- **Client TCP de Données :** Reçoit des données de suivi des yeux et mouvements de tête.
- **Processus de Mesure :** Acquiert des données pour le pipeline d'analyse.
- **Module Cerveau :** Comprend :

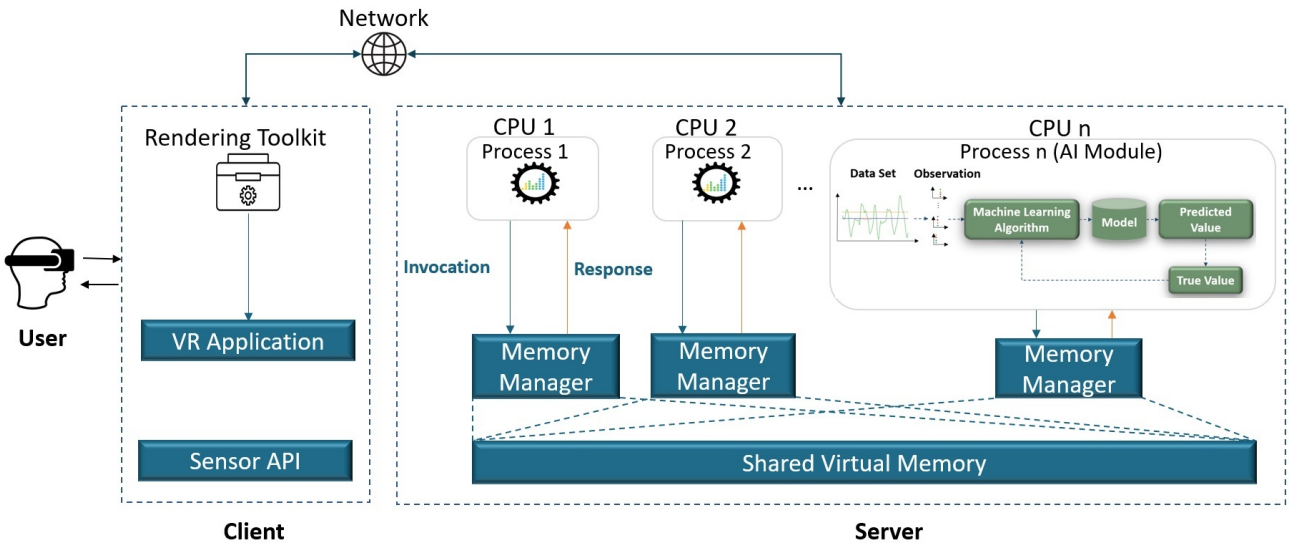


Figure C.4: Aperçu de l'architecture logicielle.

- **Module de Prétraitement :** Nettoie et normalise les données.
- **Module IA :** S'entraîne sur les données en temps réel.
- **Module de Détection/Adaptation :** Prédit des variables cibles pour l'adaptation de l'application RV.

C.5.2 Mesure Multi-Modale et Auto-Evaluation Vocale

L'avatar intelligent auto-adaptatif doit prendre en compte l'état de l'utilisateur en collectant des mesures physiologiques et comportementales liées au cybermalaise, ainsi que des auto-évaluations vocales. Les mesures comportementales, comme le suivi des yeux et des mouvements de tête, fournissent des informations en temps réel sur les réponses de l'utilisateur dans l'environnement virtuel.

C.5.2.1 Mesures Comportementales

Le suivi des yeux mesure des valeurs clés, telles que la vitesse et la position des yeux, tandis que le suivi de la tête évalue l'orientation spatiale. Des mouvements rapides peuvent indiquer des conflits sensoriels associés au cybermalaise. Cette approche utilise un système intégré dans le casque Meta Quest Pro, capturant six degrés de liberté pour des mesures précises.

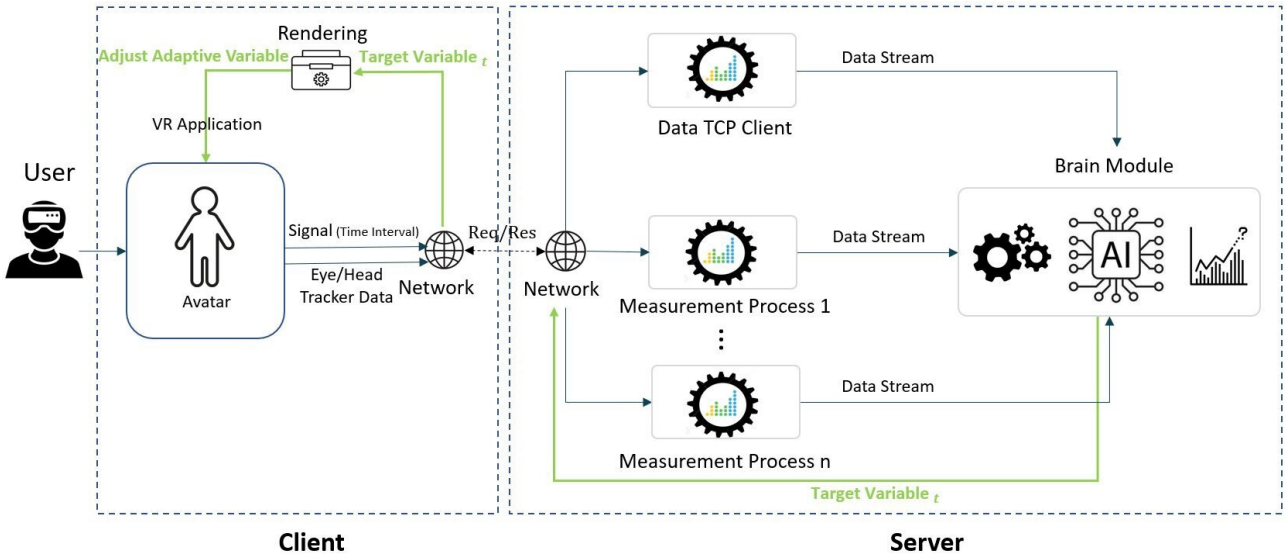


Figure C.5: Diagramme de flux de données de l'architecture SmartSimVR, un système distribué entre le client et le serveur.

C.5.2.2 Mesures Physiologiques Capturées

Les mesures physiologiques, comme l'activité électrodermale (EDA), le pouls sanguin (BVP), l'accéléromètre (ACC) et la température de la peau, sont essentielles pour évaluer et prédire le cybermalaise. Pour la collecte de ces données physiologiques, le dispositif Empatica E4 a été utilisé, fournissant des indicateurs objectifs du confort et de la susceptibilité à cette condition.

C.5.2.3 Indicateur de Classification Alimenté par la Reconnaissance Vocale Fiable

La Reconnaissance Automatique de la Parole (ASR) utilise des algorithmes sophistiqués pour transformer la parole en texte écrit. Dans cette thèse, les voix des participants ont été enregistrées et transcrites via l'ASR pour collecter des mesures adaptatives. Le système a utilisé la Fast Motion Sickness Scale (FMS) pour évaluer le niveau de malaise des participants, qui ont répondu à la question audio, *"Quel est votre score ?"* sur une échelle de 0 à 3. Un module de post-traitement a été développé pour améliorer la précision de la transcription, répétant la question si la réponse n'était pas valide.

C.5.2.4 Prétraitement des Données pour le Simulateur de Conduite RV

La synchronisation des mesures de capteurs hétérogènes est une étape essentielle dans ce système. Les données comportementales, telles que celles des trackers oculaires et de tête, sont transmises à la

simulation via un processus indépendant (Data TCP Client), avec une fréquence de 4 Hz. Pour les données physiologiques, les fréquences de transmission sont :

- EDA : 4 Hz
- BVP : 64 Hz
- TMP : 4 Hz
- ACC : 32 Hz

Un algorithme d'interpolation linéaire a été mis en œuvre pour synchroniser ces données à une fréquence standard de 4 Hz, facilitant ainsi l'intégration et l'analyse des mesures multimodales.

C.5.2.5 Apprentissage en Flux avec River : Adaptation Personnalisée dans le Simulateur de Conduite

Le prétraitement des données, notamment l'ingénierie des caractéristiques et la transformation, est essentiel pour les tâches d'apprentissage en flux. Une technique courante est la standardisation des caractéristiques, réalisée avec le StandardScaler, qui ajuste les données à une moyenne de 0 et une variance unitaire.

Nous avons utilisé l'outil d'apprentissage automatique River, avec la régression logistique comme algorithme de base, optimisée par la descente de gradient stochastique (SGD) avec un taux d'apprentissage de 0,01. La méthode de détection ADWIN a également été intégrée pour gérer les dérives dans les données.

C.6 Conception de l'Expérience, Analyse des Données et Résultats

Dans le cadre de l'expérience sur le cybermalaise, un protocole précis était essentiel. Deux études pilotes ont été menées pour affiner la procédure et identifier les points de blocage, garantissant un protocole final robuste.

L'hypothèse centrale est que l'utilisation d'un avatar intelligent auto-adaptatif dans un simulateur de conduite peut réduire le cybermalaise, améliorant ainsi la sécurité et la confiance des utilisateurs. Pour évaluer cette hypothèse, nous analysons les données en examinant les signaux dans les domaines

temporel, fréquentiel, temps-fréquence et topologique. Une analyse statistique complétera cette approche, quantifiant les tendances observées et les corrélations entre les scores SSQ et les signaux physiologiques.

Cette approche devrait fournir des insights profonds et améliorer la sécurité et le confort des utilisateurs dans les environnements virtuels immersifs.

C.6.1 Procédure de Collecte de Données

L'étude comportait deux sessions de simulation de conduite : l'Avatar Intelligent, qui détectait le cybermalaise en temps réel et adaptait l'environnement, et l'Avatar Non-Intelligent, avec des paramètres statiques. Chaque participant a testé les deux conditions, avec un intervalle de 10 jours pour minimiser les effets de transfert. Les protocoles étaient identiques afin de comparer directement les résultats de cybermalaise.

Avant le début, les participants ont reçu une introduction au système et ont signé un consentement éclairé. Ils ont rempli un questionnaire sur leur état de santé et leur susceptibilité au mal des transports. Pendant l'expérience, ils ont utilisé l'échelle FMS pour rapporter leur niveau de malaise toutes les minutes. La durée de l'expérience variait de 10 à 12 minutes, avec une attention particulière portée à leur bien-être. Si un participant signalait des symptômes, le système adaptait automatiquement les paramètres pour améliorer son confort.

C.6.2 Participants

Une étude pilote avec six participants (5 H, 1 F) a établi le protocole. Pour l'étude principale, 30 participants (15 H, 15 F, moyenne d'âge 23.43 ans) ont été recrutés via des tracts sur les réseaux sociaux et dans la communauté universitaire.

Tous avaient une bonne santé et une vision normale. Parmi eux, 12 ont signalé des épisodes de mal des transports. Pendant la simulation non-intelligente, quatre participants ont ressenti un cybermalaise sévère, entraînant l'arrêt de l'expérience. L'accent a été mis sur ceux ayant complété les deux sessions avec un certain inconfort pour évaluer les changements physiologiques associés au cybermalaise.

C.6.3 Évaluation de la Performance du Modèle Pendant l'Entraînement

L'évaluation du modèle s'est concentrée sur la classification binaire des données physiologiques et comportementales, utilisant les scores FMS comme indicateur. Le modèle a été formé de manière incrémentale avec des données collectées chaque minute, chaque participant fournissant un ensemble de données de dimensions $\mathbb{R}^{240\times25}$.

Les performances ont été mesurées par des métriques telles que l'exactitude et le score F1. L'exactitude médiane était de 93.11 %, indiquant une bonne classification des données. Le score F1 médian était de 89.78 %, reflétant la capacité des modèles à prédire le cybermalaise tout en minimisant les faux positifs et négatifs.

Les résultats montrent des performances solides dans la prédiction du cybermalaise, mais une variabilité dans les scores F1 suggère des différences individuelles qui méritent une exploration supplémentaire pour améliorer les capacités prédictives des modèles.

C.6.4 Analyse du Domaine Temporel

L'analyse du domaine temporel utilise les métriques suivantes pour caractériser les réponses des participants :

1. Moyenne
2. Écart type
3. Racine Carrée de la Moyenne (RMS)

Ces statistiques permettent d'évaluer les niveaux d'activité, la variabilité et l'intensité des signaux physiologiques et comportementaux. Les résultats montrent que l'utilisation d'un avatar intelligent réduit les paramètres oculomoteurs et physiologiques, indiquant une meilleure régulation et une réduction du cybermalaise. Les données révèlent aussi que l'avatar intelligent stabilise les réponses physiologiques, contribuant à une expérience utilisateur plus confortable.

C.6.5 Analyse du Domaine Fréquentiel

L'analyse du domaine fréquentiel utilise les métriques suivantes :

1. Densité Spectrale de Puissance (PSD)

2. Entropie Spectrale

3. Fréquence Moyenne

Ces techniques permettent d'examiner les propriétés spectrales des signaux et d'évaluer les réponses des participants à l'avatar intelligent. Les résultats montrent que l'avatar intelligent réduit la variabilité et l'énergie des mouvements oculaires et des signaux physiologiques, ce qui peut entraîner une attention plus contrôlée et une diminution du cybermalaise. L'analyse de la PSD révèle des mouvements oculaires plus stables, tandis que des valeurs d'entropie plus faibles indiquent moins de désordre dans les réponses physiologiques. Ces découvertes suggèrent que l'interface avec l'avatar intelligent pourrait améliorer l'engagement de l'utilisateur tout en réduisant la charge cognitive.

C.6.6 Analyse du Domaine Temps-Fréquence

L'analyse dans le domaine temps-fréquence utilise les métriques suivantes :

1. Puissance du Spectrogramme

2. Énergie des Wavelet

Ces techniques permettent d'explorer simultanément les caractéristiques temporelles et fréquentielles des signaux, révélant des événements transitoires souvent manqués par les méthodes traditionnelles. Les résultats indiquent que les participants interagissant avec un avatar intelligent présentent des mouvements oculaires et des réponses physiologiques plus stables, suggérant une réduction du cybermalaise. À l'inverse, l'interaction avec un avatar non intelligent est liée à une augmentation de l'excitation physiologique. Ces découvertes soulignent l'importance de concevoir des interfaces virtuelles engageantes.

L'analyse des données montre que la représentation de l'énergie des ondelettes est plus efficace pour détecter les différences entre les signaux physiologiques et comportementaux dans les études avec avatars intelligents et non intelligents. Cette méthode a identifié des différences significatives dans 72% des signaux, surpassant d'autres analyses.

C.6.7 Analyse des Données Subjectives

En plus des mesures physiologiques objectives, l'expérience utilisateur a été évaluée via le Questionnaire de Mal des Transports (SSQ). Nous avons calculé la différence entre les scores SSQ pré-exposition et post-exposition pour indiquer les changements de symptômes de cybermalaise.

Les résultats montrent des différences significatives entre les conditions d'avatar intelligent et non intelligent. Dans la condition d'avatar intelligent, le score de la sous-échelle Nausea a diminué de 43.4%, celui de Oculomotor de 35.1%, et Disorientation de 70.6%. L'ensemble du score SSQ a également diminué de 44.3%.

Ces résultats indiquent que l'avatar intelligent réduit considérablement les symptômes de cybermalaise, améliorant ainsi le confort et l'expérience utilisateur dans les environnements numériques immersifs. Cela souligne le potentiel des avatars intelligents pour faciliter l'adoption des technologies VR/AR.

C.6.8 Convergence des Mesures Objectifs et Subjectives de la Performance de l'Avatar Intelligent

L'analyse des données objectives, en mettant l'accent sur les métriques d'énergie des ondelettes, montre des diminutions significatives des mesures physiologiques et comportementales dans la condition d'avatar intelligent par rapport à l'avatar non intelligent.

Les résultats révèlent des corrélations positives fortes entre la vitesse oculaire (x et y), la vitesse angulaire oculaire (x et y), et la réaction électrodermale (EDA) avec les scores du Questionnaire de Mal des Transports (SSQ). Par exemple, pour la vitesse oculaire sur l'axe x, la corrélation est $r = 0.97$, et pour l'axe y, $r = 0.95$. Les équations de régression pour la vélocité angulaire montrent également des corrélations significatives, avec $r = 0.78$ pour l'axe x et $r = 0.98$ pour l'axe y.

De plus, l'EDA présente une corrélation de $r = 0.80$ avec les scores SSQ. Ces résultats convergents renforcent la validité de ces marqueurs comme indicateurs de la performance des avatars intelligents et de l'expérience de cybermalaise.

C.6.9 Analyse des Données Topologiques (TDA)

L'Analyse des Données Topologiques (TDA) émerge comme une méthode pour extraire des insights significatifs des ensembles de données complexes, utilisant des concepts de topologie appliquée et d'homologie persistante. TDA se concentre sur les caractéristiques topologiques, telles que les composants connexes et les boucles, pour analyser la structure des données au-delà de l'espace euclidien. L'homologie persistante, qui suit l'évolution des caractéristiques topologiques à différents seuils, est fondamentale dans TDA.

Dans notre analyse, nous avons utilisé la valeur d'homologie persistante pour quantifier la durée des caractéristiques topologiques dans les données. Cette valeur est calculée comme suit :

$$\text{persistent value} = d - b \tag{C.1}$$

Les valeurs d'homologie persistante, présentées dans le Tableau B.8, offrent des insights sur la stabilité des signaux physiologiques et comportementaux mesurés dans les études avec avatars intelligents et non intelligents. Des valeurs plus élevées indiquent des signaux plus cohérents, tandis que des valeurs plus basses suggèrent une plus grande variabilité.

Les résultats montrent que les conditions d'avatar intelligent ont des valeurs persistantes significativement plus basses pour la vitesse oculaire et la vitesse angulaire, indiquant des mouvements oculaires plus stables. Les positions et orientations oculaires étaient également plus cohérentes. Les signaux de mouvement de tête et d'accélération ont montré le même schéma, suggérant des mouvements plus naturels.

Les signaux physiologiques, comme l'EDA et le BVP, ont également présenté des valeurs persistantes plus élevées, indiquant une expérience utilisateur plus confortable. Notablement, l'augmentation des valeurs persistantes est la plus marquée pour le groupe d'homologie H_2, suivi de H_1 et H_0, suggérant une réduction hiérarchique de la complexité des dynamiques du signal. Cela indique que l'avatar intelligent stabilise efficacement des relations complexes dans les données.

C.6.10 Discussion

Notre étude souligne les avantages significatifs des avatars intelligents dans les environnements RV, notamment en matière de sécurité des utilisateurs. Les résultats confirment l'importance des

retours physiologiques pour créer des expériences utilisateur engageantes et adaptatives. Par exemple, la diminution de la vitesse oculaire et de l'activité électrodermale en présence d'avatars intelligents indique que ces métriques pourraient servir d'indicateurs précieux des états des utilisateurs, ouvrant la voie à des systèmes de surveillance en temps réel.

Nous mettons également en évidence le potentiel d'un design centré sur l'utilisateur pour améliorer la satisfaction. En créant des compagnons virtuels personnalisés, nous corroborons l'idée que des avatars réalistes peuvent influencer le comportement et les interactions sociales. Les utilisateurs se sentent plus engagés et moins anxieux lorsque les avatars répondent à leurs besoins individuels.

Pour l'avenir, il serait intéressant d'explorer l'application des avatars intelligents dans divers contextes RV, comme l'éducation ou la thérapie. Des études longitudinales pourraient évaluer les effets à long terme de ces avatars sur l'expérience utilisateur. L'analyse de métriques physiologiques supplémentaires, telles que la variabilité de la fréquence cardiaque, pourrait enrichir la compréhension des états des utilisateurs. Enfin, étudier l'impact des avatars intelligents sur les interactions sociales pourrait améliorer la présence sociale et la communication dans les espaces virtuels.

En conclusion, notre recherche contribue à la littérature sur le design d'avatars en RV en montrant l'efficacité des avatars intelligents pour améliorer la sécurité des utilisateurs et réduire le cybermalaise, soulignant la nécessité d'explorer davantage les technologies adaptatives dans ces environnements.

C.7 Conclusion et Perspective

Cette recherche vise à améliorer la technologie des avatars en RV en les rendant plus adaptés aux états des utilisateurs. La contribution principale de cette thèse est l'introduction du concept d'avatar intelligent et fiable, qui s'adapte dynamiquement grâce à des techniques d'IA.

Un cadre architectural, SmartSimVR, a été conçu pour réaliser ce concept, intégrant des fonctionnalités clés telles que la distribuitivité, la simultanéité, la mémoire virtuelle partagée et l'apprentissage en continu. Ce système permet une expérience utilisateur personnalisée et réactive.

Le simulateur de conduite utilisé comme cas d'étude a démontré que l'avatar intelligent réduit le cybermalaise en intégrant un Système de Réalité Virtuelle Adaptatif. Ce système collecte des données physiologiques et comportementales, utilise une logique adaptative pour ajuster le comportement du véhicule et atténuer les symptômes de cybermalaise.

L'analyse des données subjectives et objectives a montré une réduction significative des symptômes en présence d'avatars intelligents, validant ainsi notre approche. Les résultats soulignent également des réponses physiologiques plus stables, indiquant une expérience utilisateur plus confortable. Enfin, l'analyse de la performance a révélé des opportunités d'optimisation pour améliorer l'application RV.

C.7.1 Perspectives de Recherche

D'importantes améliorations des avatars auto-adaptatifs pourraient créer des expériences virtuelles plus immersives et réalistes. Un axe de développement futur consiste à intégrer des techniques avancées de surveillance physiologique pour évaluer plus précisément les états des utilisateurs, en dépassant la dépendance aux retours auto-rapportés, qui présentent des limites.

L'utilisation des signaux EEG comme étiquettes de classification directes pourrait améliorer la fiabilité du système adaptatif. L'EEG est efficace pour détecter des indicateurs neurophysiologiques liés au cybermalaise, permettant des ajustements en temps réel des avatars.

De plus, les avatars pourraient ajuster dynamiquement leurs comportements en fonction des niveaux de compétence, des états émotionnels et des capacités d'attention des utilisateurs. Par exemple, un avatar pourrait offrir plus d'assistance aux utilisateurs novices tout en permettant plus d'autonomie aux utilisateurs expérimentés.

L'intégration de capacités semi-autonomes permettrait également d'améliorer la sécurité et le confort des utilisateurs. Les avatars pourraient alerter ou suggérer des actions pour aider les utilisateurs à naviguer dans des situations complexes. En équilibrant l'autonomie de l'utilisateur avec des interventions intelligentes, les avatars auto-adaptatifs pourraient devenir des compagnons efficaces dans divers contextes virtuels, allant de l'éducation à la thérapie.

Enfin, l'intégration d'algorithmes décisionnels avancés pourrait permettre aux avatars d'anticiper des scénarios potentiels, d'évaluer des risques et d'optimiser l'expérience utilisateur, rendant l'environnement virtuel plus sûr et personnalisé.

Azadeh HADADI

Development of trustworthy intelligent avatars in virtual immersion

Résumé : La réalité virtuelle est de plus en plus utilisée dans des secteurs tels que l'industrie, l'éducation et la santé, mais des défis comme le cybermalaise entravent son acceptation. Cette thèse présente un avatar intelligent et fiable conçu pour améliorer l'immersion virtuelle en tenant compte des états physiologiques et des réponses émotionnelles des utilisateurs. Nous introduisons "SmartSimVR", un cadre flexible utilisant l'intelligence artificielle et la fusion de capteurs pour créer des avatars auto-adaptatifs. Mis en œuvre dans un simulateur de conduite où la voiture agit comme un avatar non humanoïde, SmartSimVR vise à minimiser le cybermalaise en apprenant continuellement des données utilisateur et en ajustant son comportement en conséquence. Nos expériences ont démontré que l'avatar intelligent réduisait significativement les symptômes de cybermalaise, avec des effets positifs observés dans les mesures physiologiques, y compris la vitesse oculaire et l'activité électrodermale. L'analyse de corrélation a également validé l'efficacité de l'avatar intelligent, fournissant des preuves solides qu'il améliore l'expérience utilisateur en RV en atténuant l'inconfort.

Mots clés : Avatar, Intelligence Artificielle, Immersion Virtuelle, Fusion de capteurs.

Abstract : Virtual reality is increasingly utilized in sectors like industry, education, and healthcare, but challenges such as cybersickness hinder its acceptance. This thesis presents a trustworthy, intelligent avatar designed to enhance virtual immersion by considering users' physiological states and emotional responses. We introduce "SmartSimVR," a flexible framework utilizing artificial intelligence and sensor fusion to create auto-adaptive avatars. Implemented in a driving simulator where the car acts as a non-humanoid avatar, SmartSimVR aims to minimize cybersickness by learning continuously from user data and adjusting its behavior accordingly. Our experiments demonstrated that the intelligent avatar significantly reduced symptoms of cybersickness, with positive effects observed in physiological measures, including eye velocity and electrodermal activity. Correlation analysis further validated the effectiveness of the intelligent avatar, providing strong evidence that it enhances user experience in VR by mitigating discomfort.

Keywords : Avatar, Artificial Intelligence, Virtual immersion, Sensor fusion.