TÉCNICO LISBOA

Universidade de Lisboa

Instituto Superior Técnico

Travel Fidelity in Virtual Environments

Daniel Pires de Sá Medeiros

Supervisor: Doctor Joaquim Armando Pires Jorge
Co-Supervisor: Doctor Alberto Barbosa Raposo

Thesis specifically prepared to obtain the PhD Degree in
Information Systems and Computer Engineering

Draft

October 2018

Universidade de Lisboa

Instituto Superior Técnico

Travel Fidelity in Virtual Environments

Daniel Pires de Sá Medeiros

Supervisor: Doctor Joaquim Armando Pires Jorge

Co-Supervisor: Doctor Alberto Barbosa Raposo

Thesis specifically prepared to obtain the PhD Degree in Information Systems and Computer Engineering

Draft

October 2018

Resumo

A locomoção em ambientes virtuais é actualmente uma tarefa difícil e não natural para ser executada. Normalmente, os pesquisadores tendem a elaborar metáforas baseadas no caminhar no solo para restringir graus de liberdade (GdL) durante o movimento. Essas restrições permitem interações mais próximas do modo como as pessoas andam na vida real possibilitando uma alta fidelidade de interação. No entanto, voar permite que elas alcancem pontos específicos numa cena virtual com mais rapidez. A nossa experiência sugere que técnicas de alta fidelidade podem melhorar a experiência de voo, mesmo que esta não seja inata para humanos, o que requer o controle simultâneo de múltiplos GdL. Por outro lado, a utilização de uma representação do utilizador pode também aumentar a eficiência de navegação. Aspectos como fidelidade gráfica da representação, ou realismo gráfico e perspectiva na qual a representação, ou avatar, é vista. De forma a investigar fidelidade de interação, nós contribuímos com o Magic Carpet, um espaço de desenho que usa um *proxy* de chão com uma representação de corpo inteiro, para evitar problemas de perca de equilíbrio e náusea cibernética. O nosso espaço de *design* permite abordar a viagem com fidelidade (ou seja, a proximidade do mapeamento virtual à sua contrapartida real), tanto nas partes de fidelidade de representação e interação. Este desenho permite a separação dos GdL, abordando as direcções de direção e controlo de velocidade, especificadas separadamente, facilitando assim o uso de técnicas com maior fidelidade de interação. Para abordar a parte de fidelidade de representação do espaço de desenho, propusemos um primeiro estudo para escolher a melhor representação adequada, variando tanto o nível de fidelidade da perspectiva (como o corpo virtual do usuário é visto, seja de uma perspective em terceira pessoa ou em primeira pessoa) e fidelidade gráfica (a proximidade da representação virtual de sua contrapartida real). Argumentamos que o uso de um avatar em primeira pessoa ainda é o mais adequado para locomoção em realidade virtual (RV) e não é afetado pelo realismo gráfico da representação, enquanto a utilização de avatares em terceira pessoa é altamente influenciada por este fator, principalmente pelo facto que nesta perspectiva, o avatar é sempre visto enquanto o utilizador interage com o ambiente virtual. Isso é particularmente perceptível quando as características do usuário do mundo real (por exemplo, roupas, cabelos) são mapeadas em sua representação

dentro do ambiente virtual.

Para validar o nosso espaço de desenho, propusémos dois estudos complementares, um para cada uma das fases da viagem. Na nossa avaliação experimental, apresentamos os resultados de ambos os estudos e indicamos as técnicas mais adequadas para serem usadas em conjunto dentro do espaço de *design* do Magic Carpet. Aplicámos medidas objetivas e subjectivas para avaliar a eficiência, o nível de presença e os efeitos colaterais das técnicas testadas em nosso espaço de projeto, tais como fadiga física e náusea cibernética. Os nossos resultados mostram que este espaço de *design* é viável para propor novas técnicas com um alto nível de interação fidelidade para voar. Embora o uso de alta interação de fidelidade não resulte sempre nas técnicas mais eficientes, os métodos que explorámos permitem um controle de velocidade mais preciso. Para demonstrar a flexibilidade da nossa abordagem, realizámos uma avaliação adicional com três técnicas passivas baseadas em Alvos para a avaliação de velocidade e transições nesse tipo de técnica. Os resultados dessa avaliação mostraram que as técnicas que fornecem tradução imediata têm melhor desempenho e que o uso de transições de fase aparentemente não melhora os factores de qualidade de viagem.

Abstract

Locomotion in Virtual Environments is currently a difficult and unnatural task to perform. Normally, researchers tend to devise ground-floor based metaphors, to constrain Degrees of Freedom (DOF) during motion. These restrictions enable interactions closer to the way people walk in real life to provide high interaction fidelity. However, flying allows people to reach specific points in a virtual scene more expeditiously. Our experience suggests that high-fidelity techniques may also improve the flying experience, even though flying is not innate to humans, which requires the simultaneous control of multiple DOF. We contribute the Magic Carpet, a family of methods that combines a floor-proxy with a full-body representation, to avoid imbalance and cybersickness issues. Our design space allows to approach travel with fidelity (i.e. the closeness the virtual mapping is to its real counterpart), both on the representation and interaction fidelity parts. We approach interaction fidelity in flying scenarios inside our design space by separating DOFs and addressing the indication direction and speed control travel phases separately. thereby promoting techniques with higher interaction fidelity. To address the representation fidelity part of the spectrum, we proposed a first study to choose the best representation suited, varying both the level of the perspective fidelity (how the user's virtual body is viewed, either from a Third-Person Perspective (3PP) or a First-Person perspective (1PP)) and graphical fidelity (the closer the virtual representation is to its real counterpart). We argue that the use of a 1PP is still the best suited perspective for travel in Virtual Reality (VR) and is not affected by the graphical realism of the representation, while the 3PP is highly influenced by this factor. This is particularly noticeable when user characteristics of the real world (e.g. clothing, hair) are mapped into their representation within the virtual environment (VE).

To validate our design space, we proposed two complementary studies, one for each travel phase. In our experimental evaluation, we present the results of both studies and identify the best suited techniques to be used in combination in the Magic Carpet approach. We applied both objective and subjective measurements to evaluate efficiency, level of presence, and side-effects of the tested techniques inside our design space, such as physical fatigue and cybersickness. Our results show that the Magic

Carpet family of methods supports novel techniques with a high degree of interaction fidelity for flying. While high interaction fidelity techniques are seemingly not the most efficient, the methods we explored allow a more precise speed control than other flying techniques. To prove the flexibility of our design space we conducted an additional evaluation with three Target-based techniques to assess both speed and transitions afforded by these techniques. Experimental Results show that methods that provide immediate user translation perform best and that phase transitions do not improve travel quality factors.

Palavras Chave

Navegação
Realidade Virtual
Interação Humano-Computador
Representação Incorporada
Ambientes Virtuais Imersivos

Keywords

Travel
Virtual Reality
Human-Computer Interaction
Fully Embodied Representation
Immersive Virtual Environments

Acknowledgements

First of all I would like to thank my family that even far (in distance) helped me through all my academic life. To all my colleagues from the VIMMI Group at INESC-ID, Daniel Lopes, João Guerreiro, Sandra Gama, Soraia Paulo, Daniel Gonçalves, Hugo Nicolau, João Pereira and Alfredo Ferreira; and most importantly to the MIG team: Maurício Sousa, Daniel Mendes and Rafael Kuffner dos Anjos. To my advisors Joaquim Jorge and Alberto Raposo for the helpful support during all this time. To CAPES Foundation-Brazil and Fundação para Ciência e Tecnologia (FCT) for the financial support through the projects IT-MEDEX (PTDC/EEISII/6038/2014), TECTON-3D (PTDC/EEISII/3154/2012) and CEDAR (PTDC/EIA-EIA/116070/ 2009), without which this work could not be possible. Lastly, I would like to thank all my new and old friends for all the good memories during this period.

“Run, live to fly, fly to live, do or die
Run, live to fly, fly to live, Aces high”
– Aces High, Iron Maiden

.

Contents

List of Figures

List of Tables

Glossary

1PP First-Person perspective. iii, xiii, xiv, 4, 14–17, 19, 20, 34, 42, 44, 49–53, 56–64, 91, 92, 97, 99, 101

3DUI Three-dimensional User Interface. 1, 3, 5, 7, 15, 18

3PP Third-Person Perspective. iii, xiii, xiv, 4, 14, 16–20, 34, 44, 45, 49–53, 55, 57–61, 63, 64, 91, 92, 97, 99, 101

6DoF Six Degrees of Freedom. xiv, 13, 22, 23, 29–33, 37, 92

AR Augmented Reality. 7, 12, 15

AT Animated teleport box. 88, 89

CAVE Cave Automatic Virtual Environment. 1, 31

DOF Degrees of Freedom. iii, 2, 3, 9, 29, 32, 37, 38, 65, 66, 68, 70

FOV Field of View. 4, 15, 18, 26

FPS Frames per second. 68

HCI Human-computer Interaction. 5, 7

HMD Head-Mounted display. 1, 2, 12, 15, 17, 21, 34, 42, 52, 64, 67, 68, 72, 76, 88, 91, 95, 96, 101

IPD Inter-pupillary distance. 15

IVE Immersive Virtual Environment. 4, 6, 8, 11, 13, 14, 19, 28, 44, 85, 86, 100

LM Linear Motion. 88, 89

LoS Level of Scale. 27

MSVE Multiscale Virtual Environment. 27, 28, 101

POI Point Of Interest. 24

RHI Rubber-Hand Illusion. 17

SSQ Simulator Sickness Questionnaire. 26, 72, 76

TP Teleport. 88, 89

UI User Interface. 1

VE Virtual Environment. iii, xiii, 1–4, 6, 8, 11–13, 16, 21–23, 25–30, 32, 33, 35, 37, 41, 45, 46, 49, 52, 58, 63, 68, 72, 73, 76, 81, 85, 86, 88, 91, 92, 95–97, 101

VR Virtual Reality. iii, 1–7, 11–13, 15–19, 21, 25, 26, 28, 32–35, 37, 42, 44, 49, 68, 72, 83, 85, 86, 88, 93, 95, 98–100

WIM World In Miniature. 21

WIMP Windows, Icons, Menu and Pointer. 21

WIP Walking In Place. 3, 26, 33, 38, 66, 77, 79–83, 92, 93

1 Introduction

Virtual Reality (VR) has recently gained traction with many new and ever more affordable devices being released. This increased popularity has originated new applications and has attracted casual consumers to experience VR. Virtual reality setups are classified in how immersed users feel inside the Virtual Environment, from non-immersive setups (monitors+keyboards) to fully-immersive setups (CAVEs and HMDs). As in conventional setups, the User Interface (UI) is an important part of the user experience and determine the success of the system being developed. When an User Interface is used to interact with a 3d environment, this is called a Three-dimensional User Interface (3DUI) and can be divided according to which action is being performed. Classic tasks in this area are of navigation, object selection, manipulation, symbolic input and system control [15].

1.1. Motivation

As VR gear has become more widely available, effective navigation inside Virtual Environment (VE)s has also risen in importance. Navigation consists in the action of moving (also known as travel or locomotion) and the cognitive process of planning where to go and how to go, also referred as wayfinding [50]. In their classic survey, Bowman et al. [13] classified travel techniques on the basis of two main criteria: 1) whether movement is controlled virtually or physically, and 2) whether the action of motion is controlled actively or passively. Quality factors of locomotion include the appropriate control of speed, accuracy, spatial awareness, ease of learning, information gathering (i.e. the user's ability to actively obtain information from the environment during travel) and presence [13]. The use of immersive setups also enable the use of techniques to interact with the Virtual Environment that resemble the way people perform these tasks in real life [18, 26, 62]. Travel techniques can be classified by how the movement is done, either physically (where people use their bodies to control movement) or virtually (using an external device) [13]; and who controls the movement, the user (in active techniques) or the system (in passive techniques). A more recent study relies on task decomposition to classify travel into three phases: direction indication, velocity specification and input conditions. These phases specify how the movement is started, continued and terminated [50].

A common way to design a travel technique is by proposing ways of interacting with the environment with fidelity. Fidelity is known by "the objective degree of exactness with which real-world sensory stimuli are reproduced" [35]. This concept can be extended to users actions while interacting with a VE, which is known as interaction fidelity [62] and may influence the effectiveness of the technique used. In travel tasks, the higher level of interaction fidelity a technique has, the closer it is to the way people walk in real life. However, restrictions of the physical space make this approach not as suited for VR. In complex environments, for example, travel targets may reside out of reach, e.g., above ground, or in remote spots of the VE [61].

Flying provides more efficient means of locomotion in unconstrained large VEs. However, limited work has been done on flying in VR as compared to other locomotion approaches. One possible explanation for this lack is that flying is unnatural to humans and difficult to control using state-of-the-art techniques since it requires simultaneously controlling many Degrees of Freedom (DOF) for translation and speed control. It is unclear whether travel remains a simple task when additional DOFs are added.

The use of Head-Mounted displays also poses the problem that this type of equipment completely occludes users' selves. In this case, a partial or complete representation of the user may be provided to better locate the user inside the VE. This can also help users to establish a higher sense of connectedness between them and the environment, helping people to establish a correct scale relation between them and the VE [46]. An important aspect to consider to evaluate the effectiveness of the representation is the sense of embodiment, which is defined by Kilteni et al. [41] as the "sense that emerges when a virtual body's properties are processed as if they were the properties of one's own biological body". When using a high fidelity technique the use of an avatar is also important as they need instant feedback [114]. The fidelity of the representation used can also be determined by how close the representation is to the real user (graphical fidelity) and how close the point of view on which the avatar is seen is to real-life conditions (perspective fidelity). Since each type of tasks need different kind of feedback, we argue that each task inside the 3DUI area is affected differently by the representation provided. We focus on travel tasks, where we claim that a more realistic representation (with a higher level of representation fidelity), is a better asset for locomotion, more prominently on the perspective aspect.For this, we propose "Magic Carpet", a design space to enable high interaction fidelity metaphors for flying in virtual scenes. This design space addresses the interaction fidelity by isolating the unnatural part of flying, the direction indication phase, where the user needs to control multiple DOFs and enable the use of techniques with higher interaction fidelity to control speed of movement, such as the Walking In Place (WIP) [18]. Inside this design space, we define travel fidelity in fully-embodied VR being composed by the interaction and representation parts. In this thesis, we aim to study how fidelity in the representation and interaction factors in flying tasks affects travel quality factors.

We present an additional study to validate our design space, by using passive, low-fidelity techniques for speed control. In these techniques, called Target-based techniques, the translation is made by the system, which can be immediate (Infinite Velocity, or Teleport) or gradual.To select the best performing speed control technique for this family of techniques, We used the best performing technique for direction control, the Hand technique, in conjunction with three Target-Based techniques for speed control. These include a novel technique called Animated Teleport Box, which incorporates elements from the Teleport technique, which translate people immediately to the desired goal, but includes a transition between initial and end positions.

1.2. Thesis Statement

Locomotion (or Travel) is an important part of the experience when interacting with a VE. This is even more true in Virtual Reality, where people can use forms of travel that are closer to the way people walk in real life. If well designed, these type of techniques can improve the connectedness between the user and the Virtual Environment and impact quality factors of the travel task. The use of techniques with a high level of interaction fidelity can be also important when using supernatural metaphors, which enhance human abilities, such as flying. However, traditional metaphors used for travel alone, are not sufficient to support flying. For this, we propose to isolate the unnatural part of flying which is the control of additional degrees of freedom and treat the two parts of locomotion separately. This separation enables the use of techniques with a high level of interaction fidelity in both direction indication and speed control phases. Also, the use of techniques with high interaction fidelity need the use of fully-embodied representations, since it requires immediate body feedback. Currently, the majority of approaches to self-representation in Immersive Virtual Environments is done by using avatars on a First-Person perspective. Also, some are the works that relates graphical realism to sense of embodiment when an user is viewed in a 1PP. A further development in this matter is the use of realistic avatars that reconstruct the real body of the user and maps into his virtual self.

> **The use of high interaction fidelity techniques can enhance quality factors of travel in a flying supernatural scenario. Moreover, the use of a representation with a high level of fidelity can also improve users' feeling of embodiment in travel tasks, specially on the perspective factor.**

1.3. Results

User representation in Travel Tasks We found that the use of a 3PP avatar, in general, does not lead to improvement in quality factors.Moreover, we recommend the use of avatars in the 3PP only in the case where obstacles reside outside of people's Field of View.We also found that, differently from 1PP which are not affected by its graphical fidelity, 3PP avatars are more affected by this aspect. In summary, we recommend the use of 1PP avatars for travel tasks, regardless of the graphical fidelity of the avatar. However, if the task require an expansion of the FOV, we recommend a 3PP avatar that reconstructs people real-selves into its virtual counterpart.

Flying in VR using high-interaction fidelity metaphors With the exploration of our Magic Carpet Design space, we were able to subdivide the unnatural part of flying, which is the synchronous control of six-degrees of freedom and address each part of the travel pipeline separately, namely direction control and speed control and propose high-interaction fidelity techniques for both phases. We validated this design space with two separate studies for each of these phases. For direction indication, we can infer that the best suited technique is the one that direction control and camera control are separate, so which the user can inspect the scene while travelling. This decoupling also improves the awareness over people's bodies, leading to a more effective way of avoiding obstacles while navigating. On the speed control phase on the other hand, we noticed that the low-fidelity techniques are the most efficient but a high-interaction technique can give a more precise control over speed of movement in flying tasks.

Effects of speed and transitions in Target-based Techniques We can claim that on this type of techniques where the system abruptly moves the person's position (also called as Infinite Velocity) they experience less discomfort but have less spatial awareness than on the Linear Motion technique. We can also assert that the usage of a transition on Infinite Velocity techniques does not affect neither performance nor cybersickness.

1.4. Contributions

The studies presented on this thesis led to the following contributions on the fields of Human-computer Interaction, Three-dimensional User Interface and Virtual Reality:

- Assessment of the impact of the level of representation fidelity and its subfactors graphical fidelity and representation fidelity in Travel tasks quality factors. On the graphical fidelity factor, we used three different representations and varied the level of graphical fidelity of these for assessing how this impacted the quality factors of the Travel technique. Regarding perspective fidelity, we utilized both first and third-person perspective fully-embodied avatars in To isolate the representation fidelity factor we chose to use a travel technique with high level of interaction fidelity : the real walking metaphor. The tasks consisted on walking while avoiding obstacles situated around the user and metrics such as time, collision time and questionnaires were used to quantify the influence of the graphical and perspective fidelity on fully-embodied avatars.

- Conception of the "Magic Carpet" design space that allows the use of techniques with high interaction fidelity in the direction indication and speed control phases for flying in Immersive Virtual Environments. This design space consisted in a floor-proxy which remains at the same position of the real floor, perpendicular to users' bodies, avoiding cybersickness and balance issues. By using this metaphor, users maintain a comfortable position to fly, while leaving the hands-free for performing further actions such as With this, we proposed two different studies that encompass both phases of travel, the direction indication and speed control phases. On each of them, we used three different techniques, from a lower to a high level interaction fidelity.

- Proposal of two novel techniques for flying in fully-embodied Immersive Virtual Environment (IVE)s. We proposed one technique for each of the phases of the flying pipeline : the Elevator+Steering for direction indication and the Speed Circle, for speed control. The Elevator+Steering technique lies on the lower part of interaction fidelity within the tested techniques for direction specification and consists of separating the control of movement in the horizontal plane by using the head, while additional buttons are used to control movement on the vertical axis. The speed circle, on the other hand, uses people's bodies in an analog way as a joystick, where This technique lies on the middle of the interaction fidelity with the other techniques, since as opposed to the Walking In Place technique, users may remain physically stationary while performing locomotion on the VE, differently from the Walking In Place metaphor [18].

- Study to investigate effects of speed and transitions on target-based travel by comparing three different techniques and how it impacts the VR experience in key aspects such as comfort and cybersickness. We compare the already established Linear Motion and Teleport techniques, against a novel technique called Animated-Teleport Box, which improves on the Teleport Technique by adding a gradual transition effect after and before translation, where a box surrounds the user, as similar to an elevator.

1.5. Publications

In this section we list the publications published during the course of my dissertation. The publications which are directly related to this thesis are :

Published:

1. Creepy Tracker Toolkit for Context-aware Interfaces.
   M. Sousa, D. Mendes, R.K. dos Anjos, D. Medeiros, A. Ferreira, A. Raposo, J.M. Pereira, and J. Jorge. In: Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces (ISS), 2017 (Core A Full paper).

2. Evaluation of Travel Techniques for Virtual Reality.
   E. Cordeiro, D. Medeiros, D. Mendes, M. Sousa, A. Raposo, A. Ferreira and J. Jorge In: Proceedings of the Encontro Português de Computação Gráfica(EPCG), 2016 (Full paper).

3. Effects of Speed and Transitions on Target-based Travel Techniques.
   D. Medeiros, E. Cordeiro, D. Mendes, M. Sousa, A. Raposo, A. Ferreira and J. Jorge In: Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST), 2016 (Core A Poster).

4. Keep my head on my shoulders!: Why third-person is bad for navigation in VR.
   D. Medeiros, R. dos Anjos, D. Mendes, A. Raposo, J. Jorge In: Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST), 2018 (Core A Full paper).

5. Magic Carpet: Interaction Fidelity for Flying in VR.
   D. Medeiros, M. Sousa, A. Raposo, J. Jorge. Submitted to the IEEE Transactions on Visualization and Computer Graphics Journal, 2019 (IEEE TVCG) (Q1 Journal Paper).

During the course of this dissertation we published papers that although not directly related to the thesis topic, helped on pushing the boundary on the HCI, 3DUI, AR and VR areas. These works were also important on the concept of the work presented on this thesis. Those are:

1. Design and evaluation of novel out-of-reach selection techniques for VR using iterative refinement.
   D. Mendes, D. Medeiros, M. Sousa, E. Cordeiro, A. Ferreira and J. Jorge Elsevier Computers & Graphics Journal, 2017 (Q2 Journal Paper).

2. PRECIOUS! Out-of-reach selection using iterative refinement in VR.
   D. Mendes, D. Medeiros, E. Cordeiro, M. Sousa, A. Ferreira and J. Jorge. In Proceedings of the 2017 IEEE International Conference on 3D User Interfaces (3DUI), 2017 (Core B Poster).

3. Mid-Air Modeling with Boolean Operations in VR.
   D. Mendes, D. Medeiros, M. Sousa, R. Ferreira, A. Raposo, A. Ferreira, J. Jorge In: Proceeding of the 2017 ACM International Conference on 3D User Interfaces (3DUI). 2017 (Core B Short paper).

4. Perceiving Depth: Optical versus Video See-through.
   D. Medeiros, M. Sousa, D. Mendes, A. Raposo and J. Jorge In: Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST), 2016 (Core A Short paper).

5. SleeveAR: Augmented Reality for Rehabilitation using Realtime Feedback.
   M. Sousa, J. Vieira, D. Medeiros, A. Arsénio and J. Jorge In: Proceedings of the ACM Intelligent User Interfaces (IUI), 2016 (Core A Full paper).

6. Interaction Techniques for Immersive CT Colonography: A Professional Assessment.
   D. S. Lopes, D. Medeiros, S. Paulo, P. Borges, V. Nunes, V. Mascarenhas, M. Veiga and J. Jorge In: Proceedings of the 21st International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'2018), Granada- Spain,2018 (Core A Full paper).

7. Usability studies on building early stage architectural models in virtual reality.
   R de Klerk, A. M. Duarte, D. Medeiros, J. P. Duarte, J. Jorge, D. S. Lopes In : Automation in Construction 103, 104-116. (Q1 Journal Paper)

1.6. Dissertation Outline

We will start with a general overview of embodiment, used classification and present a discussion about the factors that impact the sense of embodiment and presence of users inside a VE. A brief discussion is also presented relating the classification and open problems regarding these factors. Following, a description and relevant work regarding travel in immersive virtual environments. This will include a discussion with the relation between interaction fidelity of the travel technique and travel quality factors which include efficacy, spatial awareness and ease of learning. After, a discussion will be presented interconnecting the two main research topics of this thesis: representation fidelity and travel fidelity in IVEs. Following, we include a brief description of exploratory work made during this thesis and their contributions to the main research topics of this thesis.

In the third chapter, we will present a study about representation fidelity and its sub-components, graphical and perspective fidelity, and its relation to travel tasks. To isolate the interaction fidelity of the study, we chose to use the technique with the higher level of interaction fidelity : the real walking. We varied the level of graphical fidelity in this test, ranging from a humanoid abstract avatar, then a human mesh avatar to a user's real-time point-cloud representation.

Then, on the fourth chapter we present the study on interaction fidelity for flying tasks. For using high-interaction fidelity techniques we explore our proposed "Magic Carpet" design space.This design space enables the use of high-interaction fidelity techniques for both phases of travel : direction specification and speed control. The use of this space enables to isolate the unnatural part of flying, the control of multiple DOFs to indicate direction, which enable the use of high-interaction fidelity techniques on both phases. Each of them is studied separately, with the first study focusing on the part of the direction indication and the second, the phase of speed control. For each phase we evaluate three different techniques in both studies. We present an additional study to further validate our design space using passive, low-fidelity techniques for speed control to choose the best performing techniques. In this study, we use the best performing technique for direction indication, the hand technique. For speed control, three techniques were evaluated to assess the influence of speed and transitions.

In addition, we present a summary discussion of the topics that have been presented on this thesis. Finally, a summarized list of conclusions is provided, focusing on the contributions and limitations of the work described in this thesis, while addressing clear guidelines for future work.

2 Related Work

This work relates to two interconnected domains of previous research, about travel and representation of people inside immersive virtual environments (IVEs), which are also referred as VR settings.

Firstly we introduce the concept of Representation fidelity and its sub-components Perspective and Graphical fidelity and how these factors affect the VR experience. Then, we define and classify travel in Virtual Environments VEs. We then relate the concept of interaction fidelity for travel techniques and how this relates to quality factors [13]. Finally, we interconnect all of the concepts presented and show an in-depth discussion about the most important work in these areas.

2.1. Representing the user on the Immersive Environment

Many are the factors that affects the VR experience, being Presence the most important of them. Presence [99] relates to the feeling of "being there" on the VE and is an important factor for providing a good experience in immersive setups. The feeling of presence is related to the concept of proprioception, which is the ability to sense stimuli arising within the body regarding position, motion, and equilibrium. Many are the ways of assessing how much users feel "in there" on the VE, some of them through questionnaires [40, 116] and through derived measures such as brain activity [83] and galvanic-skin response [56], which measures the stress of the experience through skin response.

An important part of the experience is how users are represented on the virtual scene. As opposed to CAVE-like systems [23], the use of Head-Mounted display technology occludes users' real self, compromising the overall virtual-reality session. A way of overcoming this problem is by using a fully-embodied representation of the user within the VE [97, 98]. Another related concept related to presence in Virtual Environments is the Sense of Embodiment, which is defined by Kilteni et al. [41] as the "sense that emerges when a virtual body's properties are processed as if they were the properties of one's own biological body". The sense of embodiment affects the way one interacts with virtual elements [41] and is an important aspect to enhance the illusion of being there in the Virtual Environment [8]. This concept is subdivided in three components: the sense of agency, i.e. feeling of motor control over the virtual body; (ii) the sense of body ownership, i.e. feeling that the virtual body is one's own body; and (iii) self-location, i.e. the experienced location of the self.

The use of a fully-embodied representation in conjunction with fully-immersive VR technology facilitates users to establish depth judgments between their real selves and the environment. Ries et al. [86] relates the improvement in distance estimation using a tracked avatar with the enhanced sense of presence it promotes. Unlike conventional display setups, where people are represented by an infinitesimal point in space, following the classic camera pinhole, full-immersive VR setups use stereoscopic technology to better perceive the world in three dimensions. This type of technology uses stereoscopic displays, generating a different image for each eye, that helps people to understand both monocular and binocular depth cues. Monocular cues are the depth cues perceived by one eye such as distance, occlusion and size;and binocular the ones that need both eyes to be perceived, such as convergence and shadows [25]. According to Cutting et al. [24], the relative importance of different depth cues is determined by the distance of the objects to the user. There are three different areas: Personal space (0 to 2 meters), action space (2 to 20 meters) and vista space (more than 20 meters). In the personal space, binocular disparity provides the most accurate depth judgments. A classic task to evaluate distance judgments is through open loop tasks, with procedures such as the blind-walking [109]. This type of task have been previously used to assess depth perception both on AR [109]

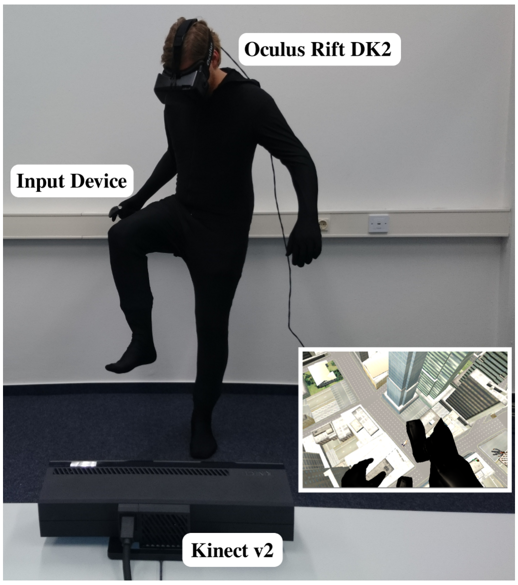

Figure 2.1: User being captured by a Microsoft Kinect and its representation counterpart on the virtual environment (Source: Langbehn et al. [46]).

and VR settings [70].

Mapping users' real movements in an avatar typically utilizes expensive equipment that uses optical tracking-based solutions. These kind of systems, such as the Vicon (Figure 2.2), are capable of tracking the full body of the user so that their body movements can be fully mapped with six degrees of freedom (6DoF) to their virtual representation within the IVE with millimetric accuracy. However, in order to provide full-body tracking users need to to wear special suits with reflexive markers attached to them [105].A cost-effective way to provide full-body tracking is by using the Microsoft Kinect, an off-the-shelf depth sensor that track body joints positions and orientations without the need to use special markers. The low precision of the joint position tracking and the fact that the Kinect sensor does not offer a way to manage occlusions, which is a huge problem in VR settings.Toolkits that combine information from various kinects were already proposed, but these toolkits are focused in context-aware scenarios and do not provide full-body tracking [58, 94, 117].

The level of realism of the avatar also plays an important part on the VR experience and how it relates to the sense of embodiment of an user [57]. A common problem on this matter is the uncanny valley [73], which states that the acceptability of an artificial character will not increase linearly in relation to its likeness to human form. Instead, after an initial rise in acceptability there will be a pronounced decrease when the character is similar, but not identical to human form. This phenomenon can be perceived in robots, third-party artificial characters (that are not controlled by the user) and also when people control their representation inside a Virtual En-

Figure 2.2: Vicon Motion Capture System.

vironment. Following the original study from Mori [73] that used static characters to define the Uncanny Valley, Piwek et al. [78] conducted a study with animated artificial characters.In this study, the authors tested seven different characters with varying levels of realism, from a more abstract (including dead characters such as a toy robot, a mannequin and a zombie) to a high-poly human character. Each of the characters was displayed in both static and animated forms. The animated characters were shown with increasing levels of motion quality (i.e. closer to the way real humans move). After the exposure from each character form, users were asked how would they classify them regarding the character human likeness with a 9-point Likert scale. Results from this study showed that when animated, the effect of character realism in the deepest part of the valley became more acceptable. In Immersive as in conventional setups the way people's avatars are rendered affects their perception in various ways, namely presence, embodiment, task efficiency and effectiveness. We define these elements as being part of a concept named Representation Fidelity, which is the level of closeness (or exactness) to the way people are represented and seen in the real-world. Representation fidelity has two components : graphical fidelity, which relates to how close to the user the avatar is and 2) perspective fidelity, being how their bodies are viewed in Immersive Virtual Environment (IVE). As such, the perspective fidelity varies from a Third-Person Perspective (3PP), where the camera is normally positioned behind users' heads and First-Person perspective (1PP), where the virtual camera is placed close to people's eyes, emulating how they see their bodies in real life. Regarding graphi-

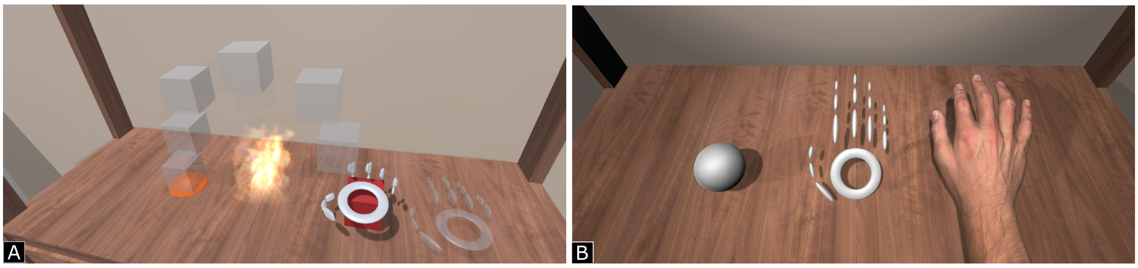

Figure 2.3: Argelaguet et al. [1]: (A) Pick-and-place task while avoiding obstacles (B) Representations used in the pick-and-place task. Abstract with no fingers (Left), Abstract Hand with fingers (Middle) and Realistic Hand (Right)

cal fidelity in Virtual Reality setups, Lugrin et al. [54] showed that when using a virtual mirror, there is a slight improvement in sense of embodiment in highly detailed user representations. Although, the authors did not noticed a considerable Uncanny Valley effect, as the user still maintained a high sense of embodiment when using a stylized avatar. The Uncanny Valley is also studied in HMD-based VR setups, where the real body of the user is completely occluded. Works by the same authors [56, 57], used a 1PP to simulate a hazardous situation, where parts of the virtual body and parts of the scenario would catch fire. As a result they noticed a potential Uncanny Valley effect, where avatars with high resemblance with the human form decreased in acceptance. Although many studies about the realism of representation in Virtual Reality setups are described in the literature, they are mostly focused on the embodiment and presence component parts. An interesting topic would be how the relation between this aspect would benefit classic 3DUI tasks such as selection, manipulation and navigation tasks. In fact, due to the nature of each of these tasks, each of them has its own particularities, so the impact of the realism of the representation is specific for each of them. Argelaguet et al. [1] for example, presented a representation study applied to a pick-and-place selection task (Figure 2.3). Their findings indicate that the use of a more abstract representation lead to improved sense of agency and increase task efficiency in this type of task. However, as seen, the effects of graphical fidelity in this study are still limited to the First-Person perspective in a partial, hands-only, type of representation in manipulation tasks.

The perception of depth of users is also affected by the type of equipment used on both VR [16] and AR equipment [3, 87, 106].Because of this, some parameters need to be adjusted for a more faithful distance estimation: Inter-pupillary distance (IPD) (separation of the eyes of the user) and Field of View (FOV), on Head-Mounted displays; and image separation, on stereoscopic displays,need to be adjusted for a comfortable experience. Misuse of these parameters can lead to misunderstanding

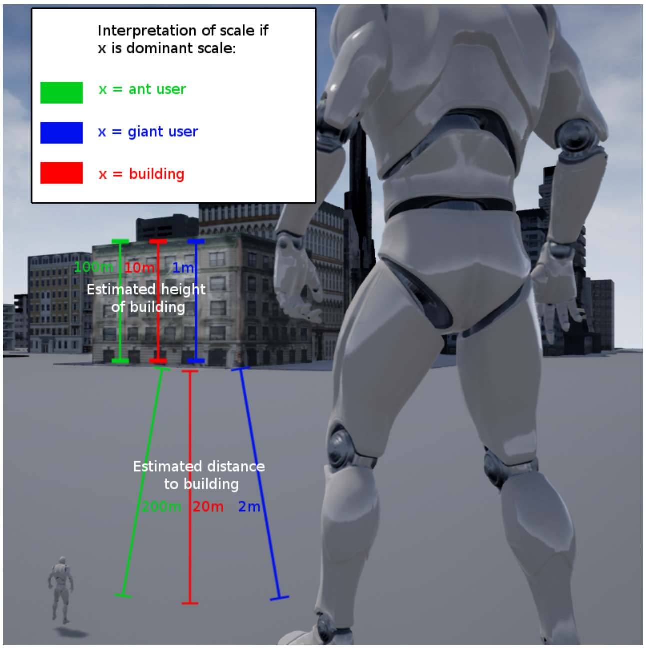

Figure 2.4: Without a common scale level the participants' notion of distances and sizes differs largely. Source : Langbehn et al. [46].

of scales on VEs [2, 19, 21]. A correct estimation of distance (or depth perception) is also an important part of the spatial awareness of users and how they interact with their environment.

Regarding avatars, studies indicated that users' depth judgments depend on characteristics like body shape [5], age [75] and gender.Having a virtual avatar that is faithfully scaled can provide users with more cues as to how they fit into the virtual world. A virtual body can supply them with a reference of recognizable size and connectedness to the VE [39, 85]. The ability for users to see their virtual feet when immersed is a way to improve their pose and ground themselves in the VE [70].If the user is incorrectly scaled, distances can be overestimated or underestimated (Figure 2.4) and when there is no virtual representation of the user, distance judgments are solely based on correct scaling of the environment. Langbehn et al. [46] combined different scale relations (Figure 2.8) and concludes that the dominant scale perception depends on (in order of importance) correct scale of the virtual self, realism of the scene and presence of other avatars in the environment. However, studies indicated that the use of virtual body in immersive setups may still cause distance underestimation [82].

Another factor that may influence the VR experience is the perspective in which the virtual body is viewed. On the First-Person perspective the virtual camera is placed near the avatar's eyes, simulating a real-life condition.The Third-Person Per-

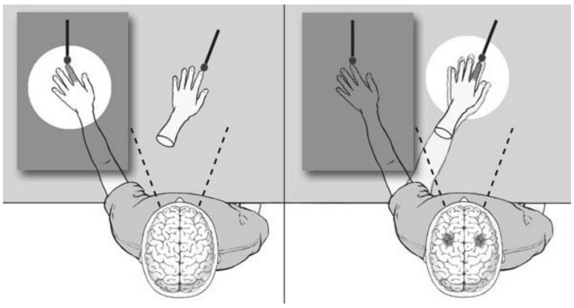

Figure 2.5: Rubber-Hand Illusion [9]

spective, on the other hand, places the virtual camera outside the virtual self (normally behind the avatar's head) giving users an external view of their virtual-self. This representation is widely used in games to improve spatial awareness in conventional displays [30, 93]. The use of a Third-Person Perspective may compromise the naturalness of the interaction but can improve awareness of their surroundings. Additionally, artificial bodies can still provide embodiment when a body is viewed in a different point-of-view. A classical extra-corporeal experience is known as the Rubber Hand Illusion (RHI) [9]. In this illusion, one of the users' real hand is hidden from view and an artificial hand placed in its place. Then, the artificial hand is stimulated and users feel that this artificial limb is their own. This illusion has a similar counterpart in VR setups, which is called Virtual Hand illusion, and can be induced by visuotactile [96] and visuomotor synchrony [91, 119].

The Rubber-hand Illusion has also proven to work with full body [77, 59]. Additional work by Ehrsson et al. [31] and Leggenhager et al. [51] proved that people can still feel embodied in VR when they see an image of their own body from a different point of view. In VR, using orthogonal third person viewpoints has been explored and was recommended to help setting the posture of a motion controlled virtual body [10]. This underestimation is also present when the avatar is seen on a third person perspective [89].Users reported that they adjusted their distance assessment more quickly in their personal space (0 to 2 meters from user) by using a third person perspective.

Salamin et al. [90] used an augmented-reality setup with a displaced camera and a HMD (Figure 2.7-C) to show that the best perspective depends on the performed action: First-Person perspective (1PP) can improve object manipulation precision (Figure 2.7-A), while the Third-Person Perspective (3PP) can improve performance in moving actions(Figure 2.7-B). Work by the same author also showed that people preferred the 3PP in comparison to 1PP and additionally 3PP required less training in a ball catching scenario [89].

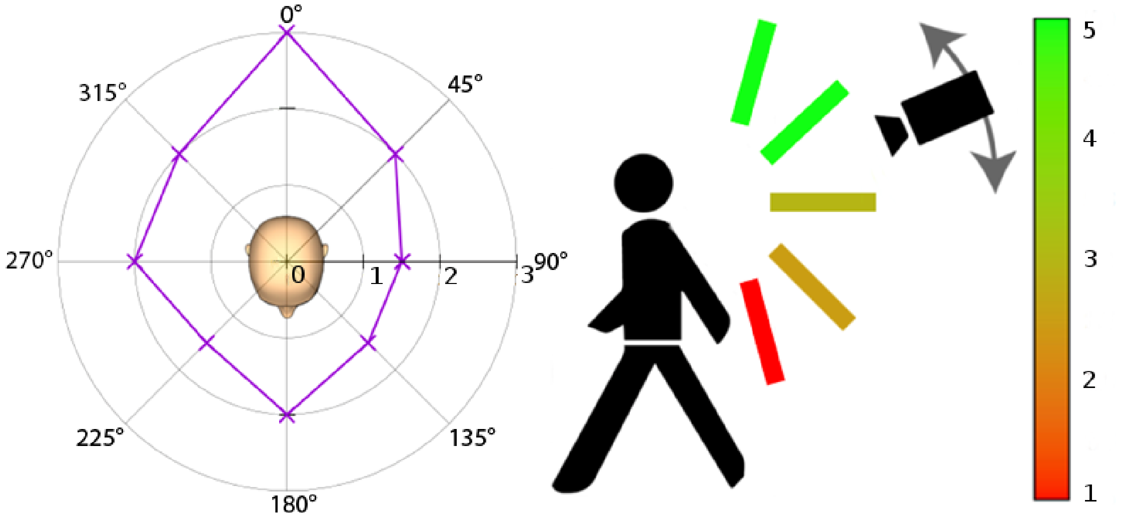

Figure 2.6: User preferences according to Kosch et al. experiment [43] regarding camera position and orientation. The Left side indicates user preferences along the x-axis. Right side indicates orientation preferred on the y-axis

Figure 2.7: (a) Simulated first-person perspective (b) Simulated third-person perspective (c) Setup camera used to position the displaced camera facing the user [89]

For determining the best position and orientation for the camera in a Third-Person Perspective on a real environment, Kosch et al. [43] used a video-camera that was positioned on top of a stick. The position and orientation of the stick was modified along the x and y axis. Users performed a test which consisted in two phases : 1) walking along a pre-defined path and 2) walking around an environment while avoiding obstacles. The study indicated that users preferred the camera position far from their body and above their heads, to augment their Field of View (FOV) and be able to see their feet [43].

Perspective studies in VR generally focus on how an artificial body affects the embodiment of a user from a Third-Person Perspective. Although the effects of this perspective in the embodiment are an important aspect of the overall experience, few works focus on how these effects affects classical 3DUI tasks such as navigation, selection, and manipulation. Work by Boulic et al. [10] used a large-screen

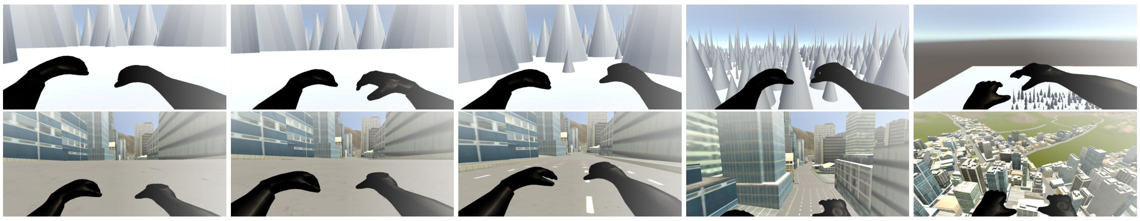

Figure 2.8: Without a common scale level the participants' notion of distances and sizes in the pilot study differed largely and hindered a meaningful collaboration in spatial tasks [46].

display to determine that the use of third-person orthogonal avatars helps to set a virtual avatar's posture. Debarba et al [29] on the other hand, showed that users can accomplish reach tasks with a high sense of embodiment using both 1PP and 3PP, but have reduced accuracy in 3PP avatars. Additionally, Monteiro et al. [72] used both 1PP and 3PP avatars and suggest the use of avatars in 3PP in order to reduce cybersickness related side-effects. The 3PP is also found to be safer when compared to 1PP in harmful situations [11, 28]. For example, Debarba et al. [28] used the classic Meehan et al. [66] pit scene in which the user followed an initially wooden floor pit and when the user went up on a wooden ramp, the pit fell revealing a large whole in the ground. This work tested three different avatar conditions : 1PP, 3PP and a condition where the participant could alternate between the First- and Third-Person Perspective. Results showed that the three conditions elicited a high sense of embodiment, but users felt a more elevated subjective sense of body-ownership related to the threat in the 1PP when compared to the other two tested conditions. Regarding graphical fidelity in third-person avatars, studies on this matter are still limited to non-rigged avatars and indicate that an avatar with a realistic human-shaped form increases the sense of body-ownership, producing a full-body illusion [59]. Nonetheless, 1PP is still more efficient than the 3PP and considered more natural by users, both on reach [29] and travel tasks [37].

Although studies show a slight improvement in spatial awareness in displaced see-through systems in 3PP, the use of artificial bodies in VR can produce different results. In VR, the use of a different perspective coupled with a virtual representation that not match the users' bodies may aggravate how people make distance judgments when the avatar is animated [71], compromising users' spatial awareness. Mohler et al. [71] proved this by comparing three conditions in a blind-walking test : user without a body, with a static body and with an animated body. Results from this test indicate that the animated body diminish the known underestimation problem found in IVEs both in first and third-person views. Previous work [37] claimed improvements in spatial awareness in VR with 3PP avatars over 1PP, but

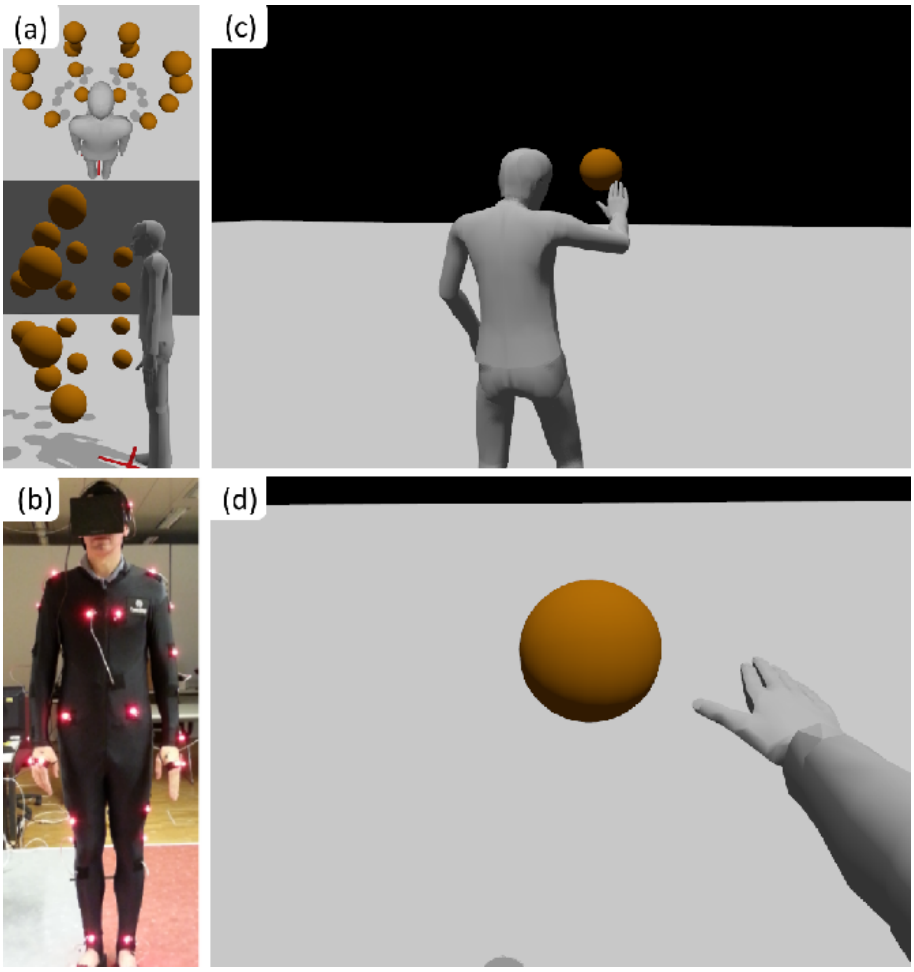

Figure 2.9: Experimental setup used in Debarba et al. [29]. (a) Location of targets that subjects have to reach; (b) Motion tracking suit and immersion equipment; (cd) Illustration of what subjects saw during the reaching task in 3PP and 1PP respectively.

their results are limited to subjective responses and, while showing a slight tendency towards 3PP being better, they have no statistical significance. Further objective metrics are needed to assess not only participants' preferences, but also objective measures of this type of representation. Moreover, since users' bodies are always visible when a Third-Person Perspective is used, we theorize that the realism of the representation have a bigger influence in both the sense of embodiment and spatial awareness in this perspective. Therefore, the use of a real-time reconstruction of people's real bodies can be an important factor for establishing a high sense of embodiment with a third-person view.

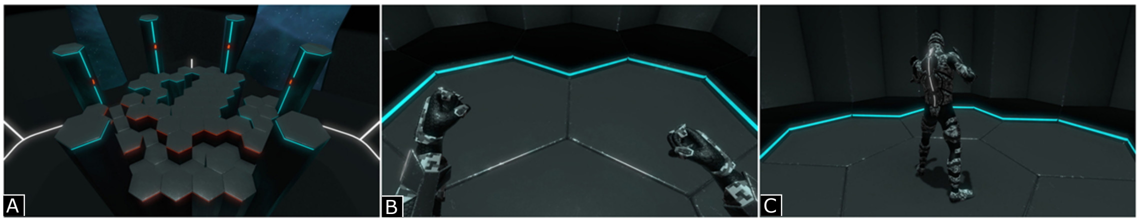

Figure 2.10: Representations used in Gorisse et al. [37] : A) Virtual Environment B) First-Person Person Perspective C) Third-Person Perspective

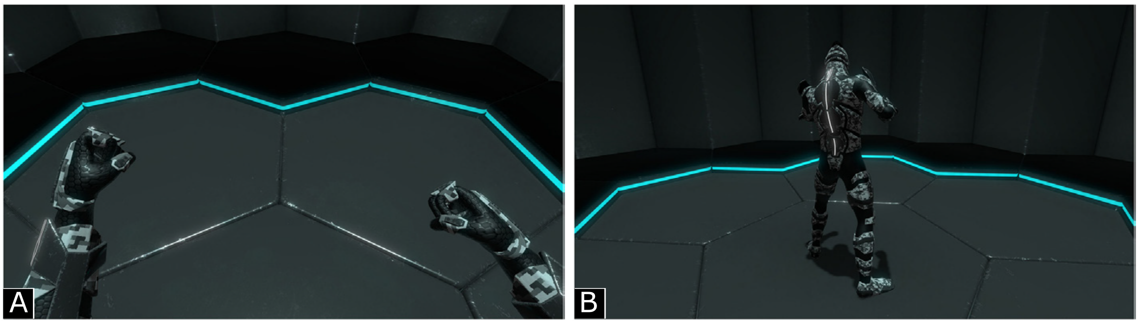

Figure 2.11: Representations used in Morisse et al. [72] : A) Virtual Environment B) First-Person Person Perspective C) Third-Person Perspective

2.2. Travel Techniques

Travel is the part of navigation where users perform the action of moving from one place to another in a given direction in the VE.Travel techniques can be of Exploration, when users have no specific goal to navigate through the environment and Search, when users have a goal and may or may not rely on additional information to assist him to get to his goal. The cognitive part of the navigation is called wayfinding [50] and involves the spatial understanding, planning and decision making. In this classification, Travel may also be classified as a form of manipulation, because it is the manipulation of a virtual camera or users' views within the VE. An example of this is the World In Miniature (WIM) [107], where users navigate and manipulate objects on a reduced version of the VE. The use of VR setups enable users to perform faster and following a more fluid path to navigate in large environments using a HMD in comparison with conventional WIMP setups [88] To effectively evaluate travel techniques, Bowman et al. [13] identified a set of quality factors. These quality factors included the use of appropriate speed, ability to the user be as close as possible to the desired target (accuracy) , spatial awareness it provides, the facility of a novice user to use the technique (ease of learning), cognitive load of the technique (ease of use), user's ability to actively obtain information from the VE during travel and sense of being inside (or within) the VE (presence).

There are many classifications for travel techniques in the literature [68][15][4]. One of those [15], divided the techniques into active navigation tasks, where users directly control the locomotion inside the VE; and passive (such as Target-based techniques [64]), where the movements are controlled by the system. Another taxonomy classifies the techniques by the way the navigation occurs in the virtual environment [68], either in a physical or virtual way. In physical navigation users control rotation and translation moving their bodies tracked by a dedicated system with 6DoF; in virtual techniques remain stationary while the movement is done,people control movements via a specific interaction device, such as a Flystick [22] or tablet [65] that can be tracked to determine the direction of movement. According to LaViola et al. [50] both classifications are complementary, making it possible to combine different techniques of either category in one system. For example, users can control direction using their gaze, while controlling speed with a joystick.

The physical travel category intends to emulate natural movements of the human body. One of the first uses of this type of technique was the walking metaphor. Although this is the most natural form of navigation, it presents some problems such as the limitation of space.One way to solve this problem is called Walking In Place [99]. In this technique the user emulates the gesture of walking without moving, decreasing the limited physical space needed, but compromising the realism of interaction [15]. Another possibility is the use of Redirected Walking Techniques [17], in which the spatial limitation is overcome by interactive and imperceptible rotations of the Virtual Environment around the user. The rotation causes the user to walk continually toward the furthest wall of the room without the user noticing the rotation. This could also be overcame by omnidirectional treadmills which uses special hardware to enable the user to walk in all directions [110] (Figure 2.15-A).

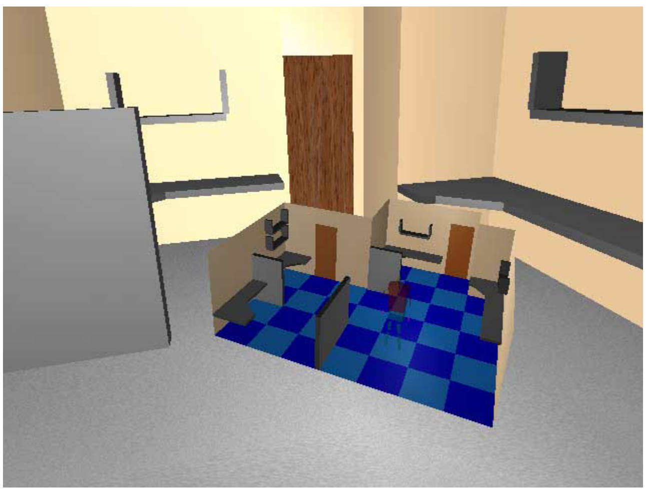

Figure 2.12: World In Miniature [107].

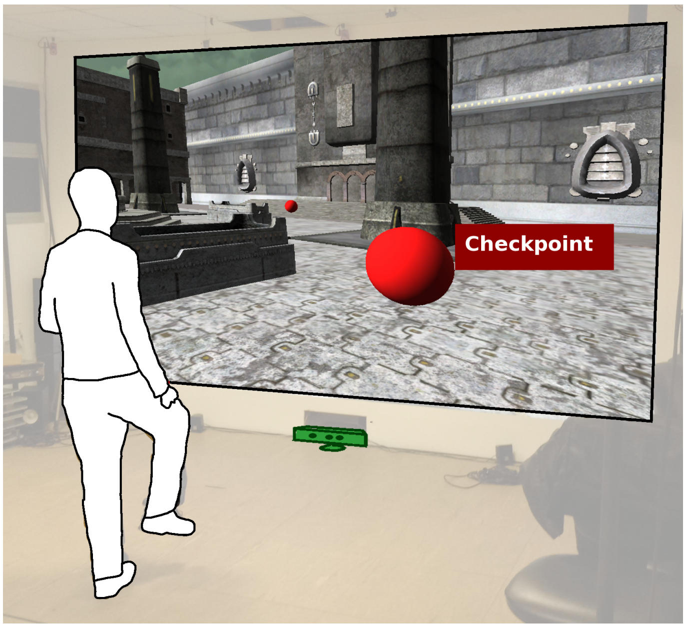

Figure 2.13: User walking-in-place to control locomotion in a VE [18]

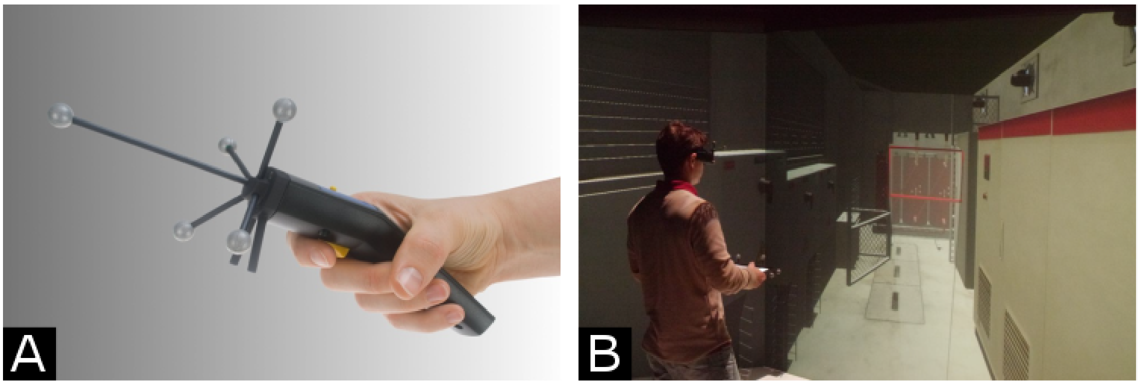

Figure 2.14: Devices used to control virtual techniques : (A) Flystick (B) Tablet with 6DoF Tracking

Some uses of physical navigation techniques are those that use the Microsoft Kinect. This device is able to track user movements and use them to make interactions with the virtual environment [81]. A example of travel technique which uses the Kinect is the Virtual-Circle [26]. This technique resembles the analog controls of a gamepad, but using users' bodies. When they go outside a virtual circle of fixed radius he is translated according to his walking vector (Figure 2.15-B).

Another way of classifying travel techniques is by task decomposition [13], which divided them by 1) direction specification 2) speed control and 3) input conditions [13] (Figure 2.19). On the direction specification phase users use parts of the body or other devices (such as mouses and joysticks analog axis) to indicate direction movement. The speed control phase consists on controlling the speed to reach a goal.

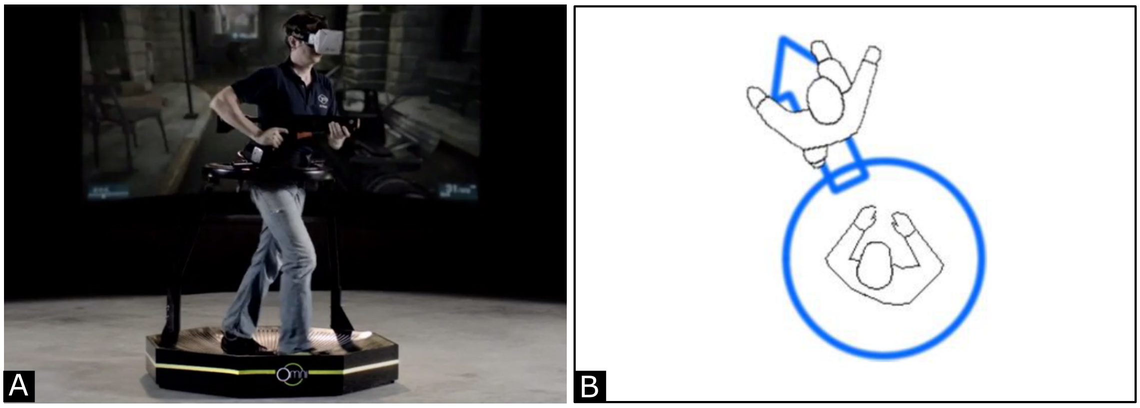

Figure 2.15: (A) Omni Treadmill (B) Virtual Circle [26]

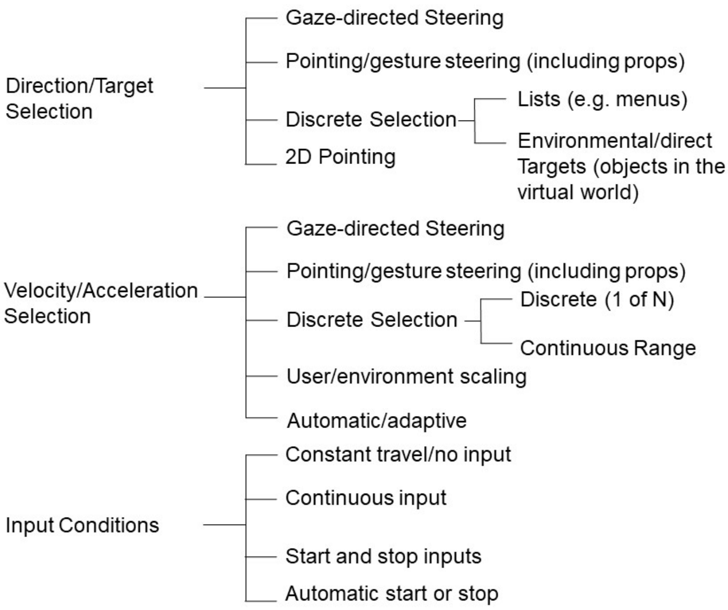

Figure 2.16: Travel classification by decomposition. Adapted from Bowman et al. [13]

This control can be made solely by the user (using a joystick), by the system (such as Target-based Techniques [80]) or jointly by the system and the user. When the speed is "jointly" controlled, the system determine the speed of travel based on the distance between the objects of the scene and the observer, and users can use additional controls to constraint navigation. An example of this is the Drag ń Go Technique [69] where the user selects a point-of-interest POI, and then users are translated according to a straight line between the users and the POI. Users can then constraint navigation using an smartphone touch-screen (Figure 2.17). The input conditions phase

There are works that attempt to mix physical and virtual navigation techniques in

Figure 2.17: Drag 'N Go Technique [69]

a virtual environment. A very common way to mix the two forms is to use users' movements to control the point of view on the virtual environment, deforming the projection matrix of the environment according to users' position within the immersive environment, using specific devices such as the ART-Flystick for navigation, manipulation and selection tasks. In one of these works, Cirio et al. [22] combines different metaphors with Real Walking to keep users in a secure position in relation to a CAVE immersive environment, where a tracked device can be used to perform the navigation task.

One of the factors that makes navigation difficult in VEs is user disorientation. Smith et al. [100] considered that the two leading causes are the absence of visual cues and problems with navigating too close to or through objects. Another factor that could lead to disorientation on VEs is the lack of control while travelling. Travel techniques where the users do not physically translate their bodies, such as Steering Techniques, can allow the maintenance of a satisfactory spatial awareness. This is justified by having some control over the movement of the user's head to explore the VE [12]. Even though, Target-based techniques (where the system shifts users' position automatically based on predetermined speed parameters) may be more appropriate for users who are more susceptible to simulator sickness in comparison with Steering techniques [80].

Regarding travel on Virtual Reality setups, McManus et al. [63] reported that users perform locomotion tasks faster and with more accuracy when an animated self-avatar is present in the VE. Moreover, a self-avatar viewed at first person perspective is known to be an important aspect in establishing embodiment [77].

Another crucial issue when considering travel technique design is cybersickness. Cybersickness can be distinguished from motion sickness, in that users are in a stationary position but has a sense of self-motion through the moving environment. A common form of assessing cybersickness is through the Simulator Sickness Questionnaire (SSQ), where need to rank nausea, oculomotor and disorientation questions using a 4-Value Likert Scale. LaViola et al. [48] suggest that there are too many causes for cybersickness and that there is no reliable approach for eliminating this problem. They also note that all theories regarding the causes of cybersickness should take the individual into account.

Lin et al. [52] conducted a study that concluded that VR-based dynamic environments can physiologically and realistically cause motion sickness on its users. They describe that these phenomena were almost the same as those induced in a real environment. When comparing the onset of cybersickness on users of two distinct roller coaster experiences [27], the one with higher level of graphic realism was considered to be the one that caused a faster onset of nausea.

So et al. [101] investigated the effects of navigation speed on the level of motion sickness with head-steered techniques in immersive virtual environments. They report that nausea and vection (illusion of self-motion in the opposite direction caused by wide field of view in the absence of physical motion) sensation increased with the raise in navigation speed from $3m/s$ until $10m/s$ where they stabilized until the $59m/s$. They conclude that the mean offset time of vection was earlier than that of nausea, which they say is consistent with the knowledge that cybersickness is a type of vection-induced motion sickness.

Fernandes and Feiner [32] referred to VR sickness as a direct factor from the diminished field of view (FOV) present in the head-mounted displays available. To examine this effect, they dynamically change a seated user FOV while they explore a VE. Even with a small number of participants, the FOV restrictors helped the users feel more comfortable and able to stay in the VE for longer. These restrictors were not noticed by the majority of the users. Plouzeau et al. [79] study suggested merely that the addition of the vibration which mimicked the action of walking in a virtual environment may decrease cybersickness. However, studies indicated that the Walking In Place metaphor, a natural travel technique, can cause increased cybersickness effects.

Young et al. [118] presented a study to determine if users' subjective response to Simulator Sickness Questionnaires is a consequence from pre- and post-test measurements. Their results indicate that motion sickness after being immersed in a VE are considerably higher when the users were instructed to fill out both pre and post questionnaires than when only a single post test questionnaires was answered.

However the use of natural metaphors for travel are not that easy as the VE grow on size and complexity, such as the example of multiscale virtual environments (MSVEs). Multiscale Virtual Environments (MSVEs) are virtual environments which encapsulate different levels of scale within the same environment [7, 34, 49, 76]. On oil fields (a type of engineering scenario) for example (Figure 2.18), there are elements varying in a scale of $1:10^7$ from the smallest object (an oil tube with a 15cm radius) to the largest (a seismic object with possibly kms of extension in all three dimensions). One of the most common classifications differs on how the system handles transitions between the different levels of scale : Level of Scale (LoS) [42], where the transition between the different levels of scale are made by scaling users or the environment; and automatic speed adjustment techniques [61, 113, 92, 38] on the other hand, the transition between levels of scale is smoothly made by adjusting speed while navigating.

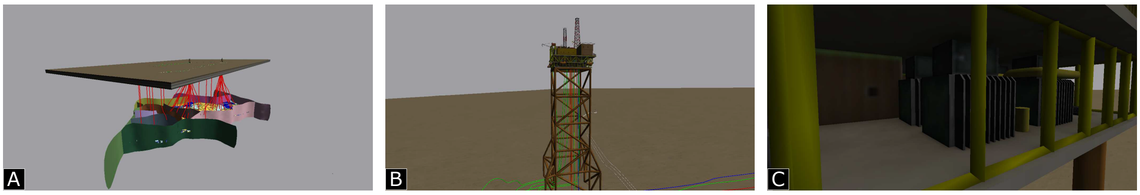

Figure 2.18: Example of three different levels of scale within an oilfield model. (a) oilfield (b) view of the whole oil-platform (c) inside the oil-platform [113]

A related problem is how users perceive scales inside a Virtual Environment. This is even more latent when there are elements with diverging levels of scale, such as Multiscale Virtual Environments [121]. McCrae et al. [60] stated that users can become disoriented when there is a representation of objects in different scales. It is also unnatural to the user to perceive the difference between levels of scale, as the scale level of human beings is very limited, ranging from centimeters to hundreds of meters [120]. Another problem is that the user, when represented inside the VE, does not have a size, as he is represented by an infinitesimally small point in 3D space. In this sense, the human experience of interacting with a physical 3D environment does not assist users in reasoning about absolute scale in a virtual environment. This is partly because of the lack of representation of the user on the VE, so the size judgments are entirely based on deductive reasoning and judgments [36]. It is also known that the user finds it easier to navigate in a large-screen display in comparison with a normal monitor because he does not see the borders so much. A way of overcoming this problem is by providing the user with a full-sized human representation of their bodies.

Another problem in complex environments such as the Multiscale Virtual Environments is that sometimes the user needs to reach specific areas of the VE. In the specific case of Multiscale Virtual Environments, while navigating between the different levels of scale, the user normally is not constrained to a floor, needing them to use a flying metaphor to reach specific parts of the virtual scene.

2.3. Interaction fidelity and Travel

Level of Fidelity is the concept which "relates the level of exactness of real-world experiences to the ones reproduced by a computing system" [63]. When dealing with users actions, authors use the concept of Interaction Fidelity, which is the "the objective degree of exactness with which real-world interactions can be reproduced" [62]. In Travel techniques, the similar the travel technique is to real-walking, the higher the level of interaction fidelity it has. However, physical space constraints and the presence of obstacles in the physical spaces make the real-walking metaphor not always desirable. Also, the big-size of some virtual environments can make the physical effort of reaching a goal inconvenient and lead to physical fatigue.

Previous studies pointed the relation between interaction fidelity and quality factors [62, 108] which found that the level of fidelity does not directly affects those factors. More importantly, McMahan et al. [62] also pointed that the impacts of fidelity of interaction are dependent on the application scenario. However, this relation was only observed in close-to-real scenarios, limited by a ground plane.Moreover, the inclusion of additional degrees of freedom could make the control of travel not such an easy task.

Regarding travel on VR setups, McManus et al. [63] reported that users perform locomotion tasks faster and with more accuracy when an animated self-avatar is present in the VE. Moreover, a self-avatar viewed at first person perspective is known to be an important aspect in establishing embodiment [77]. The level of interaction fidelity of travel is also known to influence the level of presence [99, 114] and distance estimation [53] in Immersive Virtual Environment (IVE)s.

Although effective, physical travel metaphors are often restricted to a ground plane. But in some cases, such as inspection explorations, there is the need to have more degrees of freedom to explore the virtual environment. An interesting application of flying metaphors for travel is in Multiscale Environments, because when navigating through different levels of scale the user may need to control additional degrees of

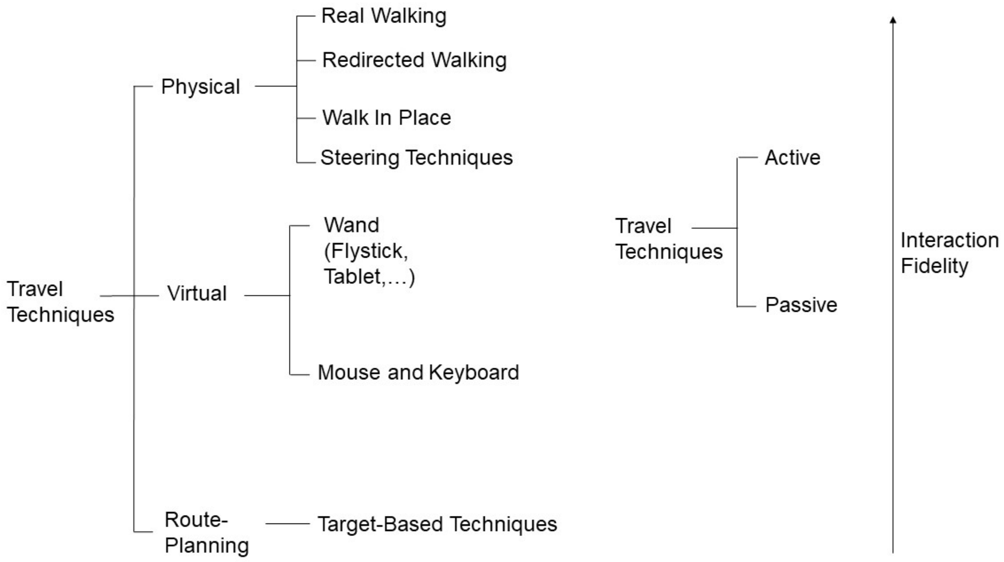

Figure 2.19: Taxonomy Used for Travel Techniques. The arrow direction shows the increase in the level of interaction fidelity of the technique

freedom to reach specific points of the Virtual Environment (VE).Since flying in the scene is not natural to humans, the control of direction, translation and speed of movement could be a difficult task. The use of high-interaction fidelity travel techniques can overcome this problem. However, due to the unnaturalness of this metaphor, high interaction fidelity techniques may not be directly mapped into flying tasks.

Virtual techniques and techniques with high interaction fidelity, such as real walking and redirected walking, are considered to elicit better results than flying. This advantage is observed only in human-scale environments [114] and in instances where the desired travel destination is in a ground-constrained location [33]. However, due to the supernatural quality of some large VEs, such as multiscale VEs [61], travel targets may reside out-of-reach, e.g.: above ground or in remote spots of the VE. Therefore, traditional, ground-based techniques alone are not sufficient to effectively support travel.Thus, flying metaphors provide the most flexible technique to navigate arbitrary virtual environments.However, the flying metaphor is still unnatural to humans. It requires people to control simultaneously 6DoF related to control of movement (rotation/translation), while concurrently controlling the additional DOF related to speed. This control is far from what people are used to in real-life. Previous works try to mitigate this issue by employing complex contraptions to emulate flying machines, such as a paraglider [102](Figure 2.20-A), a space

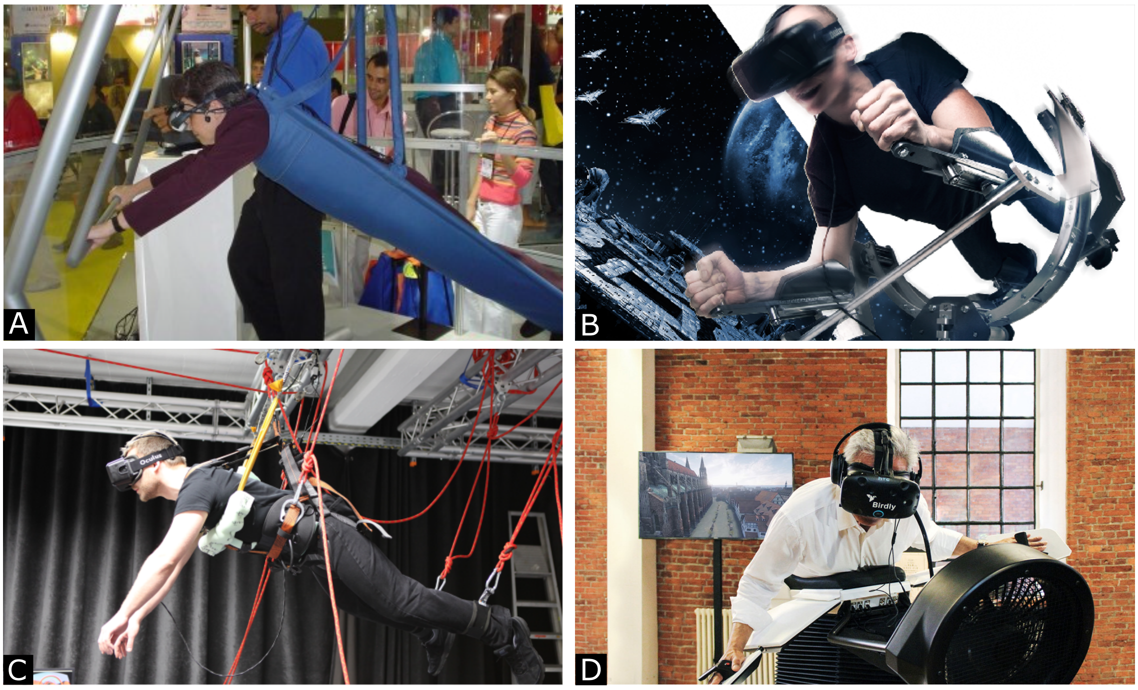

Figure 2.20: Setups for Flying in VR: (A) Hand-Glider [102] (B) Icaros VR (C) Zero-Gravity Simulator (D) Birdly

ship[1](Figure 2.20-B) or a zero-gravity simulator [44, 45](Figure 2.20-C) to indicate direction of movement using their bodies, while controlling speed with additional buttons. Birdly [84], on the other hand, uses a complex setup to enable users to fly by flapping mechanical wings, that can also control speed of movement (Figure 2.20-D). While these devices offer efficient ways to fly in VEs, in more intricate scenarios, when reaching a certain location, users need to further interact with the VE, either by selecting, manipulating or creating content. There, these setups restrain users' actions to one task and thus are hardly suited to richer interaction contexts.

Techniques with a higher degree of interaction fidelity have been proposed to overcome this conflict. However, most studies are focused on the direction-indication stage. Notably, work by Chen et al. [20], showed that users perform better with a physical technique (gaze-steering) as compared to a virtual technique to indicate direction in a flying scenario. Chen et al. [20] tests two techniques for 6DoF direction control: a virtual joystick-based flying technique, with a physical head-based controller technique. On both techniques the authors use 6DoF separation to control people's orientation and translation. The authors divide the direction control in two phases : 1) Coordinate System definition: either using an analog-axis,on the Joystick, (Figure 2.21-A) or head-orientation(Figure 2.21-B); and 2) axis selection, using a button or analog-axis on the joystick (Figure 2.21-A) or body movements,

[1]Icaros: Virtual Reality Fitness Experiences. Available in: `https://www.icaros.com/`

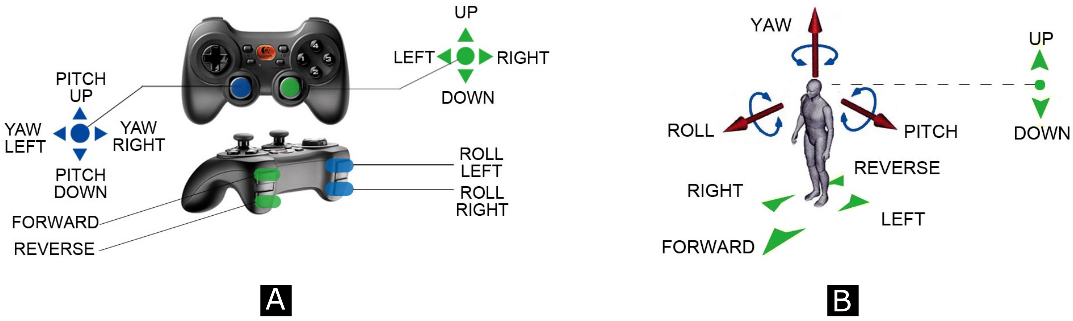

Figure 2.21: Chen et al [20] tests two different techniques for 6DoF direction control: (A) Joystick-Based Controller. (B) Head-Based Controller

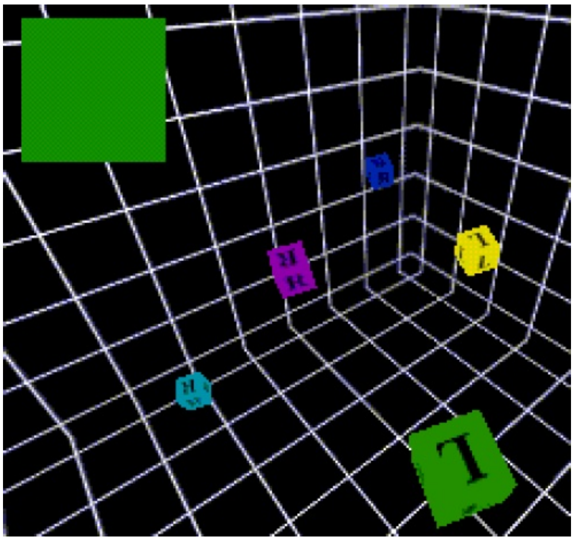

Figure 2.22: Virtual Environment used on Bowman et al. [13]

on the head-based (Figure 2.21-B). Results from this study show that the use of body movements' to control direction can lead to more precise movements and may lead to reduced cybersickness.

On a similar vein, users can specify direction by using their bodies either standing standing [112] or sitting down [6]. Work by Tong et al. [112] presented a technique that use leaning to control forward/backward movement while arms movement (up/down) control the movement in the z-axis. However, the authors do not provide a test to assess quality factors of the proposed technique against a baseline technique.

Techniques with a higher degree of interaction fidelity have been proposed to overcome this conflict. However, most studies are focused on the direction-indication stage. Chen et al. for example, compared a 6DoF joystick controller (Figure 2.21-A) with a Head-steering metaphor (Figure 2.21-B) in a CAVE setting. Results show that the use of a physical metaphor (head-steering) outperformed the virtual technique in time completion and path quality. Users also reported a decreased sense of cybersickness and a moderate increase in presence with the head steering technique. On a similar vein, users can specify direction by using their bodies either standing standing [112] or sitting down [6]. The Point-and-Fly technique, for example, uses the orientation of a 3D Wand to indicate direction while using the horizontal distance between head and hand to determine speed of movement [33] Tong et al. [112] proposed a technique that use leaning to control forward/backward movement while arms movement (up/down) control the movement in the z-axis. However, the authors do not provide a test to assess quality factors of the proposed technique against a baseline technique.

Separation of degrees of freedom has also been proposed to mitigate the unnatural behaviour of the flying metaphor. This approach is a common strategy for improving precision in 3D object manipulation in VR [67]. Since travel can be classified as a form of manipulation, as discussed by LaViola et al.[50], this strategy is an option for controlling direction of movement in flying scenarios.

A first attempt on this matter was the ChairIO [6] which consisted of a stool, that allowed the user to control rotation on the horizontal plane by rotating the stool or to lean in a chosen direction to control both the direction and speed of movement. Additional pressure sensors provided limited movement in both directions along the Z-axis [6]. However, no evaluation was carried out to evaluate the DOF-separation strategy for flying in VEs using the proposed device. Work by Bowman et al. [13] compared both hand- and gaze-oriented steering in a 6DoF translation environments and found similar performances in both cases, although the results were still preliminary owing to the lack of obstacles and the absence of user representation in the VE. Wang et al. [115] used a leaning approach and devised two different techniques, one using a frontal stance (with the user's feet facing the VE) and one using a sidewise stance, to fly in VR. The results of this study showed generally better results for the frontal-stance technique, but among the 12 tested participants, 3 (or 25%) left the test due to severe cybersickness side effects. A related approach by Sikström et al. also showed that physical techniques can offer an improved sense of embodiment [41] in a flying scenario as compared to joystick control [95] when a virtual body is present.

Another proposed device for this matter is the CharIO [6]. In this device enable the decoupling of degrees of freedom (DOF) to fly in virtual scenes. This device consists in a stool, where users can control rotation on the horizontal plane by rotating the stool, leaning towards a chosen direction to control both direction and speed of movement. Additional pressure sensors provide limited movement in both directions

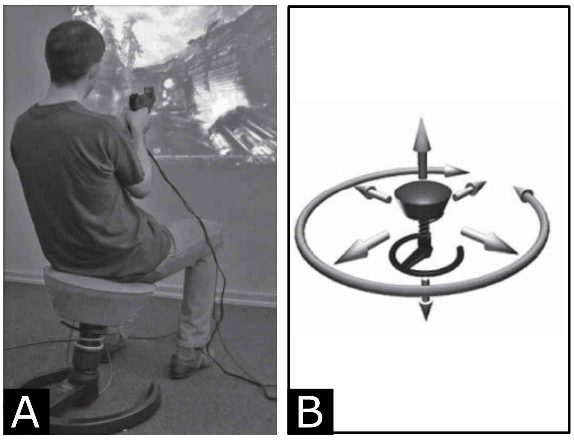

Figure 2.23: CharIO device: (A) User interacting with a virtual environment using the device (B) Possible Controls

in the Z-axis [6]. Work by Bowman et al. [13] compared both hand and gaze-oriented steering in a 6DoF translation environment, and found similar performance in both cases, although the results were still preliminary due to the lack of obstacles and the lack of user representation on the virtual environment. The use of a virtual body is also shown to improve sense of embodiment when a physical technique is used for flying in virtual scenes [95].

We propose to improve travel quality factors in flying tasks by addressing this metaphor with fidelity. To this end, we isolate travel in two phases, direction indication and specification to isolate simultaneous dof control and address each of these phases separately. This separation allows the use of techniques with a higher level of interaction fidelity such as the WIP to control motion speed.

2.4. Discussion

Representing correctly the user in the immersive experience is a part of the success of a VR session, particularly on travel tasks. The use of a fully-tracked representation improves the way users' interact and even increase efficiency and distance judgments inside a VE. Mohler et al. [71] proved this by comparing a static avatar and a fully-tracked representation against a baseline that did not provide any body feedback. This experiment showed an improved performance and preference when using the fully-tracked representation. There are three main components of the experience regarding self-representation in VEs , namely representation fidelity and its sub-components : graphical fidelity and perspective fidelity; and interaction fidelity.These factors highly impact the immersive experience.

Regarding the particular graphical fidelity sub-component, a common problem is the uncanny valley [73], which studies the acceptability of representation of people's bodies do not increase with the . Even though when an avatar is animated the effect of realism in the deepest part of the valley becomes more acceptable [78]. It is also shown that the uncanny valley affect the feeling of embodiment when the self-embodied avatar is viewed through an HMD in a first person perspective (1PP) [56, 57].

It is also possible for the user to feel an artificial body as their own, either in real scenarios [89] or virtual environments [29]. When real body movements are mapped accordingly, users can feel this body is their own [9] and also locate themselves [31] in the virtual environment. This feeling is called Visual Body Illusion and is also recommended to improve posture of a motion controlled virtual body [10]. Regarding perspective, works by Salamin et al. [89, 90] stated that the use of different perspectives can be suited for different situations, First-Person Perspective for example could be used for more precise operations such as object manipulation and Third-Person Perspective (3PP) are better suited for moving actions. The authors also indicate a slight improvement in spatial awareness when using a Third-Person Perspective in a See-through augmented reality setup. Since users' bodies are always seen in the Third-Person Perspective it may indicate that this effect can be only observed when a highly detailed representation is used that provide real-time reconstruction of users' bodies during the VR session.

This is explained because, when seeing their own body on a third-person perspective it provides the user a better spatial awareness of the virtual, or augmented, scene. However, results in the effects of graphical fidelity in third-person avatars are still limited. Maselli & Slater [59], used static avatars with head-tracking and find that realistic avatars in human form improve the feeling of embodiment in the Third-Person Perspective, but still not enough to produce a full-body illusion. Nonetheless, 1PP is still more efficient than the 3PP and considered more natural by users [29, 37].

In VR, the best representation depends on which task being executed. Debarba et al. [29], for example tested both 1PP and 3PP for reaching tests, their findings indicate a slight improvement in 3PP for objects that are closer to the user. Earlier work also seems to confirm an improvement in spatial awareness with 3PP avatars in VR. In both Gorisse et al. [37] and Monteiro et al. [72] the third-person was used to expand user's view and be able to see further parts of the virtual environment. For example, in Monteiro et al. [72] the task consisted in controlling a vehicle and in the Third-Person perspective were able to see further details in the road and respond faster when further actions such as turning were needed. Gorisse et al. [37] on the other hand, based the improvement in users' spatial awareness with the reaction of moving objects being thrown towards the user, something that was already confirmed by Salamin et al. [89]. For the navigation phase, the study [37] used solely subjective metrics to assess the spatial presence.

Travel is the part of navigation which consists in taking the user from a point A to a point B in a given direction in the virtual environment. These subset of techniques can be classified regarding on how the users travel [68], in an active way (where the user actively translates himself in the VE) or passive (where the system is responsible for the movement for the user reaching his goal). A complementary classification proposed by Bowman et al. [15] include the way the movement is made, physical : when the person use his body movements to travel and virtual, where the travel is made by a device such as an ART flystick and a tracked tablet [65]. Another possible classification is by task decomposition, which divides travel techniques in three categories : direction specification (or indication), velocity specification (or speed control) and input conditions, which indicates how a movement is continued and terminated (continuous, discrete,etc) [15]. A related concept is interaction fidelity [62] which indicates how close the designed technique is to its real counterpart, which is in this case, the way people walk in real life. But, physical constraints of the real space and the scale of the VE may limit their usage. A way to solve this is the Walk In Place[99]. In this technique the user emulates the gesture of walking without moving, decreasing the limited physical space needed, but compromising the realism of interaction [15]. Another way of overcoming the problems of natural navigation is by mixing elements from both physical and virtual classifications. Cirio et al. [22] combines different metaphors with Real Walking to keep the user in a secure position in relation to a CAVE immersive environment, where a tracked device can be used to perform the navigation task. The use of high fidelity travel techniques is recommended, but quality factors are not always improved as interaction fidelity level increases [62]. But, when a high fidelity display is used, the use of high-interaction fidelity metaphors are known to improve efficiency [62]. Also, recent work by Langbehn et al. [47] suggest an improvement in travel quality factors, such as presence and spatial awareness, when a high-interaction fidelity is used, when compared to a virtual low-fidelity technique.

Flying is not the most natural efficient way to travel in VR [112], but , traditional travel metaphors are not sufficient to reach targets which reside above ground. Such

Paper	Task	Devices	
		Input	Output
Mori et al.	Uncanney Valley Study	Mouse & Keys	Monitor
Mohler et al.	Embodiment Study	Body Movements	HMD
Lugrin et al. 2015a	Embodiment Study	Body Movements	HMD
Lugrin et al. 2015b	Embodiment Study	Body Movements	HMD
Argelaguet et al.	Embodiment Study	Body Movements	HMD
Salamin et al. 2008	Walk, Avoid Obstacles	Body Movements	Video See-Through
Salamin et al. 2010	Embodiment Study	Body Movements	Video See-Through
Debarba et al	Reach Task	Body Movements	HMD
Maselli & Slater	Embodiment Study	Body Movements	HMD
Monteiro et al	Avoid Obstacles & Walk	Body Movements & Joystick	HMD
Gorisse et al.	Game	Body Movements	HMD

Paper	Avatar Representation		
	Type	Graphical Fidelity	Perspective Fidelity
Mori et al.	-	Realistic	High (1PP)
Mohler et al.	Full-body	Mesh/Realistic	High (1PP)
Lugrin et al. 2015a	Full-body	Mesh/Realistic	High (1PP)
Lugrin et al. 2015b	Full-body	Mesh/Realistic	High (1PP)
Argelaguet et al.	Hands-Only	Realistic	High (1PP)
Salamin et al. 2008	Full-body & Dynamic	Realistic	High (1PP)
Salamin et al. 2010	Full-body	Realistic	High (1PP) & Low (3PP)
Debarba et al	Full-body & Dynamic	Mesh/Realistic	High (1PP) & Low (3PP)
Maselli & Slater	Full-body & Static	Mesh/Realistic	High (1PP) & Low (3PP)
Monteiro et al	Full-body	Mesh/Realistic	High (1PP) & Low (3PP)
Gorisse et al.	Full-body	Mesh/Realistic	High (1PP) & Low (3PP)

Paper	Travel		
	Classification	Technique	Interaction Fidelity
Mori et al.	Physical & Active	-	-
Mohler et al.	Physical & Active	Real Walking	High
Lugrin et al. 2015a	-	-	-
Lugrin et al. 2015b	-	-	-
Argelaguet et al.	-	-	-
Salamin et al. 2008	Physical & Active	Real Walking	High
Salamin et al. 2010	Physical & Active	Real Walking	High
Debarba et al	-	-	-
Maselli & Slater	-	-	-
Monteiro et al	Physical & Active	Real Walking	High
Gorisse et al.	Physical & Active	Real Walking	High

cases include inspection investigations in large VEs, where effective locomotion requires additional DOFs in order to reach remote points in the virtual environment above ground. Since flying is not natural to humans, controlling direction, translation and speed of movement can be trying tasks. A way of flying in the scene requires complex contraptions to emulate flying machines, such as a paraglider [102], a space ship[2] or a zero-gravity simulator [44] to indicate direction of movement using their bodies, while controlling speed with additional buttons. Birdly [84], on the other hand, uses a complex setup to enable users to fly by flapping mechanical wings, that can also control speed of movement. While these contraptions provide efficient ways to flying in VEs, in more intricate scenarios, when reaching a certain location, users need to further interact with the VE, either by selecting, manipulating or creating content. There, these setups restrain users' actions to one task and thus are hardly suited to richer interaction contexts.

Techniques with a higher degree of interaction fidelity have been proposed to overcome this conflict. However, most studies are focused on the direction-indication stage. Notably, work by Chen et al. [20], showed that users perform better with a physical technique (gaze-steering) as compared to a virtual technique to indicate direction in a flying scenario. On a similar vein, users can specify direction by using their bodies either standing standing [112] or sitting down [6]. The Point-and-Fly technique use users hands as means to control direction and the horizontal distance between head and hand to determine speed of translation [33]. Although flexible for flying tasks, the evaluation conducted only considered floor-constrained travel.

Since travel is a form of manipulation, a viable way of controlling multiple DOFs is by using the DOF-separation strategy, which is mainly used on 3D object manipulation to improve precision. A first attempt to this was the ChairIO [6], Which is a device consisting of a stool where users can control rotation on the horizontal plane by rotating the stool and the chair leaning was used to control speed and movement in the lean direction.Additional pressure sensors provided limited movement in both directions in the Z-axis [6]. Work by Bowman et al. [13] compared both hand and gaze-oriented steering in a 6DoF translation environment, and found similar performance in both cases, although the results were still preliminary due to the lack of obstacles and the lack of user representation on the VE. A related approach by Sikström et al. also showed that physical techniques can improve sense of embodiment [41] in a flying scenario as compared to joystick control [95] when a virtual body is present. A leaning interface has also been proposed to fly in VR, but 25% of the participants quit the test during the training task due to severe cybersickness

[2]Icaros: Virtual Reality Fitness Experiences. Available in: `https://www.icaros.com/`

side-effects [115].

include [115],

Since flying is not natural to humans, we propose to isolate the components of travel (as defined by Bowman et al. [13]) in two phases, namely direction indication and speed control. By decoupling these phases, it is possible to isolate the unnatural part of flying, which denotes the direction indication phase, in which people need to control additional DOFs. This separation enables the use of higher interaction fidelity techniques, such as the WIP as a means to control speed of movement.

Paper	Task	Devices	
		Input	Output
McMahan et al. 2012	Floor-constrained Travel	Body Movements/Mouse & Keyboard	Monitor & CAVE
Cirio et al.	Floor-constrained Travel	Tablet	CAVE
Langbehn et al 2017	Floor-constrained Travel	Body Movements	HMD
Freitag et al 2014	Floor-constrained Travel	Body Movements & Joystick	HMD
Soares et al.	Flying	Specific Device	Wall Display
Icaros	Flying	Specific Device	HMD
Birdly	Flying	Specific Device	HMD
Chen et al.	Flying	Body Movements (Head)	Wall Display
CharIO	Game	Specific Device	Wall Display
Bowman et al 1997	Flying	Body Movements	HMD
Tong et al 2016	Flying	Body Movements	HMD
Sikström et al 2015	Flying	Body Movements & Joystick	Monitor
Wang et al 2012	Flying	Body Movements	HMD

Paper	Avatar Representation		
	Type	Graphical Fidelity	Perspective Fidelity
McMahan et al. 2012	Real Person	-	1PP
Cirio et al.	Real Person	-	1PP
Langbehn et al 2017	Avatar Representation	Low	1PP
Freitag et al 2014	Real Person	-	1PP
Soares et al.	Real Person	-	1PP
Icaros	No Avatar	Low	1PP
Birdly	No Avatar	Low	1PP
Chen et al.	Real Person	-	1PP
CharIO	Real Person	-	1PP
Bowman et al 1997	No Avatar	Low	1PP
Tong et al 2016	Real Person	-	1PP
Sikström et al 2015	Real Person	-	1PP
Wang et al 2012	No Avatar	-	1PP

Paper	Travel		
	Classification	Technique	Interaction Fidelity
McMahan et al. 2012	Virtual & Active/Physical & Active	Joystick/Virtual Circle	Low/High
Cirio et al.	(Virtual+Physical) & Active		High
Langbehn et al 2017	Physical & Active	Redirected-Walking/Flying	Moderate
Freitag et al 2014	Physical & Active	Body Gestures	Moderate
Soares et al.	Physical & Active	Device Steering	Moderate
Icaros	Virtual & Active	Device Steering	Moderate
Birdly	Physical & Active	Device Steering	Moderate
Chen et al.	Virtual & Active	Device Steering	Moderate
CharIO	Virtual & Active	Device Steering	Moderate
Bowman et al 1997	Physical & Active	Gaze Steering/Hand Steering	Moderate
Tong et al 2016	Physical & Active	Gestures	Moderate
Sikström et al 2015	Physical & Active	Gaze+Gestures	Moderate
Wang et al 2012	Physical & Active	Leaning-Based Interface	Moderate

3
Our Approach

Travel [13] is the act of moving from a point A to a point B in a given direction. In Virtual Environment this means translating the user from an initial to a target position. Travel techniques can be divided into two criteria [15]: how the movement is controlled (either virtual or physical) and how the action is performed (either active or passive). On virtual techniques the movement is controlled using additional equipment (e.g. joysticks and wands); also, on Physical Techniques the movement is controlled by users bodies. Regarding performance, on active techniques the actions are executed by the user and passive, by the system.

A related concept is Interaction Fidelity, which is defined by McMahan et al. [62] as "the objective degree of exactness with which real-world interactions can be reproduced". In Travel techniques, the similar the travel technique is to real-walking, the higher the level of interaction fidelity it has.Physical techniques have a greater level of interaction fidelity, since the users' own bodies are used to indicate direction [62][14] and control speed of movement [18][26].

Quality factors of a travel technique include ease of use, ease of learning, increased spatial awareness, presence and efficacy [13]. Previous studies pointed the relation between interaction fidelity and quality factors [62][108] which found that the level of fidelity does not directly affects those factors. However, this relation was only observed in close-to-real scenarios, limited by a ground plane.But the inclusion of additional degrees of freedom could make the control of travel not such an easy task.

In some cases traditional travel metaphors are not sufficient for users to reach remote places on the virtual environments that reside above grounds. For that, users need to have control of additional degrees-of-freedom to enable such positions.This is normally done with the use of complex equipment that emulates machines such as a paraglider or virtual wings. However, these equipment put users in uncomfortable positions to perform further actions, such as virtual objects selection and manipulation. A way of overcoming this problem is by using parts of the body such as the head (gaze-based) [68] or hand (hand-based steering) [74] to indicate direction of movement [13]. Preliminary work by Chen et al. [20] show that the use of techniques with higher level of fidelity such as gaze-based steering are more efficient than joystick-based virtual techniques. Still, controlled Flying in virtual environments remains a difficult task, posing challenging problems. In this work we explore the "Magic Carpet" concept, depicted in Figure 3.3, a space where users can use close-to-real-walk metaphors on both phases of the travel pipeline for flying tasks. The "Magic Carpet" acts as an informative proxy of the real physical ground matching its position and rotation in order to mitigate balance issues, cybersickness and fear of heights. This space is rendered in the virtual scene below users' feet.We employed a fully-embodied representation for improved awareness inside the "Magic Carpet", since both studies were carried out using a full immersive VR using an HMD,

Fully-embodied avatars are also an important part of the travel user experience, particularly when using Head-Mounted display (HMD)s, where the equipment completely occludes the user's real self [63]. The use of an avatar is even more important when utilizing a travel technique with a high level of interaction fidelity, since it uses parts of the body as input, which need immediate body visual feedback.The VR experience when using an embodied representation depends on how this representation is presented and how it is viewed. We define this concept as "representation fidelity", which is the level of exactness of how the virtual body is similar to people's bodies in real-life (here defined as graphical fidelity) and the level of exactness in which this representation is viewed (here defined as "perspective fidelity).Works on the literature are normally focused on the graphical fidelity part [57, 56] when using a high-perspective fidelity (the 1PP) and indicate that the level of graphical fidelity does not directly affects the feeling embodiment of a virtual body. However, these works are limited to a more general scenario and it is not yet seen the effects of this particular components on quality factors of a travel scenario. The use of a

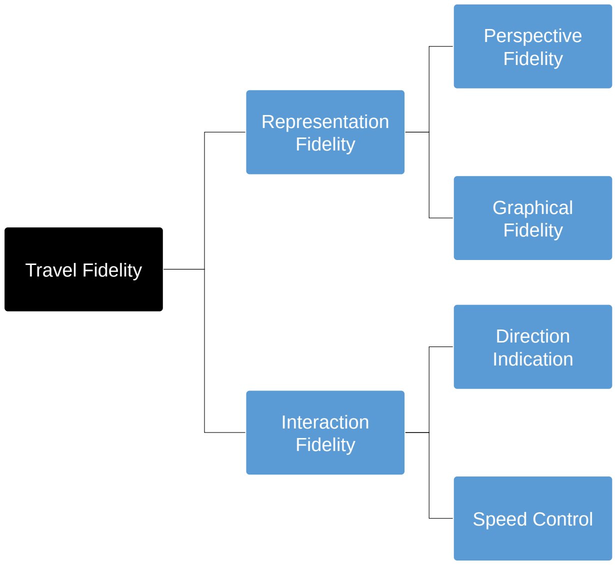

Figure 3.1: Components of Travel Fidelity : Interaction Fidelity and Representation Fidelity

low-perspective fidelity representation leads to a seamless improvement in spatial awareness both in augmented reality [89] and virtual reality [37, 72] head-mounted display setups, but still compromise efficiency of the task being performed. Also, since people's bodies are always seen when a low-perspective fidelity representation is used, we argue that the level of graphical fidelity has an improved impact over embodiment and travel quality factors.

In order to investigate both representation and interaction fidelity factors, our approach consists in subdividing the travel fidelity in two main parts. For this, we propose to encompass each of the factors in a separate study. To isolate the representation fidelity factor, we propose to use the highest interaction fidelity technique possible, the real walking. Then, to study interaction fidelity in flying scenarios we explore a design space where people can use techniques with increasing levels of interaction fidelity, the Magic Carpet Metaphor. In the following subsections we explain in further detail the concepts needed and give insight of each of the proposed studies.

3.1. Representation Fidelity

Currently, the majority of approaches to self-representation in IVEs is done by using avatars on a First-Person perspective. Also, some are the works that relate graphical realism to sense of embodiment when an user is viewed in the First-Person perspective (also known as uncanny valley). A further development in this matter is the use of realistic avatars that reconstruct the real body of the user and maps into his virtual self. The nature of the representation and its body movements mapping could improve the sense of embodiment of an user. During navigation, the level of graphical fidelity does not influence the feeling of embodiment, despite the level of interaction fidelity of the travel technique used, when a First-Person perspective is used. This is explained because since this is a natural perspective, users' rely more on their judgments and the graphical fidelity does not influence these factors in travel tasks. Although the use of first-person avatars influence on embodiment is object of great study on the literature, the use of third-person avatars is not still fully understood in Immersive Virtual Environments (IVEs). These types of avatars are commonly used in conventional monitors to improve user's spatial awareness of the involving scene. A Third-Person Perspective of the user can aid the user to expand their view of their surroundings while not compromising his sense of embodiment and the overall VR experience. The use of this perspective also does not improve spatial their spatial awareness between them and their surroundings. Also, since it is not how people see their bodies in real-life, the use of this perspective decrease travel efficiency. Additionally, since users' bodies are always viewed in a Third-Person Perspective, we argue that the level of graphical fidelity has more impact on sense of embodiment in this perspective.

To isolate the representation fidelity component, we proposed a study that enables people to use the highest interaction fidelity technique for travel: the real walking metaphor. In this study we assess how representation fidelity impacts the further study the effects of perspective (1PP and 3PP) and realism of the representation of self-embodied avatars in the sense of embodiment of users in VR setups. To this end, we use three different representations with varying level of realism (graphical fidelity) of each representation, ranging from an abstract, a realistic humanoid and a real-time point-cloud representation. The abstract representation uses spheres and boxes to represent parts of the body. The second is a realistic mesh avatar that is rigged and deformed according to tracking information. The third representation is a low cost point-cloud based avatar, using extracted video information from a person's

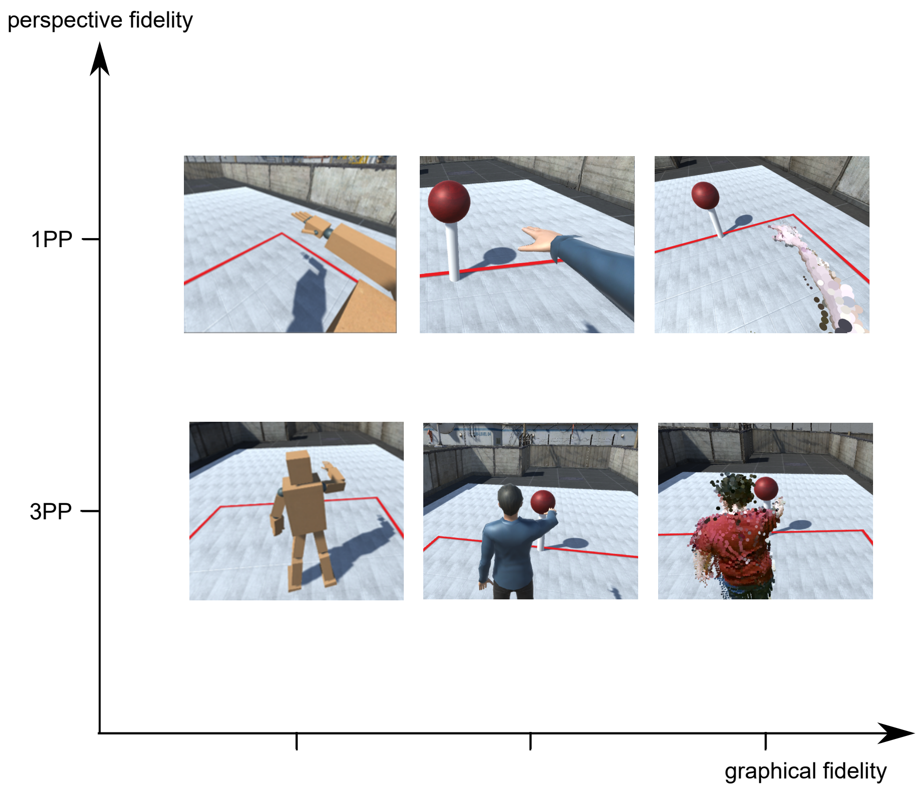

Figure 3.2: Representation Fidelity and its sub-components : Graphical (x-axis) and Perspective Fidelity (y-axis). The x and y axis are an approximated visual representation.

real body that is mapped into the virtual environment. Studies indicate that the realism of the representation improve embodiment factors, but just slightly, when using a first-person avatar [55]. But, since users' bodies are always seen when using a 3PP avatar, we argue that the realism of the representation highly impact both spatial awareness and embodiment factors. We use navigation tasks to assess both spatial awareness and embodiment factors, where users are asked to walk physically while avoiding obstacles in a VE. The obstacles were positioned around the user, near the feet of the users and at the head level.

3.2. Interaction Fidelity

Other important parameter in the immersive navigation experience is the level of interaction fidelity of the travel technique used. Being the real walking the most effective method, constraints such as the physical limits of the room compromises the experience. In complex virtual environments, users need to be able to control more degrees of freedom to reach certain parts of the virtual scene. Normally, this is made with the use of complex equipment which put users in uncomfortable positions and make them unable to perform further actions. Since flying is not a natural metaphor used by humans, we subdivide this task in two phases and separate them in two phases : the direction indication and speed control. This decoupling enable to isolate the unnatural part of flying, which is the control of multiple degrees of freedom enable the use of techniques with higher interaction fidelity in both phases. On the speed control phase people can use high-interaction fidelity techniques and its use can cause an improvement in quality factors.

As a key contribution, we propose to improve the flying experience by introducing two different studies that isolate the components of travel, proposed by Bowman et al. [13] inside our Magic Carpet design space. A first study to investigate 1) the direction indication phase, where users need to specify the direction where they should go (using a pointing gesture, their hands or joysticks). And a second study focused on 2) the speed control phase, where users manage their speed in the previously indicated direction (using their bodies or specialized interactions). This separation enables techniques with high level of interaction fidelity in both phases of travel.For example, a steering method could be used for indicating direction while using the Walking In Place metaphor for speed control [18]. Therefore, we contribute the results from two separate studies comparing varying fidelity levels travel metaphors for direction indication and speed control. Being the objective of both studies the assessment of interaction fidelity for all metaphors.

3.3. Target-based techniques

Another form to reach remote sites in the VE is with the use of Target-based techniques. In these techniques the speed of translation is determined by the system. The translation can be immediate, such as the Infinity Velocity techniques, or gradual such as the Linear Motion Technique [15]. These techniques are known to reduce cybersickness effects while still providing a good spatial orientation inside the virtual scene [80]. In this type of technique The translation of the can be immediate or gradual, where the translation is made following a pre-determined path. However, using a Teleport technique people may lose the context of the surroundings during translation and may cause disorientation during translation. We argue that the use of transitions before and after the translation diminish the sense of disorientation while navigating. Also, we argue that Infinite Velocity techniques are still more efficient and cause less cybersickness than gradual techniques.

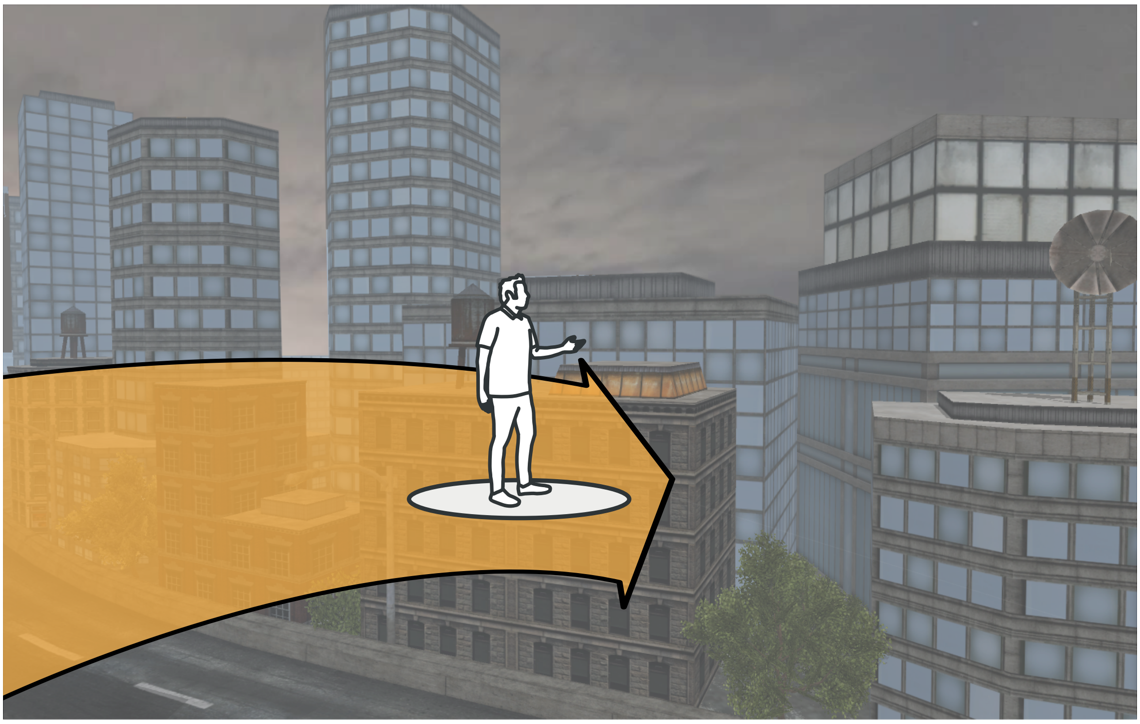

Figure 3.3: Vision of the explored design space ("Magic Carpet")

4

Assessing Representation Fidelity for Travel

In this chapter we further study perspective (1PP and 3PP) and graphical fidelity effects on the sense of embodiment of users in Virtual Reality setups when using a fully-embodied representation. To this end, we use three different representations varying the level of graphical fidelity of each representation, from an abstract to a realistic humanoid representation. The abstract representation uses spheres and cylinders to represent parts of the body. The second is a realistic mesh avatar that is rigged and deformed according to tracking information. The third representation is a low cost point-cloud based avatar, using extracted video information from a person's real body that is mapped into the Virtual Environment. To compare sense of embodiment, efficiency and ease of use for each representation we designed and evaluated performance on three different natural tasks based on previous work [89]. The selected tasks are navigation tasks, where users are asked to physically walk while avoiding obstacles in a Virtual Environment. In the following sections we describe the experiments, report on measures, discuss the results obtained and propose a set of guidelines for Self-Embodied VR applications.

4.1. User Study

In this work, we assess how the perspective and representation affects efficiency and efficacy in navigation tasks. We consider First- and Third-Person perspectives and, for representation, we utilize three different avatars with increasing levels of graphical fidelity. These range from a stylized box-avatar, a humanoid mesh avatar and a real-time point-cloud avatar, which use depth-cameras to map users' representations inside the virtual environment. All avatars have visuomotor synchronicity in order to provide a more realistic experience. To assess efficiency and efficacy we designed three different task that consist in walking while avoiding obstacles, which differ in how the obstacles are positioned in the virtual environment. The first consisted on avoiding obstacles that are positioned around the user; on the second task and third tasks, the objects need to make changes on the vertical plane to surpass them, by going over (Task 2) or below obstacles (Task 3). In our test, we chose a 2x3 within-subjects design with the perspective and avatar used as independent variables. The dependent variables used in our studies were 6-likert scale entries of the post-test questionnaire, task time, collision time and number of collisions. To maximize our tracking space and provide a more realistic experience, we chose to use a circular path in all three conditions.

In this section we describe the main aspects of designing the test experience regarding user representation and the design of the task. The following subsections present the task concept, the avatar representations used and the setup used on the test task.

4.1.1. User Representations

We chose three different user representations for our evaluation, from a low-level perspective fidelity method, the 3PP to a high-level perspective fidelity method, the 1PP.Camera positioning in 3PP is based on previous work by Kosch et al. [43], in which the camera is positioned above user's head for improved spatial awareness.

In all the used representations, the depth sensors' joints positions and rotations were mapped directly into the avatars using direct Kinematics. Skeleton tracking was performed using the "Creepy Tracker" toolkit from Sousa et al. [103]. This toolkit provides reliable markerless tracking using Kinect sensors, and ensures us to follow users in the area necessary for the study (4 meters by 4 meters). A surrounding bounding box to each joint was used as a basis for collision detection between the users and the obstacles on all tasks. The bounding boxes used for the collision are also used as the basis for rendering the abstract avatar and are used for collision purposes in all avatars.

4.1.1.1. Abstract

The first avatar was a simplified avatar representation which was composed by abstract components. Spheres were used to represent each joint, and boxes for each bone connecting joints and the head. These boxes were scaled according to the user and were also used on the other representations for collision detection. Figures 4.1-A and 4.1-B show this representation in both First and Third Person Perspectives (1PP and 3PP), respectively.

4.1.1.2. Mesh

The second representation is a realistic mesh avatar resembling a human being. This representation did not include animation for individual fingers, since they are not tracked by the Kinect sensor. Figures 4.1C and 4.1D show this representation in the First and Third Person Perspectives (1PP and 3PP), respectively.

4.1.1.3. Point Cloud

This body representation was based on a combination of separate streams of point clouds from Microsoft Kinect sensors, that were broadcast over the network by the "Creepy Tracker" toolkit [103]. These point clouds were in the same coordinate system as the skeleton information, which was also used for collision detection. When using the avatar in the 1PP, head information was discarded to avoid visual occlusion.

Figures 4.1-E and 4.1-F show this representation on the first and third-person views.

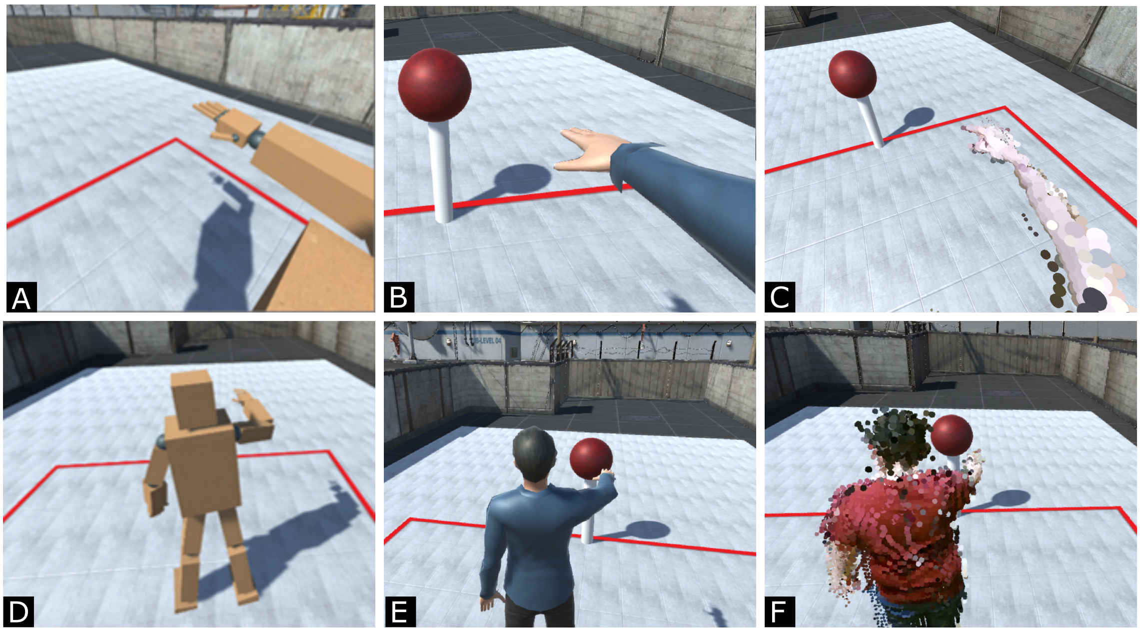

Figure 4.1: Self-representations used in our study. (A) 1PP Abstract Avatar (B) 3PP Abstract Avatar (C) 1PP Mesh Avatar (D) 3PP Mesh Avatar (E) 1PP Point-Cloud Avatar (F) 3PP Point-Cloud Avatar

4.1.2. Methodology

For assessing the effects of Representation and Perspective Fidelity, we used a 2x3 factor Within-Subjects Test Design. The test was divided into eight stages: 1) introduction to the study and application of a pre-test questionnaire; 2) explanation about the tasks and each of the users representations 3) adjustment of the device for comfort; 4) calibration procedure; 6) task execution; 7) application of post-test questionnaire; 8) and a semi-structured interview.

At first, we explained the test objectives. Then, the users completed a pre-use questionnaire to raise the participants profile regarding previous experience with related technologies (HMDs, virtual avatars, etc).

Subsequently, we showed a brief description of the tasks and representations used. Furthermore, we executed the calibration procedure. This procedure was performed to calibrate the tracking system between the HMD and the depth-sensors. Then, in order to familiarize the users with the procedures, they performed a task in a training scenario, where they could freely explore the VE and familiarize themselves with the setup and each of the representations.

After performing the training task, the users reached a fixed object in the environment and performed the test task. Then a questionnaire was given to the users to

Table 4.1: Questionnaires used in this study.

Question	
Q1	It felt like I was in control of the body I was seeing (Agency)
Q2	...that the virtual body was my own (Body Ownership)
Q3	...as if my body was located where I saw the virtual body to be (Self-Location)
Q4	... if I had more than one body
Q5	it was easy to walk in the virtual environment
Q6	it was easy to avoid obstacles in the virtual environment (Task 1)
Q7	it was easy to go over the obstacles in the virtual environment (Task 2)
Q8	it was easy to go under the obstacles in the virtual environment (Task 3)
Q9	I felt fatigue

gather information about their experience using each of the representations. These steps were done for each of the combination of the test conditions (perspective and representation) of a total of 12 permutations. We permuted the order of representations used and the order of perspectives, so if a user performed the order Abstract-Point-Cloud-Mesh in the 1PP he would do it at the same order when using the 3PP. The order of the avatar representation was changed in every test, following a Balanced Latin square arrangement, to avoid biased results. After performing each representation-perspective the users filled in a 6-Point Likert Scale Questionnaire (Table 4.1) to assess embodiment, easiness of completion of each of the tasks and fatigue issues.

4.1.3. Virtual Environment

The selected environment was based on the Stealth Scene, which was obtained on the Unity Asset Store[1]. This scene was modified to remove visual clutter, to not interfere with the goals of the test by capturing user's attention.

We also included in the environment a representation of the Kinect's tracking limits, which was marked with a red square, where the user could walk freely.

4.1.4. Tasks Description

In order to isolate different aspects of navigation tasks that we wanted to evaluate, the test was divided into three tasks. For each of the tasks, users would go through

[1] http://unity3d.com/store

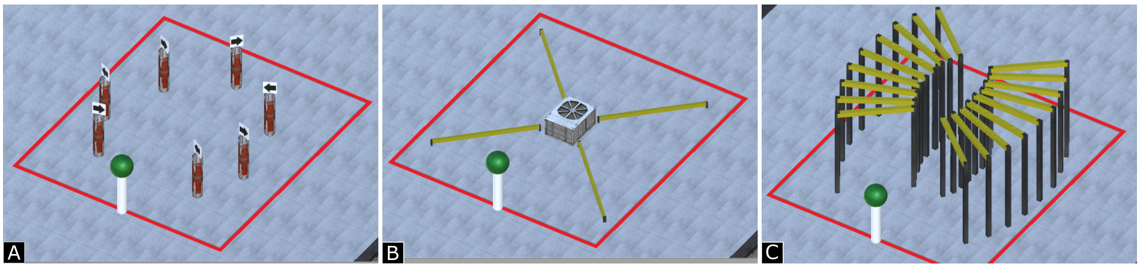

Figure 4.2: The proposed tasks for our evaluation. The green lollipop marks the initial position of the user.

the test until they reached the starting point again, marked by a green colored sphere, triggering the start of the next task. These were chosen based on natural tasks such as walking while avoiding obstacles based on previous work [89]. To maximize tracking space, we arranged the objects along a circular path, where the user walks following an anti-clockwise path until reaching the initial point.

In the following subsections we present and explain in further detail each of the proposed tasks.

4.1.4.1. Task 1 (Barrels Task)

In this task, users needed to go around the barrels as indicated by the signs on top of them . Figure 4.2A illustrates the first Task.

4.1.4.2. Task 2 (Bars Task)

In the second task, users needed to avoid each of the yellow bars by raising their feet (or jumping) until they reach the initial point (Figure 4.2B).

4.1.4.3. Task 3 (Tunnel Task)

In this test users needed to go under the two tunnels until they reached the initial point. This tunnel is adjusted according to the user's height, which was estimated using the distance between the head and the toe when the user started the test. The ceiling of the tunnel was placed 12 centimeters below the user's height (Figure 4.2C).

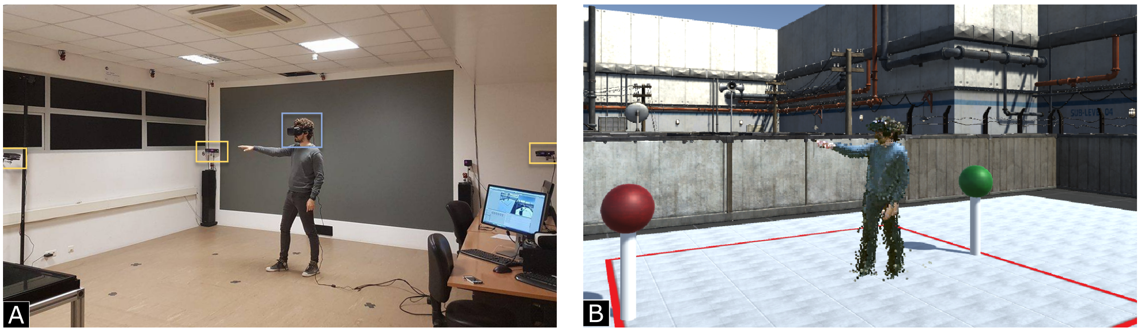

Figure 4.3: The setup used for our study. Figure A shows the laboratory and one user performing a test, and Figure B the virtual world mapping.

4.1.5. Setup

The physical setup chosen for our study can be seen on figure 4.3, where some of the Kinect sensors used for body tracking and point cloud reconstruction can be seen. A wide-baseline setup was used due to two main reasons; firstly the fact that the Microsoft Kinect sensor has a limit on its effective range (0.4m to 4.5m, with skeletons losing reliability starting on 2.5m), and in order to properly evaluate a navigation task, a bigger space was needed. When the user was at the limits of the sensors operating range, the quality of the experience would be compromised, so a wide-baseline setup guarantees the whole body of the user was always visible by at least one camera. Secondly, since a 3PP was presented as one interaction paradigm, the whole body of the participant had to be visible at all times in order to avoid holes in the representation. A narrow baseline or single sensor setup would capture just half of the participant's body, greatly compromising the experience.

Five Kinect sensors were fixated on the walls of the laboratory where the study was being held, covering an area of approximately 4 x 4 meters. The placement of the kinect sensors was chosen based in such a way that user's bodies are always visible.

4.2. Results

In this section, we present the main observations made during the tests as well as the difficulties and suggestions from users about the test task. To assess the difference between the three user embodied representation both in first and third-person perspectives, we collected both objective and subjective data, in the form of logs and inquiries respectively, during the evaluation sessions. For the continuous variable, i.e. time, we used Shapiro-Wilk test to assess data normality. Since all samples were normally distributed, we used the Two-way Repeated Measures ANOVA for finding main effects between the two variables used, namely perspective and representation. When found main effects, additional One-way Repeated Measures ANOVA for multiple comparisons, and the Paired-Samples T-Test test between two samples, to find statistically significant differences. When comparing more than two samples, we applied the Bonferroni correction. Presented significance values are corrected.

In the following subsections we present the analysis made based on the results of the questionnaires and log files data obtained during the test.

4.2.1. Subjective Responses

As a result of the Two-way Repeated Measures ANOVA we found statistical significance between perspectives on Embodiment factors - Agency ($F(1,26)= 7.499$, $p= 0.011$), Body Ownership ($F(1,26)= 4.489$, $p= 0.044$) and Self Location ($F(1,26) = 9.755$, $p=0.004$). Statistical significance was also verified for easiness of walking ($F(1,26)=17.827$, $p<0.001$), completing the Barrels Task ($F(1,26) = 0.549$, $p<0.001$), Bars Task ($F(1,26)=4.23$,$p=0.005$) and Tunnels task ($F(1,26)=65.768$, $p<0.001$).

We also found interaction between variables perspective and representation on sense of agency ($F(1.969,51.191)= 3.884$ $p=0.027$), sense of body-ownership ($F(1.558, 52)=7.839$, $p=0.001$) and sense of self-location ($F(1.972,51.272)=4.889$,$p=0.011$). Also, the feeling of having two bodies ($F(1.668,51.683)=6.896$ $p=0.002$), easiness of walking ($F(1.971,51.234)=4.086$ $p=0.014$) and completing the tunnels task ($F(1.925, 50.057)=12.826$ $p<0.001$). To further investigate this interaction, we made two different comparisons based on the data collected through the questionnaires, between representations on the same perspective and representations between perspectives.

4.2.1.1. Perspective

When comparing between representations in the 1PP, we found no statistical differences in any of the questions using a One-way ANOVA. The only two exceptions were found in Q4, the feeling of having more than one body ($F(1.909,49.622)=9.869$ $p<0.001$) and easiness of completing task 3 ($F(1.917, 49.835) = 3.503$ $p=0.04$) for

Table 4.2: Results from the questionnaires collected in the second experiment, presented as median (interquartile range) values.

	1PP			3PP		
	Abstract	Mesh	Point Cloud	Abstract	Mesh	Point Cloud
Q1	5(1)	5(2)	5(2)	5(2)	4(1)	5(2)
Q2	5(1)	5(2)	5(3)	4(3)	4(2)	5(2)
Q3	5(1)	5(2)	5(2)	4(4)	4(3)	5(3)
Q4	2(2)	2(2)	3(3)	3(4)	3(3)	3(3)
Q5	5(2)	5(1)	5(2)	3(2)	4(2)	4(2)
Q6	5(2)	5(1)	5(1)	4(2)	4(2)	4(2)
Q7	5(2)	5(1)	5(1)	4(3)	4(2)	4(2)
Q8	6(1)	6(2)	5(2)	3(2)	4(2)	4(2)
Q9	2(3)	2(3)	2(2)	2(3)	2(3)	2(3)

the 1PP. Post-hoc paired t-tests showed that users felt as if they had more than one body with the Point-Cloud avatar when comparing with both the Abstract ($t(26)=-0.811$ $p<0.001$) and Mesh ($t(26)=0.004$ $p=0.012$) avatars.

We found a greater number of statistically significant statements when comparing between representations in the 3PP. By running the One-way ANOVA we found statistical significance on the 3PP in agency ($F(1.771,46.033)=5.25$ $p=0.008$), body-ownership ($F(1.934, 45.531) = 9.314$ $p<0.001$), self-location ($F(2, 51.087)= 3.812$ $p=0.029$) and easiness of completing task 3 ($F(1.925, 49.835) = 3.503$ $p=0.001$).

With the results of the post-hoc tests we noticed that users attributed a higher sense of embodiment, specifically on the sense of agency to the Point-Cloud Avatar when comparing to the Abstract Avatar ($t(26)= -2.595$, $p= 0.045$) and when comparing the Point-Cloud with the Mesh Avatar ($t(26)=-2.672$ $p=0.039$). Statistical significance was also found on Sense of Body-Ownership, with Abstract statistically worse than the Point-Cloud ($t(26)=-3.798$ $p=0.003$); a higher sense of self-location was sensed with the Point-cloud in comparison with the Abstract avatar ($t(26)=-2.55$ $p=0.017$). Regarding task 3, we found statistical significance in the 3PP ($F(1.879,48.844)$ $p=0.001$), with users also finding easier to execute the Tunnel Task (Task 3) using the Abstract avatar when comparing to the Mesh ($t(26)=-3.349$ $p=0.006$) and Point-Cloud ($t(26)=-3.365$ $p=0.006$) avatars.

4.2.1.2. Representation

When comparing perspectives between the different representations we found overall better results with the 1PP on all representations. On the Abstract avatars, users felt a stronger sense of embodiment in the First-Person Perspective in all its components: agency ($t(26)=3.514$ $p=0.006$), body-ownership (q2) ($t(26)=3.776$ $p=0.003$) and self-location (q3) ($t(26)=2.848$ $p<0.001$). They also felt that they had more than one body in the 3PP ($t(26)=-2.926$ $p=0.021$), found it easier to walk in the VE ($t(26)=6.176$ $p<0.001$) and to perform the Barrels task($t(26)=5.827$ $p<0.001$) and Tunnels Task ($t(26)=7.963$ $p<0.001$).

Regarding Mesh Avatars, users felt a stronger sense of body-ownership ($t(26)=2.89$ $p=0.024$) and self-location ($t(26)=2.848$ $p=0.024$) with the First-person perspective and also less feeling of having two-bodies ($t(26)=-2.591$ $p=0.045$). Users also found it easier to walk with the 1PP ($t(26)=2.842$ $p=0.027$). About task easiness, they found it easier to avoid obstacles in the Barrels task ($t(26)=2.769$ $p=0.03$) and Tunnel task($t(26)=4.352$ $p<0.001$).

Lastly, on Point-cloud avatars, no difference was found regarding sense of embodiment and its sub-components. About task easiness, users only found it easier to perform the Tunnel task with the First-Person perspective ($t(26)=2.69$ $p=0.036$).

4.2.2. Task performance

In this subsection we present the analysis of results collected from users during the evaluation session. For assessing task performance of the users between the different representations we collected data through logs. We counted the time to assess the efficiency of the representation, the number of obstacles hit and the collision time to evaluate spatial awareness. Figures 4.4 and 4.5 show the total and collision time for each task in both perspectives and representations, respectively. The number of obstacles hit can be found on Table 4.3.

In the following sub-sections we present the results obtained for each of the metrics used (time, number of obstacles hit and collision time) for each of the sub-tasks.

Table 4.3: Obstacles hit per task. Median number of obstacles hit (inter-quartile range).

	1PP			3PP		
	Abstract	Mesh	Point Cloud	Abstract	Mesh	Point Cloud
Task 1	4(2)	4(2)	4(2)	5(1)	5.5(2)	5(2)
Task 2	4(0)	4(1)	4(2)	4(1)	4(0)	4(2)
Task 3	5(7)	5(9)	10.5(12)	12(6)	12(8)	16(7)
Task 3 (Just Bars)	0(1)	0(2)	0.5(3)	2(8)	1(3)	7(16)

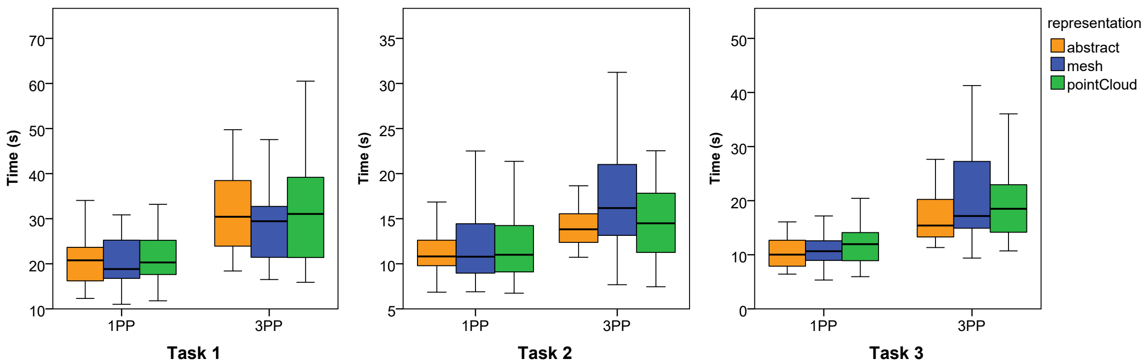

Figure 4.4: Performance time of Avatars in First-Person perspective and Third-Person Perspective divided by task. median, first and third interquartile ranges (boxes) and 95% confidence interval (whiskers). Orange represents the Abstract avatar, Blue the Realistic Mesh Avatar and Green, the Point-Cloud Avatar.

4.2.2.1. Barrels Task

Number of collisions

We found statistical significance on Barrels Task regarding number of objects collided on the Perspective factor($F(1,23)=24.636$ $p<0.001$, with the 1PP having a smaller number of objects hit. This behaviour was observed both with the Abstract ($t(23)=-2.497$ $p=0.06$) and Mesh ($t(23)=-3.657$ $p=0.009$) Avatars.

Collision time

When running a two-way repeated measures ANOVA, we found statistical significance on the perspective factor ($F(1,23)=26.592$ $p<0.001$), with better results on

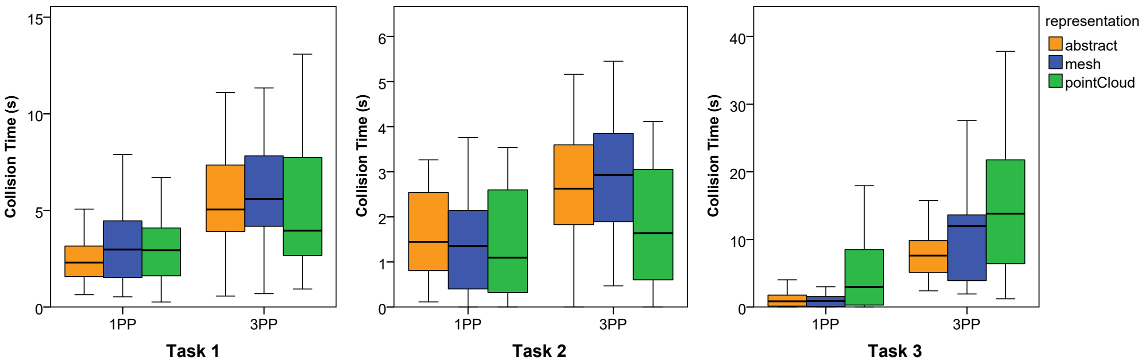

Figure 4.5: Total Collision time of Avatars in First-Person Perspective (1PP) and Third-Person Perspective (3PP) divided by task. median, first and third interquartile ranges (boxes) and 95% confidence interval (whiskers). Orange represents the Abstract avatar, Blue the Realistic Mesh Avatar and Green, the Point-Cloud Avatar.

the 1PP in all representations (Abstract: t(26)=-4.201 p<0.001; Mesh:t(23)=-4.35 p<0.001; Point-cloud: t(23)=-3.6 p=0.002).

Completion time

We only found statistical significance in the perspective factor (F(1,20)=76.686 p<0.001), with the 1PP being more efficient than 3PP in all cases (Abstract: t(21)=-7.818 p<0.001; Mesh: t(21)=-6.555 p<0.001; Point-cloud: t(23)=-6.336 p<0.001).

4.2.2.2. Bars Task

Number of collisions

We did not find any statistical significance for the number of collision in the Bars Task.

Collision time

A two-way ANOVA pointed statistical significance in both representation (F(1.879, 41.34)= 3.456 p=0.04) and perspective (F(1,22)=17.574) p<0.001 factors, but with no interaction between variables. When grouping representations by perspective, we found statistical significance on the 3rd Person Perspective (F(1.655,38.072)=3.7 p=0.042, post-hoc tests indicated less collision time with the Point-Cloud representation (t(23)=3.022 p=0.018). Comparing the perspectives in each representation, we found better results in the 1PP in the Point-Cloud avatar (t(23)=-3F.136 p=0.015).

Completion time

The First-Person Perspective was also the most efficient on this Task (F(1,16)=49.364 p<0.001), in all cases (Abstract: t(16)=-5.898 p<0.001; Mesh:t(22)=-4.260 p<0.001; Point-cloud: t(22)=-3.779 p=0.003).

4.2.2.3. Tunnel Task

Number of collisions

For this task, we considered two possibilities: the number of objects collided and the number of horizontal bars collided.

Regarding number of objects collided, we found statistical significance on both representation (F(2,43.077)=7.832 p=0.001) and Perspective (F(1,23)=39.606 p<0.001). We also found statistical significance between representations on the First-Person perspective (F(2,22)=7.150 p=0.004). Post-hoc test indicated that fewer objects collided with the Abstract in comparison with the Mesh Avatar (t(23)=-3.761 p=0.003).

When considering just the collision with the tunnels, we found statistical significance on representation (F(2,46)=15.858 p<0.001), perspective (F(1,23)=16.935 p<0.001) and interaction between factors (F(2,46)=4.591 p=0.015). Comparing between representations on both perspectives we found statistical significance on both 1PP (F(2,46)=6.124 p=0.012) and 3PP(2,46)=12.306 p<0.001. In the 1PP, users collided less with the tunnels using the Abstract avatar in comparison with the Point-Cloud avatar (t(23)=-2.802 p=0.03). In the 3PP, the Mesh had better results in comparison with both Abstract (t(23)=-2.890 p=0.024) and Point-Cloud Avatars (t(23)=-4.831 p<0.001). The comparison between the different representations in each perspective showed less collisions in the 3PP both on Abstract (t(23)=-3.744 p=0.003) and Point-cloud avatars.

Collision time

When running a two-way repeated measures ANOVA we found statistical significance in the perspective factor in favor of the 1PP (F(1,23)=36.841 p<0.001).

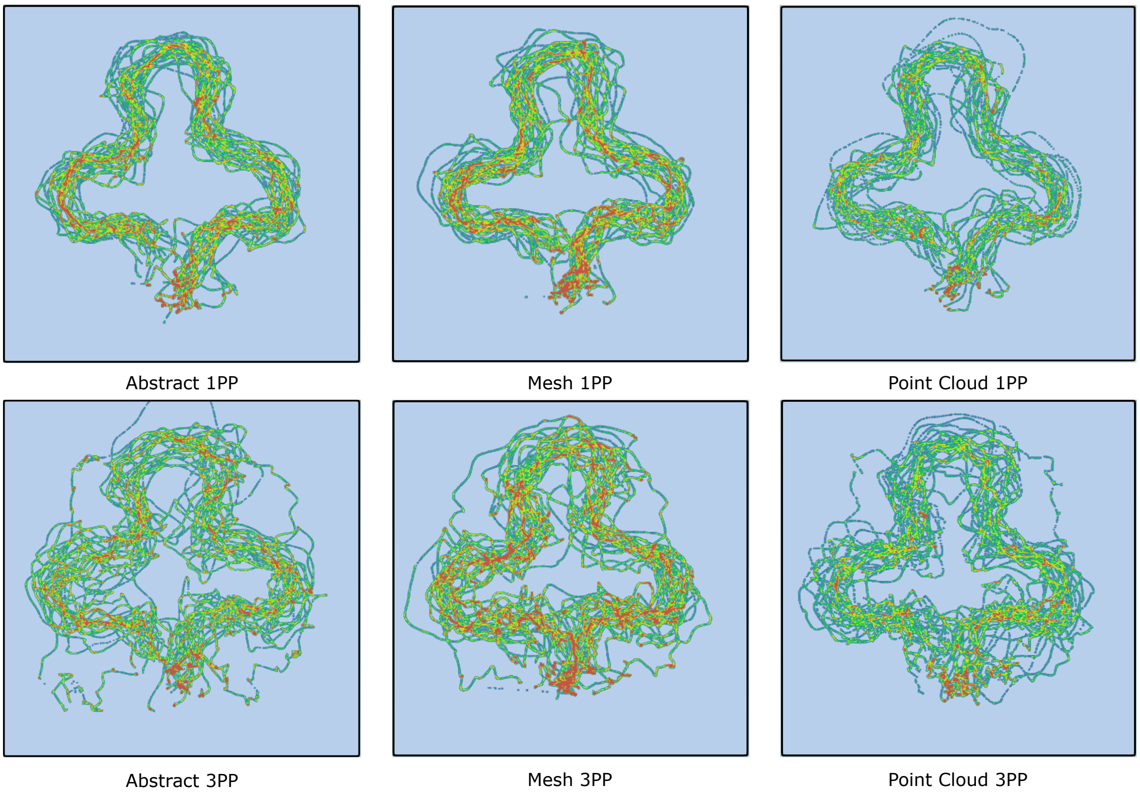

Figure 4.6: Heatmaps representing of users' paths for Task 1 separated by representation and perspective.

Post-hoc tests indicated better results in all representations (Abstract: t(23)=-4.333 p<0.001; Mesh: t(23)=-3.858 p=0.003;t(23)=-4.871 p<0.001).

Completion time

The First-Person Perspective was also the most efficient on this Task (F(1,16)=49.364 p<0.001), in all cases (Abstract: t(23)=-4.856 p<0.001; Mesh:t(23)=-6.305 p<0.001; Point-cloud: t(23)=-7.563 p<0.001).

4.3. Discussion

From an overall analysis of the results, we can verify that the 1PP was found to be more suited for travel tasks. The performance results were significantly better for all representations (time, collision, collision time), showing that it not only allowed users to perform the tasks faster, but with higher precision. Users also felt a higher sense of embodiment in this perspective, and felt it was easier to complete the tasks when compared to the 3PP. These observations can be easily explained by the fact that it is a more natural point of view to which they are used to, while self-location in the 3PP was found to be significantly harder.

However, this confirmation does not correspond to earlier work. Debarba et al. [29] have reported a similar sense of embodiment in the 3PP when compared to the 1PP. In their work, the majority of the interaction time is limited to reaching tasks. The navigation phase of the test is limited to reach the test area. In the remaining time, the majority of users' bodies remain stationary. This may explain the difference in embodiment factors. In our study, users stay in movement most of the time, so the relationship between people and the environment is always in motion. We verified that this aspect affects embodiment and all of its components, particularly in the sense of self-location.

Earlier work also suggested an improvement in spatial awareness with the use of third-person avatars [37, 72, 89]. Some differences in these papers explain the different results from our work. In both Gorisse et al. [37] and Monteiro et al. [72] the third-person was used to expand user's view and be able to see further parts of the virtual environment. For example, in Monteiro et al. [72] the task consisted in controlling a vehicle and in the 3PP people were able to see further details in the road and respond faster when further actions such as turning were needed. Gorisse et al. [37] on the other hand, base the improvement in users' spatial awareness with the reaction of moving objects being thrown towards the user, something that was already confirmed by Salamin et al. [89]. For the navigation phase, the study proposed by Gorisse et al. [37] used solely subjective metrics to assess the spatial presence. Another point to be considered is that the obstacles presented in the VE were only used to affect the performance. One example can be seen in the video provided where users were seen stepping out of the limits of the VE, something that was not analyzed in that study. In our case, our study focus on the relations between users and the VE. This is highly influenced by how an user make distance (or spatial) judgments, which indeed affect the feeling of spatial awareness. This difference can clearly be seen in Figure 4.6, where the paths taken in the 1PP for the first task are more fluid to the expected trajectory to avoid the proposed obstacles. Also, in this perspective we found no influence on the graphical fidelity on both efficiency and spatial awareness.

As seen, the use of an avatar in the 1PP is indeed the most efficient and effective in navigation tasks.The only exception was found when the Point-Cloud avatar was used. Since this representation uses a reconstruction of users' bodies inside the VE, they feel an equivalent sense of embodiment with both 3PP and 1PP.Also, no statistical significance was found between representations in the 1PP.This equivalence was also found in spatial awareness factors (collision number and collision time). This representation had a smaller collision time in the second task when compared to its alternatives in the 3PP. About travel efficiency we did not find statistical difference between representations. By analyzing users' path we also noticed a more fluid path with the Point-cloud avatar in the 3PP when compared to the other representations, which indicates an improved perception of their surroundings and how they make spatial judgments (Figure 4.6). This fact was more noticeable with the the representation with higher level of graphical fidelity, the Point-Cloud.Also, when comparing between the representations in the 3PP we also found a higher feeling of embodiment in all of three sub-components of the sense of embodiment: agency, body-ownership and self-location..

Some particularities although were found in Task 3, regarding spatial awareness and task easiness. On the Tunnels task, we noticed a higher amount of collided time and objects collided with the Point-cloud (average=10.46s) in comparison with the abstract (average=5s) and mesh (avg=7.67s). This may indicate a higher effect on the perceived distance compression provided by the HMD, related to users' height. With a highly detailed avatar, people's tend to make distance judgments in a similar way to how they make in real-life and due to the perceived compressed space, more errors may occur, particularly in the z-axis.

5 Assessing Interaction Fidelity for Flying in VR

Natural locomotion is not always suitable, and flying in VEs is not natural to humans. Furthermore, the increase in the number of DOFs creates more problems than it solves. Because flying is not natural to humans, we propose to isolate the components of travel into two phases, namely, direction indication and speed control. By decoupling these two phases, we enable the isolation of the most unnatural aspect of travel in this metaphor, namely, the control of the additional DOFs, and the use of techniques with a higher level of interaction fidelity for speed control.

To this end, we present a new interpretation of the *Magic Carpet* metaphor that combines a fully embodied user representation with a virtual floor proxy to improve travel quality effects [13], spatial awareness and the sense of embodiment and to prevent side effects such as cybersickness, fear of heights, and imbalance issues. In each phase of travel, we use continuous input control, where the start and end of movement are specific to each technique, as described below.

We performed two user studies, one for each phase, to ascertain the most suitable combination of methods. In the first study, to choose the best-suited technique for indicating direction, we evaluated three different techniques. The first technique tested for this phase was the novel technique “Elevator+Steering”, which uses the DOF separation strategy. This is a common way to improve the accuracy of 3D object manipulation [67]. In accordance with this technique, the control of the DOFs was decoupled into a horizontal translation, based on the projection of the user’s gaze onto the horizontal plane, and movement along the Z-axis, based on the concurrent use of additional buttons. The second technique was a gaze technique, in which the indicated direction is controlled by the user’s gaze, and the third was a hand technique, in which the user’s hands are used to indicate where to go. In contrast to the work of Bowman et al. [13], the presence of a full-body representation and obstacles in the scene enabled an in-depth investigation of various travel quality factors, namely, efficiency, cybersickness and, most importantly, spatial awareness.In the hand technique, the user indicates direction with his or her hands while still being able to use his or her head to inspect other parts of the VE. During the execution of the second study, to assess speed control, we used three different techniques with varying levels of interaction fidelity. In order of increasing interaction fidelity, the tested techniques for speed control were a joystick-based technique, the speed circle technique – a novel technique for controlling speed based on previous work [62][26] – and the WIP technique [18]. Both the speed circle and WIP techniques can be regarded as high-interaction-fidelity techniques. Because of the number of tested techniques, we employed the best-performing direction indication technique identified in the first study in conjunction with the tested techniques for speed control.

In the following sections, we present the common test design for both studies, followed by descriptions of the techniques implemented for each trial and by a detailed analysis and discussion of the elicited results.

5.1. Study 1 : Direction Indication

In this experiment we tested three different techniques regarding direction indication in a flying task.

In the following subsections, we present the techniques implemented for direction indication and the results obtained in terms of qualitative and quantitative metrics, followed by a detailed discussion. The qualitative metrics were obtained based on questionnaires issued to evaluate user preferences and comfort issues. Additionally, the questionnaires evaluated the sense of embodiment and its subcomponents, including the sense of agency (the feeling of having control of the virtual body), sense of body ownership (the feeling that the virtual body is their real body), and the sense of self-location (the feeling that the virtual body was located at the same place as the real body) [41]. To evaluate task performance, we gathered data based on logs and collected the total task time, total collision time, and user path length. In summary, the dependent variables used in this test were the previously described total task time, 6-likert scale responses (for subjective metrics), cybersickness questionnaire, collision time and path length and the independent variable the direction indication technique.

We recruited 18 participants for this study, two of which were female. The ages of the users varied from 21 to 35 years (average = 25, standard deviation = 3.5). Regarding experience, the majority of the users had previous experience with 3D applications such as games and modelling systems. The majority of them had previous experience with Head-Mounted Displays (83.3% or 15 participants) and 77% (14 participants) with previous experience with Kinect usage.

In the following subsections we present the techniques implemented for direction indication, present the participants involved, the results obtained regarding qualitative and quantitative metrics. The qualitative metrics were obtained through questionnaires to evaluate user preferences and comfort issues. To evaluate task performance we gathered data through logs, the metrics used for this matter were: total time of task, total collision time and path length.

5.1.1. Setup

As a visualization platform we used the Oculus Rift DK2 HMD.We also used the Sony Playstation Move Navigation Controller joystick to enable the user to trigger navigation through the scene using buttons and to control the speed of navigation using the analog stick.

To track users' movements we used the Creepy Tracker toolkit [104], using five Kinects which are connected wirelessly to a central application.This central hub was then responsible for merging users' body positions into a shared coordinate system and sending information to the client application. To minimize the effects of network communication, both the central hub application and client applications were running in the same desktop computer.

The five sensors were fixed on the walls of the laboratory where the study was being held and covered an area of approximately 4 x 4 m^2. We chose a subregion of the tracking area of each Kinect with a size of 1.2 x 1.2 m^2 to ensure more reliable tracking outcomes and arranged the Kinects using a wide-baseline arrangement.

The five sensors were fixed on the walls of the laboratory where the study was being held, covering an area of approximately 4 x 4 m. We chose a subset of the tracking space of the Kinect (2.5 x 2.5 m) to avoid tracking problems and arranged the Kinects such was that the participant was always facing a sensor so their his or her body was always visible.

Due to the Kinect's limitations in terms of tracking the hands and head, we used 10 Flex3 OptiTrack cameras, placed on the ceiling and operating at 100 FPS. The Creepy Tracker had an average delay of 76 milliseconds relative to the OptiTrack system due to the Kinect's limitations. Although this latency might hinder real-time performance in VR without some mitigating factor, combining the tracker's positional data with the orientation provided by current HMDs appears sufficient to satisfy the illusion of being present in a VE. Additional markers were placed on the Playstation Move to enable tracking of the hand rotation with 3 Degrees of Freedom.

5.1.2. User Representation

The user was mapped onto an abstract humanoid avatar (Figure 5.1A). This representation was chosen for both male and female participants. The user's joint positions and rotations as obtained from the Kinect sensor were mapped directly onto the avatar using direct kinematics. Because of the Kinect's hand tracking limitations, we attached a reflexive rigid body to the control wand to enable rotation tracking using the OptiTrack optical system[1]. The collected data were then mapped onto the avatar's hand to control rotation of the dominant hand of the participant.

To provide a comfortable experience and to avoid evoking a fear of heights, in all techniques, we positioned a plane below the user's feet. This plane represented a subarea of the available tracking space measuring 3 x 3 meters (Figure 5.1B), where the participant could walk freely. The orientation of this plane was fixed and always the same as that of the real floor, perpendicular to the user's body.

[1]https://www.optitrack.com

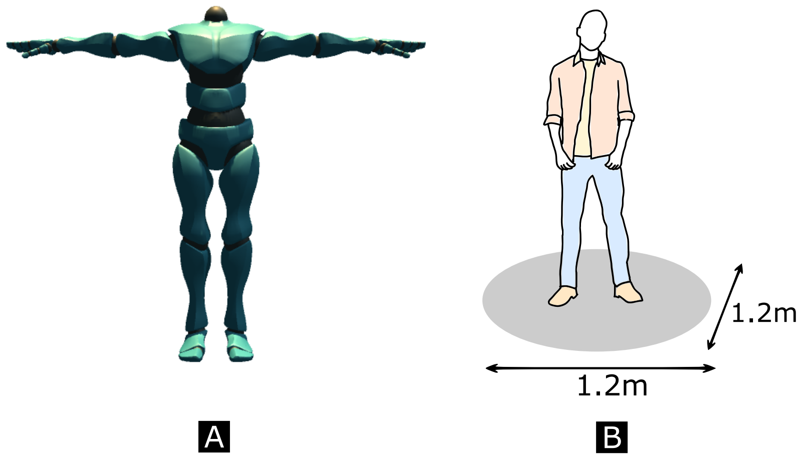

Figure 5.1: User representation: A) Avatar and B) Magic Carpet.

5.1.3. Virtual Environment

The selected environment was based on a city scene obtained from the Unity Asset Store[2]. This scene was modified to remove visual clutter that might otherwise distract the user's attention and thus interfere with the goals of the test.

Figure 5.2: Virtual Environment used on the experiments with a ring

5.1.4. Task Description

During this experiment, the user was asked to fly through rings, indicating the direction of movement using one of the proposed techniques. To guide users through the scene, we positioned rings to indicate the desired path. Only one ring was shown at a time, and once crossed – successfully or not – that ring disappeared, and a

[2]http://unity3d.com/store

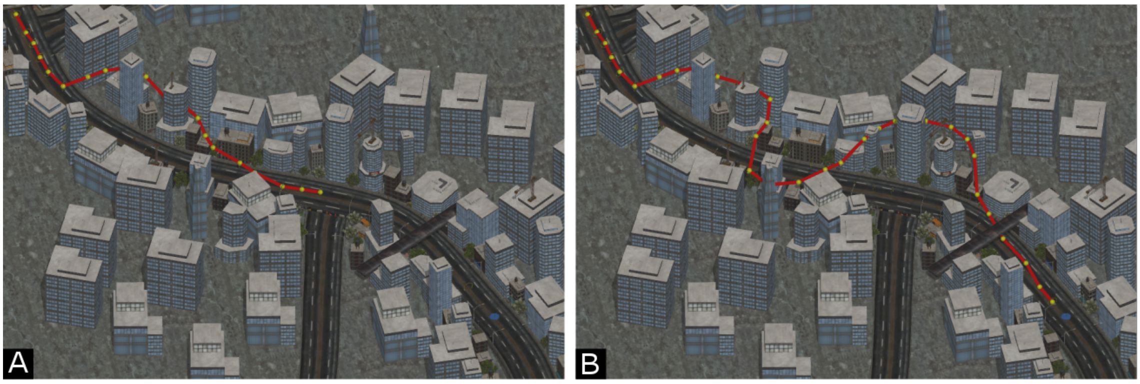

Figure 5.3: User tasks in the virtual environment: (A) Path used in the first experiment, with 22 rings and a total length of 180 m. (B) Path used in the second experiment, with 34 rings and a total length of 350 m. The red line indicates the path for each experiment, and the yellow dots indicate the positions of the rings.

new ring appeared (Figure 5.3). If the new ring was not visible to the user, an arrow was shown in the middle of the screen to indicate the next ring's position. This test consisted of 22 rings along a path with a length of 180 m and included abrupt changes in the Z-coordinate to best evaluate the users' attention and the effectiveness of direction indication for each of the tested techniques (Figure 5.3B). Once the direction had been chosen, the user pressed a button on the control wand and was then translated in the chosen direction at a constant speed of 3 m/s. The user could also modify his or her direction while traveling. To stop moving, the user needed to release the trigger button.

5.1.5. Implemented Techniques

Two different techniques were implemented that differ in how the user uses his or her body to indicate the direction of movement with his or her head or hands, as described in previous studies [13][33]. This work extends the work of Bowman et al. [12] by placing obstacles in the virtual scene and by using a fully embodied representation to improve the spatial awareness of the user. In addition, we developed the novel technique referred to as Elevator+Steering, which uses the DOF separation strategy commonly employed in 3D object manipulation [67]. Because travel is a form of manipulation, this approach can facilitate the control of additional DOF that is necessary when flying in a virtual scene.

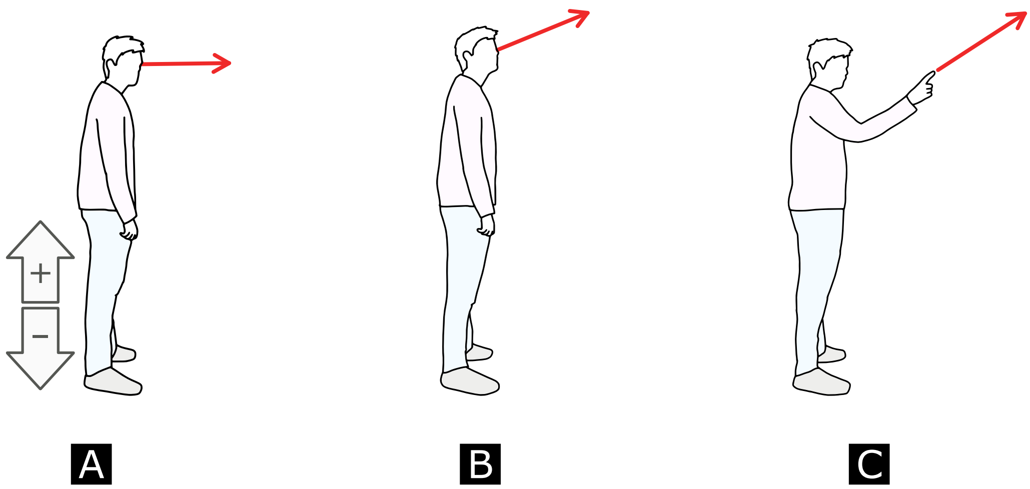

Figure 5.4: Direction Indication techniques implemented : (a) Hand Technique (b) Gaze Technique (c) Elevator+Steering Technique

5.1.5.1. Elevator+Steering

In this technique, the direction of movement is based on the projection of the participant's gaze orientation onto the horizontal plane. Additional buttons control the direction of travel in the vertical plane (Figure 5.4A) during travel. Another additional button is used to trigger movement.

5.1.5.2. Gaze-Oriented

In this technique, the direction of movement is based on the rotation of the head of the participant (Figure 5.4B). An additional button is used to trigger movement.

5.1.5.3. Hand Steering

In this technique, the direction of movement is based on the orientation of the dominant hand of the participant (Figure 5.4C). An additional button is used to trigger movement.

5.1.6. Methodology

Each test followed a within-subjects test design and was divided into eight stages: 1) introduction to the study and administration of a pretest questionnaire, 2) explanation of the task and each of the techniques, 3) adjustment of the device for comfort, 4) calibration procedure, 6) task execution, 7) administration of a post-test questionnaire, and 8) a semistructured interview.

First, we presented the participants with short descriptions of the tasks and the techniques used. To collect user profiles, we asked the participants to fill out a pretest questionnaire about their backgrounds and experience with navigation methods in VR settings.

Next, we presented the users with the calibration task. In the calibration procedure, each user was positioned at a fixed point in our laboratory and was asked to raise his or her hand. This procedure was performed to calibrate the tracking system between the HMD and depth sensors, thus associating the user's movements with the virtual avatar. Then, to familiarize the user with the procedures, each user performed a training scenario in which he or she could freely explore the VE and familiarize him- or herself with the setup. This training scenario was presented before the testing of each of the techniques, and no time limit was imposed.

After performing the training task, each user performed the calibration procedure and then the test task. Then, a questionnaire asking about several user experience issues was given to the user. To assess user preferences, fatigue and the sense of embodiment, we asked each user to fill out a post-test questionnaire. Additionally, before the test and after the execution of each of the techniques, we asked each user to fill out the SSQ to assess cybersickness issues.

These steps were performed for each combination of test conditions. The order of the test conditions was changed in every test, following a Latin square arrangement, to avoid biased results.

5.1.7. Results

In this section, we present the main observations made during the first experiment as well as difficulties and suggestions from participants regarding the test task. To assess the differences among the three techniques for direction control, we collected both objective and subjective data in the form of logs and questionnaires, respectively, during the evaluation sessions. For the continuous variables (collision time, total time, and users' total traveled path lengths), we used the Shapiro-Wilk test to assess data normality.

Because the samples were not normally distributed, we used the Friedman nonparametric test to identify the main effects. Once the main effects had been found, we performed additional Wilcoxon signed-rank post hoc tests with Bonferroni correction to assess the statistical significance between each pair of variables. For the questionnaires, we also used the Friedman nonparametric test to identify the main effects and Wilcoxon signed-rank post hoc tests with Bonferroni correction for each pair of variables.

In the following subsections, we present the analysis of the results from the questionnaires and log files obtained during the tests.

5.1.7.1. Subjective Responses

By analyzing the data from the questionnaires, we identified statistically significant differences regarding the ease of indicating direction (Q2: $\chi^2(2)=25.24$, $p<0.001$), moving around the VE ($\chi^2(2)=11.677$, $p=0.003$) and reaching the rings (Q4: $\chi^2(2)=19.24$, $p<0.001$).

We can infer that the users found it easiest to indicate the direction of movement (Q2) using the hand-steering technique and found it most difficult with the elevator technique. This is explained by the statistical significance results found by comparing the gaze and hand techniques ($Z=-2.414$, $p=0.016$), the elevator and hand techniques ($Z=-3.601$, $p<0.001$), and the elevator and gaze techniques ($Z=-2.635$, $p=0.008$). Statistical significance was also found for Q3 with regard to the finding that users felt it was more difficult to move using the elevator technique than using the hand ($Z=-3.286$, $p=0.001$) and gaze ($Z=-2.919$, $p=0.004$) techniques. The participants found it more difficult to reach the rings (Q4) with the elevator technique than with the gaze technique ($Z=-2.810$, $p=0.005$). Additionally, statistical significance was found with regard to avoiding obstacles (Q5); users found it easier to avoid them with the hand technique than with the elevator technique ($Z=-3.397$, $p=0.001$).

Regarding embodiment (Q8–Q10), we did not identify any statistically significant differences among the tested techniques. Additionally, we did not identify significant differences between the tested pairs of techniques in regard to the ease of walking inside the circle (Q1), feeling of safety (Q6), or fear of heights (Q7). However,

Table 5.1: Results obtained from the questionnaires in the direction indication experiment, presented as median (interquartile range) values. Here, * indicates statistical significance.

	It was easy to...	Elevator	Gaze	Hand
Q1.	... walk inside the circle.*	5 (1)	5 (1)	5.5 (1)
Q2.	... indicate direction of movement.	4 (1)	5 (1)	6 (0)
Q3.	... move around the virtual environment.*	5 (1)	6 (1)	6 (0)
Q4.	... reach the rings.*	5 (2)	6 (1)	6 (1)
Q5.	... avoid obstacles.*	5 (1)	5 (1)	6 (1)

	I felt...	Elevator	Gaze	Hand
Q6.	... safe inside the circle.*	5.5 (1)	6 (1)	6 (1)
Q7.	... fear of heights.	1 (2)	1 (1)	1 (1)
Q8.	... I was in control of the body I was seeing (Agency)	5 (2)	5.5 (1)	5 (1)
Q9.	... that the virtual body was my own (Body Ownership)	4.5 (2)	5 (0)	5 (2)
Q10.	...as if my body was located where I saw the virtual body to be (self-Location).	5 (1)	5 (0)	5 (1)

the participants felt that the elevator technique significantly affected their fear of heights compared to the gaze technique and significantly affected their sense of self-location compared to the hand technique. The questionnaire results are summarized in Table 5.1.

5.1.7.2. Task Performance

To assess differences in user task performance among the different representations, we collected data based on logs. The data collected in this phase included the total task time, total collision time with objects, and path length. We chose the chest point as the reference point for calculating the total distance traveled because this is the most reliable joint provided by Kinect sensors.

Regarding the total task time, we identified statistically significant differences ($\chi^2(2) = 8$, p=0.018). Specifically, significant differences were noted between the gaze and elevator techniques (Z=-2.621, p=0.009), with the gaze technique requiring a shorter amount of time to complete the task, and between the hand and gaze techniques (Z=-2.417, p=0.016), with the hand technique showing an advantage. The data regarding the total path length can be found in Figure 5.5A. Regarding the total collision time, we also found statistically significant differences ($\chi^2(2)$=17.33, p<0.001), with the hand technique showing a clear advantage compared to the elevator technique (Z=-3.461, p=0.001) and the gaze technique (Z=-3.201, p=0.001). A summary of the

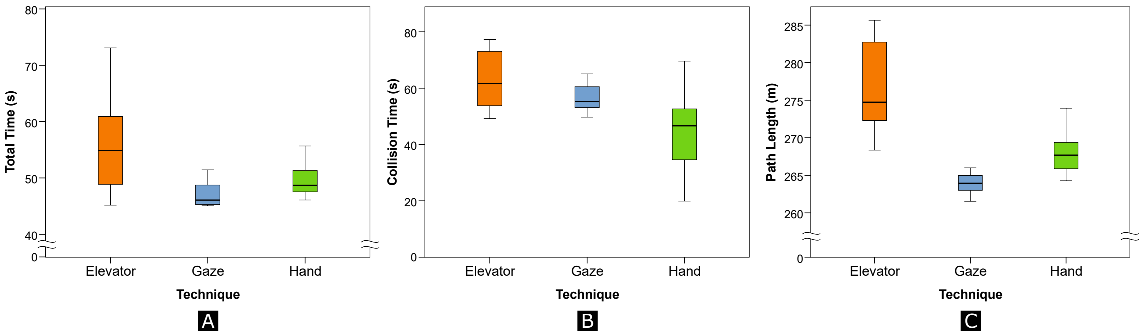

Figure 5.5: Results obtained from the direction indication experiment for (A) total time, (B) collision time, and (C) path length. In each plot, the Elevator+Steering technique is represented in orange, the gaze technique in blue, and the hand technique in green.

total collision time results can be found in Figure 5.5B.

We also found a statistically significant difference regarding the path length ($\chi^2(2)=$ 21.53, p<0.001) (Figure 5.5C) between the gaze and elevator techniques (Z=-3.574, p<0.001), with the gaze technique being associated with the shorter path length. Statistically significant differences were also found between the hand and elevator techniques (Z=-3.101, p=0.002), with the hand technique having the shorter path length, and between the hand and gaze techniques (Z=-3.337, p=0.001), with the gaze technique having the advantage.

5.1.8. Discussion

We found that the Elevator+Steering technique elicited the worst results in our tests. It was the least efficient technique (in terms of total time) because the users spent most of their time colliding with objects. It was also the technique with the longest traveled distance among the three tested techniques.

The results for the gaze- and hand-oriented steering techniques were similar to those found by Bowman et al. [13]. The hand technique had the advantage in terms of efficiency (total time), and the gaze technique the advantage in terms of the distance traveled. The users indicated that the hand technique allowed them to be more aware of the presence of a virtual body. This was attributed to the fact that the users spent a shorter amount of time colliding with objects with the hand technique. With this technique, they had increased awareness of their virtual bodies and could focus on performing the task, but they encountered difficulties in avoiding obstacles. The participants also found it easier to indicate the direction of movement and to navigate in the VE using the hand technique, as indicated by the questionnaires. Another advantage of the hand technique was that it provided the possibility of traveling in a different direction than the direction in which the user was looking, thus enabling him or her to inspect the virtual scene while traveling.

5.2. Study 2 : Speed Control

In the second study, we assessed the impact of the use of close-to-real techniques for controlling speed when flying in VEs. The test design and methodology used in this test were similar to those in the previous study. We also used the same VE and presented a task similar to that in the previous experiment but following a different path (Figure 5.3C). This path was longer, measuring 330 m, and contained abrupt changes in the Z position. We also incorporated more complex maneuvers, such as U-turns, to force users to carefully control their speeds while they flew. Similar to the previous experiment, the subjective SSQ was used to assess cybersickness. Based on the results of the previous evaluation, we employed the hand technique as the technique for indicating direction in combination with all the proposed speed control techniques. Additionally to the dependent variables used in the direction indication we added the speed variation, the percentage of time spent in translation (flying time), and the percentage of time during which the carpet remained stationary (idle time).

For this experiment, we recruited 18 participants; four were female. The ages of the users varied from 21 to 35 years, with an average age of 25 and standard deviation of 3 years. Regarding experience, the majority of the users had previous experience with 3D applications, such as games and modeling systems. The majority also had previous experience with HMDs (88.8 % or 16 participants) and with Kinect usage (72.2 % or 13 participants).

In the following subsections, we present the implemented techniques and outline the details of the obtained results, followed by an in-depth discussion.

5.2.1. Techniques implemented

We tested the speed control capabilities of three different techniques. The techniques ranged from a low level of interaction fidelity (joystick) to a high level of interaction fidelity (speed circle and WIP). To indicate the direction of movement, we used the hand technique, which elicited the best results in the previous study, in combination with all of the tested speed control techniques. For all techniques, the speed was constrained to a maximum of 5 m/s.

5.2.1.1. Joystick

In this technique, the speed was controlled with an analog stick similar to those traditionally used in games (Figure 5.8A). The vertical axis of the joystick was used to control the speed of movement.At the middle of the stick, the speed was zero, and when the stick was moved along the vertical axis, the speed increased until it reached its upper limit. For the comparison of the outcomes elicited with the different techniques, we chose to include only movements along the positive axis of the joystick.

5.2.1.2. Speed Circle

The speed circle technique is an adaptation of the virtual circle metaphor [26][62], in which the human body is used as an analog stick. We utilized a mapping identical to that in the joystick technique but used the position of the hip joint as the input for controlling the user's speed (Figure 5.6). In the neutral zone, which was represented by a green circle in the middle of the speed circle, the movement speed was zero. Outside the neutral zone, the movement speed was determined by the projected distance of the user from the center of the circle. To prevent negative speeds, the circle was divided into two different halves relative to its center, which were updated according to the user's projected forward direction. When the user stepped in the negative half of the circle, the movement stopped (Figure 5.7). Additionally, in U-turn-like movements, the user could adjust his or her position while turning or walking toward the opposite side of the circle. This limitation was explained to the users during the pretest and training scenarios. Despite the use of the torso position as the input to control the speed of translation, we used the orientation of the hand as the means of indicating the direction of movement, as in the joystick technique.

To avoid users stepping outside the ground plane circle when their bodies were within the maximum speed zone, we extended the spatial extent of the zone by 0.5 m in instances when participants reached the border of the circle. This extended

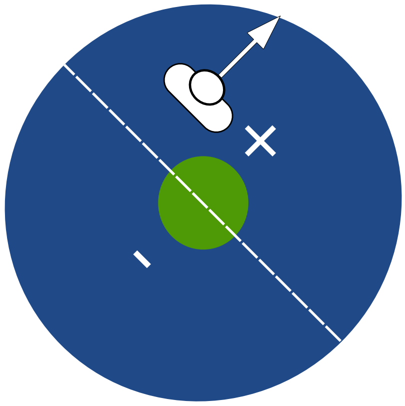

Figure 5.6: Division of the speed circle into positive and negative halves. This division was updated according to the orientation of the user.

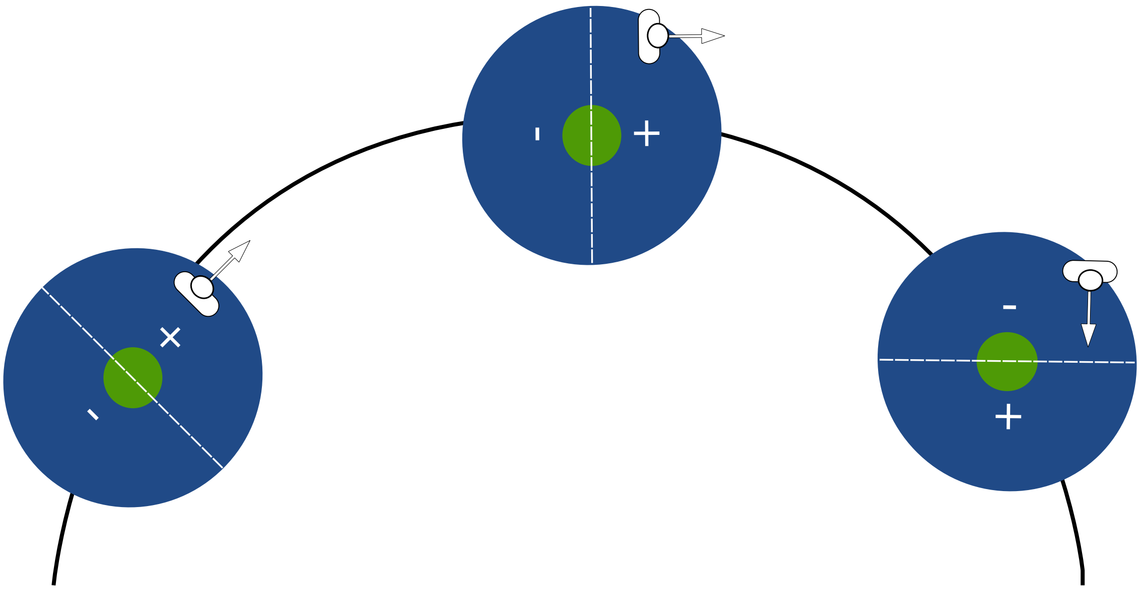

Figure 5.7: Example showing how the movement stops when the user is performing a U-turn.

circle was rendered in yellow to differentiate it from the conventional circle.

5.2.1.3. Walking In Place

Our WIP approach was adapted from that of Bruno et al. [18], which is optimized for data gathered from commodity depth cameras, because this approach employs the knee movements of the user to determine the gait speed. However, in contrast to Bruno et al., we used the hand orientation of the user to determine the overall travel direction because this evaluation scenario was not restricted by a large-scale wall display. To reduce fatigue, we limited the movement needed to reach the maximum speed. To accomplish this, we set a maximum threshold speed of 5 m/s.

5.2.2. Results

Similar to the first experiment we used both objective and subjective data to compare the three techniques for speed control.

We used the the Shapiro-Wilk test to assess data normality. We then applied a repeated measures ANOVA test with Greenhouse-Geisser correction to identify significant differences in normally distributed data and Friedman's nonparametric test with a Wilcoxon signed-rank post hoc test for non-normally distributed data. In both cases, Bonferroni correction (corrected significance = significance x 3) was used in the post hoc tests.

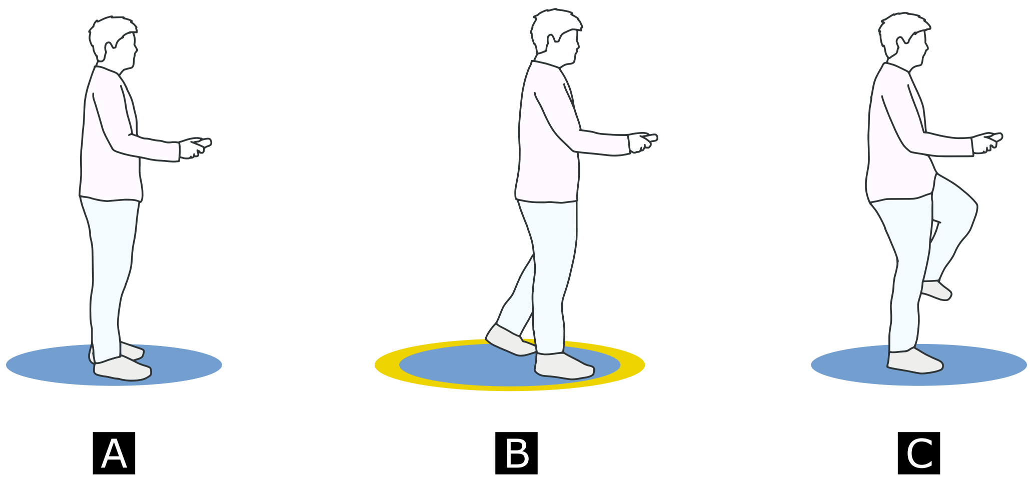

Figure 5.8: Implementation of the speed control techniques: (A) joystick, (B) speed circle, and (C) walking in place. An extra circle (shown in yellow) was rendered when the user reached the border of the default circle.

In the following subsections we present the analysis made based on the results of the questionnaires and log files obtained during the test.

5.2.2.1. Task Performance

In addition to the previously described data (total time, total collision time, and path length), we gathered additional information, such as the speed variation, the percentage of time spent in translation (flying time), and the percentage of time during which the carpet remained stationary (idle time).

Based on the analyzed results, we found statistically significant differences with regard to the total time ($\chi^2(2)$= 27.11, p<0.001), flying time percentage ($\chi^2(2)$=13.765, p=0.001), idle time percentage ($\chi^2(2)$= 10.53, p=0.005), and path length ($\chi^2(2)$= 14.33, p=0.001).

Regarding time, the users completed the task in a shorter time with the joystick (average time=64.9 s) than with either the speed circle (average time=97.73 s, Z=-3.724, p<0.001) or WIP (average time=81.7 s, Z=-3.724, p<0.001) technique. We also noted that the movement was less fluid using the WIP technique (average flying percentage=71.2%, average idle percentage=8.18%), as indicated by the smaller idle time percentage compared to the speed circle (average=88.35%, Z=-3.053, p=0.002) and joystick (average=92.7%, Z=-3.124, p=0.002) techniques and the higher flying time percentage compared to the speed circle (average=11.65%, Z=-3.053, p=0.002) and joystick (average=8.18%, Z=-3-385, p=0.001) techniques. Moreover, we found that the joystick technique resulted in a shorter path length (average length=479.57 m) compared to both the speed circle (average length=494.77 m, Z=-2.765, p=0.006) and WIP (average length=492.6 m, Z=-3.201, p=0.001) techniques.

5.2.2.2. Subjective Responses

The results showed statistically significant differences regarding the ease of walking inside the carpet (Q1: $\chi^2(2)$=10.61, p=0.005), controlling the speed (Q3: $\chi^2(2)$= 25.34, p<0.001), moving around the environment (Q4: $\chi^2(2)$=21.55, p<0.001), reaching the rings (Q5: $\chi^2(2)$=16.74, p<0.001), avoiding obstacles (Q6:$\chi^2(2)$=14.52, p=0.001), and coordinating movements (Q7: $\chi^2(2)$=10.67, p=0.005). There were also statistically significant differences related to feeling safe inside the circle (Q8: $\chi^2(2)$=17.53, p<0.001), the sense of agency (Q9: $\chi^2(2)$=7.190, p=0.027), fatigue (Q12: $\chi^2(2)$=9.8, p=0.007), and the fear of heights (Q13: $\chi^2(2)$=15.056, p=0.001).

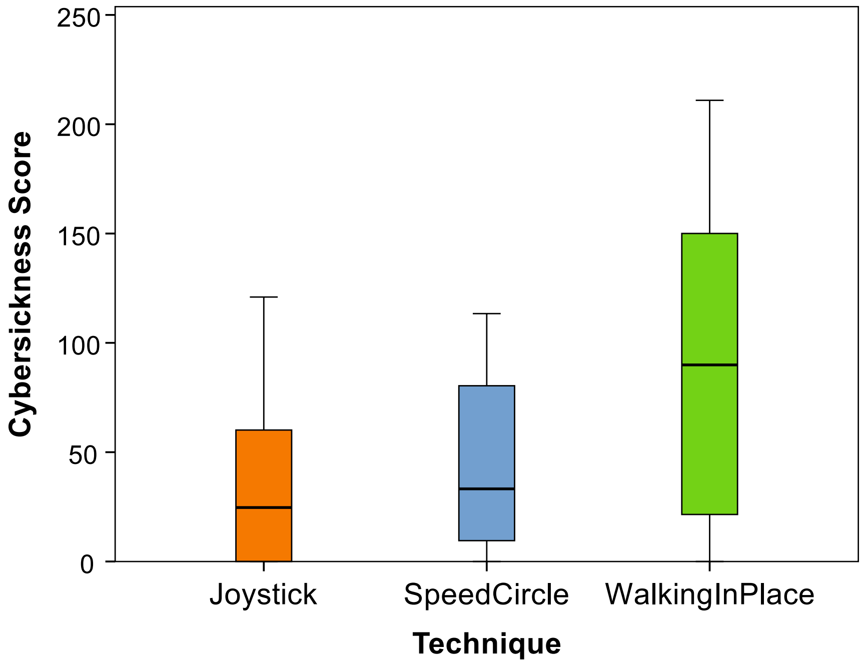

Figure 5.9: Cybersickness Score.

Table 5.2: Results from the questionnaires collected in the second experiment, presented as median (interquartile range) values. Here, * indicates statistical significance.

	It was easy to...	Joystick	Speed Circle	WIP
Q1	... walk inside the circle.*	6 (0)	5 (2)	5 (3)
Q2	... indicate direction of movement.	6 (1)	6 (1)	6 (1)
Q3	... control speed of movement.*	6 (0)	5 (1)	3 (2)
Q4	... move around the VE.*	6 (1)	5 (1)	4.5 (2)
Q5	... reach the rings.*	6 (0)	6 (1)	4.5 (2)
Q6	... avoid obstacles. *	6 (1)	5 (2)	4.5 (3)
Q7	... coordinate movements.*	5 (1)	5 (1)	4 (3)

	I felt...	Joystick	Speed Circle	WIP
Q8.	... safe inside the circle.*	6 (1)	6 (1)	4.5 (3)
Q9.	... fear of heights.	1 (2)	1 (1)	1 (4)
Q10.	... fatigue	1 (1)	1 (2)	3 (4)
Q11.	... I was in control of the body I was seeing (Agency)*	6 (1)	6 (1)	5 (3)
Q12.	... that the virtual body was my own (Body Ownership)	6 (1)	5.5 (1)	5 (2)
Q13.	...as if my body was located where I saw the virtual body to be (self-Location).	6 (1)	5.5 (1)	5 (2)

We also found statistically significant differences regarding the cybersickness score (F(2,50)=4.378, p=0.018).

Regarding the use of the Magic Carpet, the participants found it easiest to walk around the carpet (Q1) using the joystick technique (Z=-2.899, p=0.004) and felt

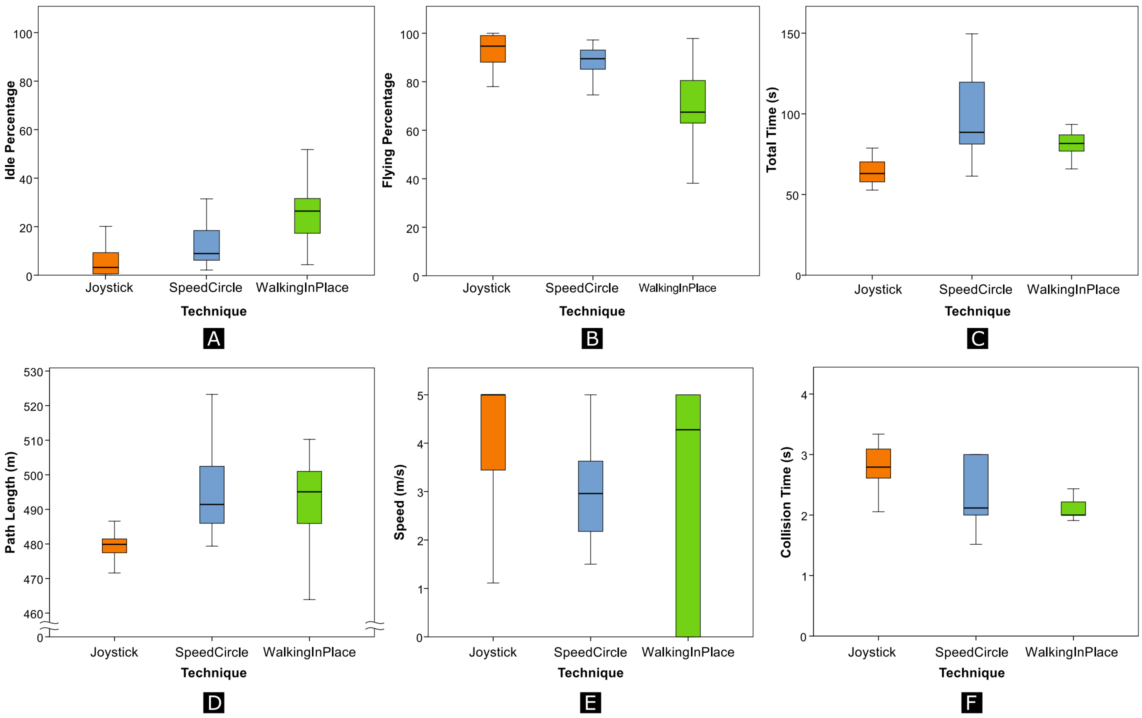

Figure 5.10: Results obtained from the speed control experiment for (A) idle time, (B) flying time, (C) total time, (D) path length, (E) speed variation, and (F) collision time. In each plot, the joystick technique is represented in orange, the speed circle technique in blue, and the walking-in-place technique in green.

less safe within the carpet when using the WIP technique in comparison to both the joystick (Z=-2.979, p=0.003) and speed circle (Z=-2.915, p=0.004) techniques. When asked about speed control (Q4), the participants reported finding it easier with the joystick than with the WIP (Z=-2.750, p<0.001) and speed circle (Z=-2.750, p=0.006) techniques and more difficult overall with the WIP technique compared to the speed circle (Z=-3.016, p=0.003) and joystick techniques.

Using the WIP technique, the users also found it more difficult to move around the environment (Q4) (joystick: Z=-3.471, p=0.001; speed circle: Z=-2.593, p=0.01) and to reach the rings (Q5) (joystick: Z=-2.822, p=0.005; speed circle: Z=-2.946, p=0.003). For obstacle avoidance (Q6), the users overall preferred the joystick technique over the speed circle (Z=-2.547, p=0.011) and WIP (Z=-3.095, p=0.003) techniques. Regarding embodiment (Q11-Q13), we found statistically significant differences only with regard to the sense of agency with the WIP technique, with which users felt they had less control over their virtual bodies compared to the speed circle technique (Z=-2.555, p=0.011). We did not find statistically significant differences among the tested techniques with regard to the fear of heights (Q9). The users also felt more fatigue with the WIP technique (Q10) than with either the joystick (Z=-2.699, p=0.007) or speed circle (Z=-2.840, p=0.005) technique.

Additionally, they found it more difficult to coordinate movements (Q7) with the WIP technique in comparison to the joystick technique (Z=-2.609, p=0.009). The elicited results are summarized in Figure 5.2.

Regarding cybersickness issues, the users indicated additional side effects related to the user experience with the WIP technique (average score=88.12), with severe cases of "Stomach Awareness", "Vertigo", "Dizziness with Eyes Closed", "Nausea", and "General Discomfort" (one instance of each). This finding could be explained by the statistically significant differences found in comparison with the joystick (average score=39.52, t(17)=-3.265, p=0.005) and speed circle (average score=43.84, t(17)=-3.021, p=0.008) techniques.

5.2.3. Discussion

The technique that performed the best was the joystick technique, which had the lower level of interaction fidelity. This may be explained by the fact that is a technique in which users are already familiarized with games that use this metaphor. In the joystick technique, users remained at most of the time at maximum speed. However, we noticed that people effectively had a more refined speed control with the Speed Circle, reducing the speed near curves and increasing it in straight lines. A clear limitation of our proposed technique for speed control, the Speed Circle, needed people to adjust their positions inside the space to make abrupt movements such as U-turns (Figure 5.7). This restriction also does not allow moving backwards with this technique. When making an U-turn for example, people would then rest inside the negative side of the circle (Figure 5.7) and would then start walking backwards. This limitation however, as shown on the results, made people perform a more precise control over the speed of movement. With this we can recommend this speed control technique for flying tasks inside our design space.

We also found that the WIP technique was the worst performing technique and had also a negative impact on embodiment factors, mainly in the Sense of Agency component. This can be explained as this technique needs an active, continuous movement to control speed, it is the one that is mostly affected by the refresh rate of the tracker used.

We found that different high interaction fidelity techniques affect travel quality factors in different ways. So, summarizing all aspects, we can point that the use of high fidelity methods are a good alternative to fly in VR, but the use of such techniques does not always translate directly into improvement in travel quality factors.

6
Effects of Speed and Transitions on Target-Based Techniques

Emerging new technologies in the Virtual Reality facilitate a rapid development of techniques and applications for travel in immersive virtual environments IVE. Travel plays an essential part on the VR experience, where the user moves from a starting point A to a target point B. We also can divide travel in two subcategories. On *Explore* tasks the user moves freely on the VE without a predetermined goal and *Search*, where he/she has to reach a specific checkpoint. The choice of the travel technique can influence the user and cause severe side effects, essentially cybersickness [48], reduced presence and disorientation [100]. The more natural the technique, the more efficiently users can perform travelling tasks on VEs [108], especially on Explore tasks. However, constraints such as fatigue and limitations of the physical space can make it unsuitable to some situations. Indirect methods such as Target-based and Steering techniques [15] can overcome this problem by providing an approach to travel while still providing a favorable spatial orientation on VEs.

Some causes of cybersickness in VR-systems include graphical realism of the environment [27], field of view [32] and navigation speed [101]. Although steering techniques can provide an improved spatial understanding of virtual surroundings, target-based approaches can reliably overcome unwanted symptoms on inexperienced users of immersive systems [80]. In this work we aim to further investigate the effects of speed and transition in Target-based techniques, by comparing three different methods and how they impact the VR experience in key aspects such as comfort and cybersickness.

6.1. Travel Techniques

We implemented three different techniques for travel in IVEs, as depicted in Figure 6.1:

Teleport Technique (TP)

This technique [15], also known as infinite velocity, translates a person instantaneously from their current position to the next checkpoint.

Linear Motion (LM)

This technique consists of moving the user along a linear path for two seconds with a constant velocity, until the next checkpoint.The velocity choice is based on previous work [101] and varies between 30 m/s and 50 m/s depending on the checkpoint distance.

Animated Teleport Box (AT)

We developed the Animated Teleport Box technique with the objective to combat the negative effects of the Teleport technique. Two 1.5 second animations were played when a user was being translated from their current position to next checkpoint. The first one animated the Box to rise up and surround the user, and the second one executed the same animation but in the inverse direction. The box has 2.3 meters on each side so that users would not feel too claustrophobic when travelling. It was developed with the intention of not showing users that they were being moved, as a mean of decreasing the disorientation that might be felt after being teleported.

6.2. Task Design

To validate the techniques described above, we completed a user evaluation. Our aim was to understand which of the techniques were preferred and the impact of

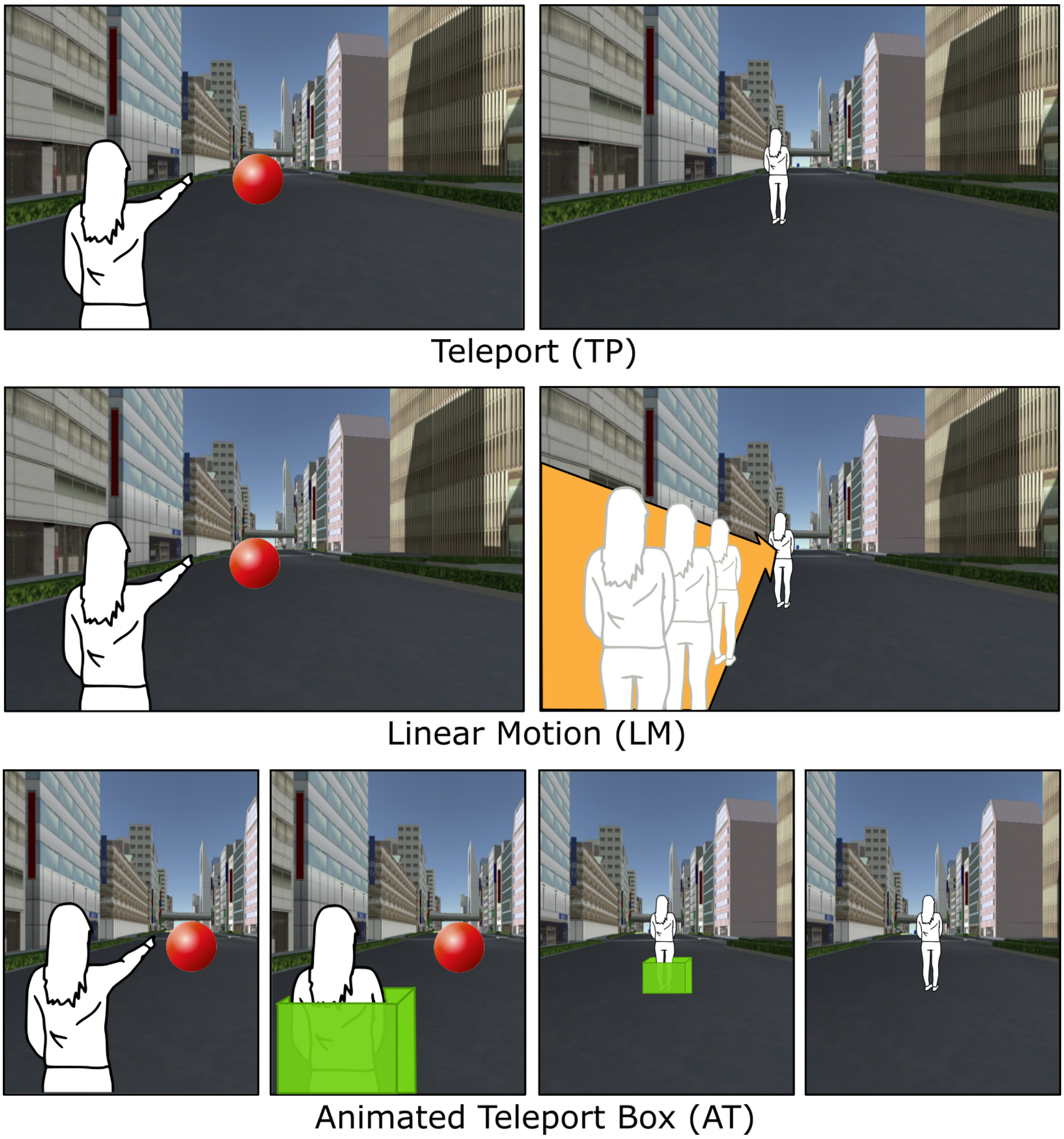

Figure 6.1: Implemented Target-Based Travel techniques.

cybersickness on users. We tested the techniques in our laboratory in a controlled environment, using a Samsung GearVR HMD with a Samsung Galaxy S7 smartphone. Users were able to freely rotate their head within the VE. 20 participants (2 of which female) completed the user evaluation, with ages ranging from 19 to 31 years old (average = 24) and 7 participants already had previous experience in VR. Each user evaluation session adopted the same protocol, starting the initial briefing with a quick explanation to the experiment and also with a description of the techniques. To avoid biased results from users becoming familiarized with the techniques and used to the environment, the techniques were presented in a partial random order, so all permutations were exhausted.

The Virtual Environment was a model of the city of Osaka, Japan (visible in Figure 6.1), which was populated with six spherical checkpoints to where the users would be travelling to. During each travel, the users were told where the next checkpoint would be (to their left or right) and were also instructed to point to said checkpoint before traveling using each of the techniques. The user had no control over the path that he would take, and would only be in charge of pointing to the checkpoints. We allowed the users an adjustment period to the environment, before travelling to the first checkpoint, to make sure they knew where they were and where they were being moved to. Each session took on average thirty minutes, which ended with a brief questionnaire about their experience. Summarizing, the dependent variables used were : 6-likert scale user questionnaire (for subjective metrics), total time and time without animation times.

6.3. Results and Discussion

Throughout data analysis, we first conducted a Shapiro-Wilk which showed that not all samples followed a normal distribution. We then used a Friedman non-parametric test to look for statistical significance between the three tested techniques. When statistical differences were found, we conducted a Wilcoxon Signed-Ranks Test to look for statistical significance on each pair of techniques with an additional Bonferroni correction. For a better comparison regarding task performance, we subtracted the animation times from the total time following the formula : $T' = T - \alpha \times (n-1)$, where T is the total time, α the path time (3 seconds in AT, 2 in LM, and zero in TP) and n the number of travels (6 in our case). Looking at Figure 6.2, we can notice a slightly better performance with AT, but without statistical significance. Because of that we can state that efficiency is similar in all the tested techniques.

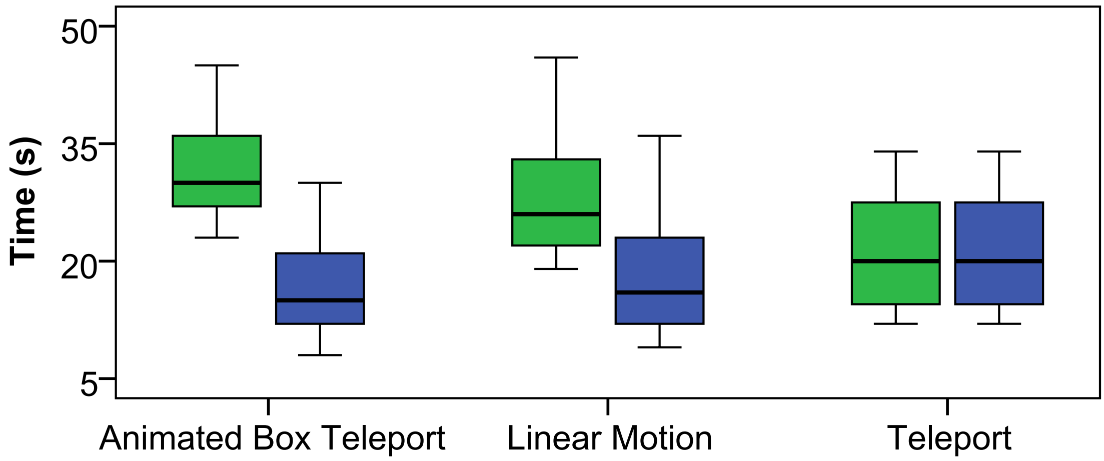

Figure 6.2: Time elapsed on each task. Green boxplots represent total time, and blue the time excluding techniques' animations.

Question	**AT**	**LM**	**TP**
It was easy	5 (1)	5 (1)	5 (1)
I was satisfied	4 (2)	4.5 (2)	4 (2)
I felt physical discomfort*	1 (1)	2 (3)	1 (1)
I felt visual discomfort	1 (1)	2 (2)	1 (1)

Table 6.1: User preferences: Median (Interquartile Range). * indicates statistical significance.

Regarding questionnaires' data (Table 6.1) we found that users felt more physical discomfort using LM (Z=-2.699, $p < 0.01$ against AT and Z=-2.386, p=0.017 against TP). Despite the discomfort caused by LM, participants stated it as their favourite technique in most cases. Due to the similarity between user preferences on both AT and TP we conducted an additional test on the total times of the test task. This test confirms a better result on such condition with TP as it does not need additional time among the movement between positions (Z=-3.114, $p < 0.01$ between AT and Z=-2.578, p=0.01 against LM).

7 Discussion

In this dissertation we conducted several studies in order to validate our research goals. As so, we aimed on the investigation of fidelity factors both in the representation and interaction parts for flying tasks.Both aspects were addressed separately, for an improved analysis of each of the components. Lastly, a study was conducted to assess the effects of speed and transitions in the speed control phase of Target-based techniques.

In travel tasks, the use of a fully-embodied representation lead to an improvement in the form people interact with VEs.After a analysis of the results regarding representation, i.e. how the user is viewed and how it is represented, we can point that the perspective aspect influences the most when navigating around the scene. In this matter, 1PP have the best overall performance (time, collision number and collision time) when comparing to 3PP avatars but also with improved precision. Embodiment factors were also enhanced when a 1PP was used. Our results contradict previous work that pointed an improved spatial awareness when a 3PP avatar is used [72, 37, 89]. These works mostly use different types of setups (such as Salamin et all. [89], which used a video see-through HMD) or present tasks where an amplified view of the surroundings is needed [37, 72]. The only exception was found when the Point-Cloud avatar was used. Since this representation uses a reconstruction of users' bodies inside the Virtual Environment, people feel an equivalent sense of embodiment with both 3PP and 1PP. This equivalence was also found in spatial awareness factors (collision number and collision time). Also, when comparing between the representations in the 3PP we also found a higher feeling of embodiment in all of three sub-components of the sense of embodiment: agency, body-ownership and self-location. With that we can state that the realism of the representation impacts mostly the 3PP, since the avatar's body is always seen in this perspective.

In relation to the interaction component, we can clearly state that the level of interaction fidelity is not directly translated into quality factors, due to the characteristic of each of the techniques used. With the use of our Magic Carpet metaphor, we were able to investigate this by dividing both direction indication and speed control in two different studies, being able to provide high fidelity means to a rather unnatural task. Also, the use of this metaphor, with a conjunct use of a full-body representation put people in a comfortable position to avoid side-effects and to be able to perform further actions. By subdividing the flying in two phases and addressing them separately we are able to isolate the unnatural part of flying, the 6DoF direction control, and provide high interaction fidelity means on both phases of travel. From results from our tests, we can clearly say that the level of interaction fidelity has a positive impact on techniques for indicating direction. Our proposed technique, the Elevator+Steering had the worst performance both in subjective and objective metrics, so we can state that the separation of degrees of freedom is not an efficient method for flying in VEs. A problem with the Elevator+Steering technique is that since camera control is a projection of people's heads' orientation vector, people needed more time to adapt to this technique. During the start of the test, people would often look up and expected to be translated on that direction, but were translated in a different direction. However, after a short time they adapted to the technique and performed the task normally. The technique with higher level of interaction fidelity, the Hand technique, outperformed the other two.

For speed control, although, we did not found a clear relation between interaction fidelity and travel quality factors. Our results contradicts previous studies that reported improvements when a high-fidelity-travel technique was used [62]. From our results, we can infer that WIP was the least suitable technique in terms of task performance in comparison to the other tested approaches. However, we can still consider it efficient in terms of collision time, path length, and total task time. In our tests, we noticed that users experienced more difficulty in coordinating the direction indication and speed control phases with this technique. Consequently, more participants stopped the speed control movement when they reached a ring, then pointed to the next ring, and then flew to it. This behavior explains the increased amount of idle time observed with this technique (Figure 5.8A). The users also stated that with WIP, it was more difficult to control the speed, and they reported more cybersickness and balance issues during the experience. Regarding embodiment, as seen from the results of the questionnaires, the users also felt less control over their virtual bodies (less sense of agency) with the WIP technique. This may be attributed to the noise in the depth sensor signals. Another interesting point reported by the participants was that they lost balance in some cases because of the weight compensation that occurred naturally during their gait in real life, for which the emulated experience did not entirely match the actual experience.

The joystick technique was found to be generally the most efficient technique for flying in immersive environments. This may be explained by the familiarity the users already had with video games. However, in most cases, the users did not finely control the speed, as indicated in Figure 5.10E, but instead maintained the maximum speed most of the time.

From observing the behavior of the participants during the test, we can infer that the participants mostly controlled their speed while they flew based on the speed circle technique (as shown in Figure 5.10E), especially when executing abrupt movements, such as U-turns. The participants often reported that they were lost within the circle. However, they quickly compensated for this issue by recentering themselves before restarting their intended travel actions. They also stated the need to periodically look down at the circle. This, however, had a negative impact on their performances in comparison to the joystick technique. They also stated that more training could improve their performance.

Ultimately, we can state that the use of high-interaction-fidelity techniques is not always the best option for flying in VR. Although the joystick technique has been proven optimal for use in such applications based on our tests, we can still identify the speed circle technique as a good alternative when more precise speed control is required. We can also report that joystick rotation does not induce increased cybersickness in flying tasks, as opposed to room-scale VR [47].

Regarding Target-based techniques, we found conflicting relations between user opinions and objective data gathered through logs. For instance, users said to prefer the Linear Motion Technique, despite being the one that produced most discomfort side-effects. Between the Infinity Velocity techniques, users did not have a clear preference between the two, but the Teleport technique lead to an improvement in efficiency. Additionally, we can point that Infinite Velocity techniques are those who produce less discomfort, but the use of transition effects does not affect neither performance nor cybersickness.

8 Conclusions

Travel is an important part of the VR experience. The incorrect use of a travel technique may compromise efficiency, efficacy of the session and produce side-effects such as cybersickness. When using a HMD, a correctly scaled fully-embodied avatar is an important factor of improving user involvement and efficiency of travel inside a VE [46]. Other factors that influence the travel experience is the representation fidelity of the avatar, which depends on the level of realism (graphical fidelity) and how the avatar is viewed (perspective fidelity) [55]. Commonly, past work limit to analyze each of these elements separately and do not specifically address these factors in a travel scenario.

In some cases, when needing to reach specific points in the scene, people may need to use a flying metaphor to perform these tasks. Due to its unnatural essence, a flying technique may require intricate equipment, which pose people in uncomfortable positions that disable them on performing further actions, such as selection and manipulation. For that, the use of high interaction fidelity techniques can be an effective way for people to fly around the VE. In this thesis, we address this problem by exploring a design space, the Magic Carpet, in which people can use high-interaction fidelity techniques in both direction indication and speed control phases. In the next sessions we summarize the main contributions of this thesis, analyze its drawbacks and pose some directions for further research in the areas of this thesis.

8.1. Dissertation Overview

To familiarize the readers with the context presented in this dissertation we made a systematic review of the two interconnected areas of research in this thesis : the representation of users in immersive setups and travel in flying tasks. Regarding both topics we presented an extensive literature review, necessary for the understanding of the topics presented in this thesis. In conjunction with the literature review, we presented important concepts such as interaction fidelity and define new ones, such as representation fidelity. The literature review was important to identify open problems in the literature and propose an approach for flying using high-fidelity travel techniques. Due to the nature of the topic being investigated, we proposed two different studies to address the representation and interaction factors separately.

The use of a fully-embodied avatar improves the way people interact with the VE and augment their sense of presence. This is even more important when dealing with setups which incorporate HMDs, which complete occludes the avatar. Normally, authors are mostly focused on the graphical fidelity of the representation. The few that are focused in the perspective part and indicate a slight improvement in spatial awareness. To better understand the representation fidelity as a whole we proposed a study that uses three different avatars with increasing level of graphical fidelity. These include an abstract cube avatar, a humanoid mesh avatar and a real-time point-cloud avatar. To enable the isolation of the representation fidelity, we utilized a technique with high level of interaction fidelity, the real walking metaphor. And, to better assess the spatial awareness component, we used three different tasks, which put obstacles in different arrangements on the virtual scene. On the first task the obstacles were disposed around the user, needing them to follow a path avoiding them. The second, the obstacles were put at users feet height, needing them to pass over them. The final task consisted in a tunnel, where users needed to duck down. The choice of a tunnel was made to avoid an increased distance underestimation between the user and the obstacles. All of the obstacles were put following a circular path, in order to maximize tracking space. Data from the user tests was gathered through logs and included time, number of objects collided and total path elapsed. Additionally, questionnaires were used to measure embodiment factors and also to gather opinions from users about the experience. Results shown that the perspective factor has a bigger impact on travel and embodiment quality factors. In most of the cases the use of a 1PP representation proved to improve efficiency in efficiency and embodiment factors. Also, the graphical fidelity had a major impact when the avatar is seen by a 3PP and no significant influence in 1PP avatars. An exception was noticed when a high graphical fidelity avatar with real-time data was used, where embodiment and easiness was similar to its 1PP counterpart.

Following, we address the interaction fidelity in flying scenarios. For this, we explored the Magic Carpet design space, that includes a fully-embodied representation with a floor proxy. This design space enable users to use high interaction fidelity techniques, while staying in a comfortable position and avoiding balance and cybersickness side-effects. Then, we addressed separately the two phases of travel: the direction indication and speed control phases. For each of these phases, a different study was conducted. In the direction indication study, we used three different techniques, varying the level of interaction fidelity in each of them. Moreover, to isolate the speed control component we used a constant speed of 3 m/s. To initiate the movement people needed to press a button, remaining in motion until this button was released. For the lowest interaction fidelity technique we proposed a novel technique called Elevator+Steering in which the degrees of freedom are separated, where the gaze-projection on the floor plane controlled the direction of movement in the horizontal plane and additional buttons on the joystick controlled movement on the vertical plane (up/down). The other techniques consisted in the gaze, where the movement was made following the orientation of users' head and Hand, where the users indicated direction of movement with their hands. On the speed control phase, we assessed three different techniques for this matter with varying levels of interaction fidelity. As the lowest fidelity technique, we used the joystick technique, which consists in controlling speed using the analog stick of a joystick. In the middle of the interaction fidelity spectrum, we proposed a new technique called Speed Circle, which resemble an analog stick of a joystick but using the body position as input. Finally, we used the Walking In Place technique as proposed by Bruno et al. [18], in which people emulate the action of movement while keeping the same physical position. We considered the Walking In Place technique to have the higher technique in comparison with the Speed Control, since in the Speed Control technique people can still remain static physically, but continue in motion in the Virtual Environment. The results show an improvement in quality factors with the increase of interaction fidelity in the direction indication component. In the speed control phase, however we did not find an improvement in quality factors with high-fidelity techniques. However, we can point that people actually controlled better speed of movement with the Speed Circle technique, but it did not translate in improved performance.

Another manner of reaching remote points in the scene is with the use of Target-based techniques. These techniques translate the users from one point to the other in the direction previously indicated. These techniques differ in how the user is translated, such as an immediate translation (with the teleport technique) or gradual (such as the linear motion technique). The use of target-based techniques are an effective way of reducing cybersickness in inexperienced users, while still having a good spatial orientation on the scene. Although, some study is presented in these techniques no work studied the effects of speed and transitions in these kind of techniques. For this, we proposed a study with three different techniques varying how the speed is handled and the presence of transitions. The first is the Teleport Technique (also called Infinite velocity) in which the user is immediately translated. The second is the Linear Motion Technique, in which users are translated with speed until reaching the target position. We then proposed the Animated Teleport Box Technique in which the speed is handled in a similar manner but include a transition effect For all of the tested techniques, we used the Hand technique as direction indication technique.

8.2. Conclusions

Choosing the appropriate travel technique is essential when designing a Virtual Reality experience. Previous research suggest that close-to-real floor-constrained techniques provide a comfortable user experience. However, in some specific cases, people need to fly to better explore virtual environments and be able to reach remote points. Indeed, flying is far from natural way people move in real life and requires the simultaneous control of multiple degrees of freedom. Yet, the supernatural quality of some large environments make flying the most efficient method to travel.

In this work we presented the “Magic Carpet” design space for flying in VR that enables the usage of a fully-embodied representation of the person along with a floor-proxy, which improves presence and avoid side-effects such as imbalance and cybersickness. Our proposed design space enables the usage of high interaction fidelity techniques, which leads to an improvement in travel quality factors in ground-constrained scenarios when a high display fidelity setup is used. To improve efficiency effectiveness, presence and embodiment of the representation used in the "Magic Carpet", we presented a study to determine the best representation to be used in this design space. To isolate the representation aspect, we used the technique with the highest level of interaction fidelity technique, the real-walking technique. To this end, we used three different representations using varying degrees of realism in both first and third-person perspectives, ranging from Abstract to a Realistic Real-time Point-Cloud Representation. To assess each representation-perspective combination, we conducted a comprehensive user study featuring a real-walking navigation task while avoiding obstacles. Among the most salient findings, we can say that the realism of the representations with a 1PP avatar did not seem to interfere with both efficiency and spatial awareness of the user. Although, the increase of graphical fidelity avatar highly affects these factors in 3PP. The use of a real-time reconstruction of the person make the sense of embodiment similar to the same representation in both perspectives.But still, the 1PP remains the most efficient and effective perspective for navigation tasks in VR. For this reason, we have chosen to use an avatar in this perspective to be part of our proposed design space.

To validate our design space, we introduced this design concept by proposing two separate user studies for each part of the flying experience, namely direction indication and speed control. In regards to the first study, we focused on direction indication by means of an assessment of two state-of-the-art techniques, the Gaze technique (which uses the orientation of the head) and the Hand technique, which uses people's dominant hand orientation to specify direction of travel. Additionally, we proposed a novel technique called Elevator+Steering, which uses DOF-separation as means to control direction. The second study focused on the speed control, where we evaluated three different techniques. We also proposed the Speed Circle, a high-interaction fidelity technique for controlling speed in flying scenarios. This approach is based on previous work on ground-constrained travel [26, 62] and enables people to use their body as a joystick to control speed of movement.

Results from the first study suggested that the Elevator+Steering technique had the worst performance among the tested techniques. The hand technique proved to be a more natural technique for this purpose, since it improves awareness of people's bodies, which can be more indicated to more complex scenes. Another advantage of this technique is that the control of movement and camera control are separate, which enabled them to look around the scene while travelling. On the Speed Control study, we found that the Joystick technique performed best, but people remained physically stationary and thus did not seem to control speed, remaining most of the time at maximum speed. We can also point that the Speed Circle technique is a good option to control speed, since people effectively control speed even in abrupt movements and tight trajectories.The results from the second study also suggest that the Walking in Place technique is not a viable option to fly in VR as people often stopped movement to specify direction. Regarding the integration of both techniques together on our design space, we can point out that users did not seem to have difficulty, except for the Walking In Place technique. Finally, we can point that our design space is indeed a viable way to fly in virtual environments and novel. We can also point the flexibility of our design space, where researchers can propose novel high-fidelity interaction techniques to fly in VR. This design space can also be used on the design of passive, low-fidelity techniques such as Target-based techniques for speed control. To validate this, we conducted an study that investigates the effects of target-based techniques regarding travel time, speed and transitions. We propose three different techniques based on previous work by varying said parameters. Through user evaluation, we found that Infinite Velocity techniques cause less discomfort. We also found that using transition effects in conjunction with these techniques does not affect either performance or cybersickness.

8.3. Future Work

This work focused on the study of the effects of fidelity in travel tasks, more importantly in flying tasks. While we succeeded in the use of high fidelity metaphors for both representation and interaction fidelity factors, we point possible directions for future works in the exploration of both factors in Immersive Virtual Environments.

Exploration of the Magic Carpet design space : In this thesis, we proposed the Magic Carpet design space, where people can use high interaction fidelity techniques for flying in VR. For that, we evaluated a set of techniques for determining the best set of techniques for direction control and speed control, where we validated our design space to fly in VR. In that matter, researchers could use state-of-the-art techniques or propose novel travel techniques to interact inside our design space.

Exploration of realistic real-time full-body avatars for travel tasks : A possible direction in the assessment of representation fidelity would be the use of more realistic real-time representations. In this work, we used a real-time point-cloud representation for assessing both graphical and perspective fidelity in travel scenes. We noticed a trend towards representations that contain real-time information for users to have a positive impact on travel quality factors, particularly on embodiment and spatial awareness when using a 3PP. Furthermore, the use of more realistic avatars can be investigated to see if their use implies improvements in the factors mentioned, as is observed when the body of the user is seen through see-through systems.

Study of the effects of representation fidelity in different tasks : In this thesis we focused on travel tasks and found the 1PP still being the best suited perspective for travel. As each task needs different kind of body feedback, we argue that the effects of representation on task effectiveness embodiment factors vary depending on the task being performed. Because of that, similar studies needs to be performed for different use-case scenarios. Additionally, different conditions should be considered for the same combinations of perspectives and realism in avatars, e.g. in collaborative settings, and social environments where communicative tasks engage different users to accomplish success.

Explore flying and user scaling factors in Multiscale Virtual Environments : An interesting use of flying techniques, as said is on Multiscale Virtual Environments, which are VEs witch contain elements with diverging levels of scale in the same environment. When users are navigating throughout the diverging levels of scale they may need to use a flying metaphor, since a ground-floor representation is not always present. Normally, works are mostly interested in how the speed must be controlled in order to establish an efficient way to navigate between the levels of scale both in conventional setups [111] and CAVE setups [2]. Although, when using a HMD coupled with a fully-embodied avatar the user or the Virtual Environment should be scaled properly to improve embodiment and reduce side-effects. In this work we focused on the conception and analysis of effects of interaction fidelity inside our Magic Carpet design space, but this concept can be extended to provide proper user scaling to navigate through this type of environment.

Bibliography

[1] F. Argelaguet, L. Hoyet, M. Trico, and A. Lecuyer. The role of interaction in virtual embodiment: Effects of the virtual hand representation. In *IEEE VR*, 2016.

[2] F. Argelaguet and M. Maignant. Giant: Stereoscopic-compliant multi-scale navigation in ves. In *Proceedings of the 22Nd ACM Conference on Virtual Reality Software and Technology*, VRST '16, pages 269–277, New York, NY, USA, 2016. ACM.

[3] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre. Recent advances in augmented reality. *Computer Graphics and Applications, IEEE*, 21(6), Nov 2001.

[4] R. Ball, C. North, and D. A. Bowman. Move to improve: promoting physical navigation to increase user performance with large displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, pages 191–200, New York, NY, USA, 2007. ACM.

[5] D. Banakou, R. Groten, and M. Slater. Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. *Proceedings of the National Academy of Sciences*, 110(31):12846–12851, 2013.

[6] S. Beckhaus, K. J. Blom, and M. Haringer. Chairio–the chair-based interface. *Concepts and technologies for pervasive games: a reader for pervasive gaming research*, 1:231–264, 2007.

[7] B. B. Bederson, L. Stead, and J. D. Hollan. Pad++: Advances in multiscale interfaces. In *Conference companion on Human factors in computing systems*, pages 315–316. ACM, 1994.

[8] F. Biocca. The cyborg's dilemma: Progressive embodiment in virtual environments. *Journal of computer-mediated communication*, 3(2):JCMC324, 1997.

[9] M. Botvinick, J. Cohen, et al. Rubber hands' feel'touch that eyes see. *Nature*, 391(6669):756–756, 1998.

[10] R. Boulic, D. Maupu, and D. Thalmann. On scaling strategies for the full-body postural control of virtual mannequins. *Interacting with Computers*, 21(1):11–25, 2009.

[11] P. Bourdin, I. Barberia, R. Oliva, and M. Slater. A virtual out-of-body experience reduces fear of death. *PloS one*, 12(1):e0169343, 2017.

[12] D. A. Bowman, E. T. Davis, L. F. Hodges, and A. N. Badre. Maintaining spatial orientation during travel in an immersive virtual environment. *Presence: Teleoperators and Virtual Environments*, 8(6):618–631, 1999.

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

[14] D. A. Bowman, E. Kruijff, J. J. LaViola Jr, and I. Poupyrev. An introduction to 3-d user interface design. *Presence: Teleoperators and virtual environments*, 10(1):96–108, 2001.

[15] D. A. Bowman, E. Kruijff, J. J. LaViola Jr, and I. Poupyrev. *3D user interfaces: theory and practice*. Addison-Wesley, 2004.

[16] G. Bruder, F. A. Sanz, A.-H. Olivier, and A. Lécuyer. Distance estimation in large immersive projection systems, revisited. In *2015 IEEE Virtual Reality (VR)*, pages 27–32. IEEE, 2015.

[17] G. Bruder, F. Steinicke, and K. H. Hinrichs. Arch-explore: A natural user interface for immersive architectural walkthroughs. In *3D User Interfaces, 2009. 3DUI 2009. IEEE Symposium on*, pages 75–82, March 2009.

[18] L. Bruno, M. Sousa, A. Ferreira, J. M. Pereira, and J. Jorge. Hip-directed walking-in-place using a single depth camera. *International Journal of Human-Computer Studies*, 105:1 – 11, 2017.

[19] F. Carvalho, D. R. Trindade, P. F. Dam, A. Raposo, and I. H. dos Santos. Dynamic adjustment of stereo parameters for virtual reality tools. In *Virtual Reality (SVR), 2011 XIII Symposium on*, pages 66–72. IEEE, 2011.

[20] W. Chen, A. Plancoulaine, N. Férey, D. Touraine, J. Nelson, and P. Bourdot. 6dof navigation in virtual worlds: comparison of joystick-based and head-controlled paradigms. In *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology*, pages 111–114. ACM, 2013.

[21] I. Cho, J. Li, and Z. Wartell. Evaluating dynamic-adjustment of stereo view pa-

rameters in a multi-scale virtual environment. In *3D User Interfaces (3DUI), 2014 IEEE Symposium on*, pages 91–98. IEEE, 2014.

[22] G. Cirio, P. Vangorp, E. Chapoulie, M. Marchal, A. Lecuyer, and G. Drettakis. Walking in a cube: Novel metaphors for safely navigating large virtual environments in restricted real workspaces. *Visualization and Computer Graphics, IEEE Transactions on*, 18(4):546–554, 2012.

[23] 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.

[24] J. E. Cutting. Reconceiving perceptual space. In H. Hecht, R. Schwartz, and M. Atherton, editors, *Looking into Pictures*. The MIT Press, Cambridge, MA, 2003.

[25] J. E. Cutting and P. M. Vishton. Chapter 3 - perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth*. In W. Epstein and S. Rogers, editors, *Perception of Space and Motion*, Handbook of Perception and Cognition, pages 69 – 117. Academic Press, San Diego, 1995.

[26] P. Dam, P. Braz, and A. Raposo. A study of navigation and selection techniques in virtual environments using microsoft kinect®. In *International Conference on Virtual, Augmented and Mixed Reality*, pages 139–148. Springer, 2013.

[27] S. Davis, K. Nesbitt, and E. Nalivaiko. Comparing the onset of cybersickness using the oculus rift and two virtual roller coasters. In *Proceedings of the 11th Australasian Conference on Interactive Entertainment (IE 2015)*, volume 27, page 30, 2015.

[28] H. G. Debarba, S. Bovet, R. Salomon, O. Blanke, B. Herbelin, and R. Boulic. Characterizing first and third person viewpoints and their alternation for embodied interaction in virtual reality. *PloS one*, 12(12):e0190109, 2017.

[29] H. G. Debarba, E. Molla, B. Herbelin, and R. Boulic. Characterizing embodied interaction in first and third person perspective viewpoints. In *3D User Interfaces (3DUI), 2015 IEEE Symposium on*, pages 67–72. IEEE, 2015.

[30] A. Denisova and P. Cairns. First person vs. third person perspective in digital games: do player preferences affect immersion? In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 145–148. ACM, 2015.

[31] H. H. Ehrsson. The experimental induction of out-of-body experiences. *Science*, 317(5841):1048–1048, 2007.

[32] A. S. Fernandes and S. K. Feiner. Combating vr sickness through subtle dynamic field-of-view modification. In *2016 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 201–210. IEEE, 2016.

[33] S. Freitag, D. Rausch, and T. Kuhlen. Reorientation in virtual environments using interactive portals. In *2014 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 119–122, March 2014.

[34] G. W. Furnas and B. B. Bederson. Space-scale diagrams: Understanding multiscale interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 234–241. ACM Press/Addison-Wesley Publishing Co., 1995.

[35] S. J. Gerathewohl. *Fidelity of simulation and transfer of training: a review of the problem.* Department of Transportation, Federal Aviation Administration, Office of Aviation Medicine, 1969.

[36] M. Glueck and A. Khan. Considering multiscale scenes to elucidate problems encumbering three-dimensional intellection and navigation. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 25(04):393–407, 2011.

[37] 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.

[38] D. Hildebrandt and R. Timm. An assisting, constrained 3d navigation technique for multiscale virtual 3d city models. *GeoInformatica*, 18(3):537–567, 2014.

[39] V. Interrante, B. Ries, and L. Anderson. Distance perception in immersive virtual environments, revisited. In *IEEE Virtual Reality Conference (VR 2006)*, pages 3–10. IEEE, 2006.

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

[41] K. Kilteni, R. Groten, and M. Slater. The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, 21(4):373–387, 2012.

[42] R. Kopper, T. Ni, D. A. Bowman, and M. Pinho. Design and evaluation of navigation techniques for multiscale virtual environments. In *IEEE Virtual Reality Conference (VR 2006)*, pages 175–182. IEEE, 2006.

[43] T. Kosch, R. Boldt, M. Hoppe, P. Knierim, and M. Funk. Exploring the optimal point of view in third person out-of-body experiences. In *Proccedings of 10th International Conference on Pervasive Technologies Related to Assistive Environments (PETRA 2016).* ACM, 2016.

[44] D. Krupke, P. Lubos, L. Demski, J. Brinkhoff, G. Weber, F. Willke, and F. Steinicke. Evaluation of control methods in a supernatural zero-gravity flight simulator. In *Proceedings of the GI-Workshop VR/AR*, 2015.

[45] D. Krupke, P. Lubos, L. Demski, J. Brinkhoff, G. Weber, F. Willke, and F. Steinicke. Control methods in a supernatural flight simulator. In *2016 IEEE Virtual Reality (VR)*, pages 329–329, March 2016.

[46] E. Langbehn, G. Bruder, and F. Steinicke. Scale matters! analysis of dominant scale estimation in the presence of conflicting cues in multi-scale collaborative virtual environments. In *2016 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 211–220. IEEE, 2016.

[47] E. Langbehn, P. Lubos, and F. Steinicke. Evaluation of locomotion techniques for room-scale vr. joystick, teleportation, and redirected walking. In *Proceedings of the Virtual Reality International Conference (VRIC)*, 2018.

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

[49] J. J. LaViola Jr, D. A. Feliz, D. F. Keefe, and R. C. Zeleznik. Hands-free multi-scale navigation in virtual environments. In *Proceedings of the 2001 symposium on Interactive 3D graphics*, pages 9–15. ACM, 2001.

[50] J. J. LaViola Jr, E. Kruijff, R. P. McMahan, D. Bowman, and I. P. Poupyrev. *3D user interfaces: theory and practice*. Addison-Wesley Professional, 2017.

[51] B. Lenggenhager, M. Mouthon, and O. Blanke. Spatial aspects of bodily self-consciousness. *Consciousness and cognition*, 18(1):110–117, 2009.

[52] C.-T. Lin, S.-W. Chuang, Y.-C. Chen, L.-W. Ko, S.-F. Liang, and T.-P. Jung. Eeg effects of motion sickness induced in a dynamic virtual reality environment. In *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 3872–3875. IEEE, 2007.

[53] Q. Lin, J. Rieser, and B. Bodenheimer. Stepping over and ducking under: The influence of an avatar on locomotion in an hmd-based immersive virtual environment. In *Proceedings of the ACM Symposium on Applied Perception*, pages 7–10. ACM, 2012.

[54] J. L. Lugrin, M. Landeck, and M. E. Latoschik. Avatar embodiment realism and virtual fitness training. In *2015 IEEE Virtual Reality (VR)*, pages 225–226, March 2015.

[55] J.-L. Lugrin, J. Latt, and M. E. Latoschik. Anthropomorphism and illusion of virtual body ownership. In *Proceedings of the 25th International Conference on Artificial Reality and Telexistence and 20th Eurographics Symposium on Virtual Environments*, ICAT - EGVE '15, pages 1–8, Aire-la-Ville, Switzerland, Switzerland, 2015. Eurographics Association.

[56] J.-L. Lugrin, J. Latt, and M. E. Latoschik. Avatar anthropomorphism and illusion of body ownership in vr. In *2015 IEEE Virtual Reality (VR)*, pages 229–230. IEEE, 2015.

[57] J.-L. Lugrin, M. Wiedemann, D. Bieberstein, and M. E. Latoschik. Influence of avatar realism on stressful situation in vr. *2015 IEEE Virtual Reality (VR)*, pages 227–228, 2015.

[58] N. Marquardt, R. Diaz-Marino, S. Boring, and S. Greenberg. The proximity toolkit: Prototyping proxemic interactions in ubiquitous computing ecologies. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, UIST '11, pages 315–326, New York, NY, USA, 2011. ACM.

[59] A. Maselli and M. Slater. The building blocks of the full body ownership illusion. *Front Hum Neurosci*, 7, 03 2013.

[60] J. McCrae, M. Glueck, T. Grossman, A. Khan, and K. Singh. Exploring the design space of multiscale 3d orientation. In *Proceedings of the International Conference on Advanced Visual Interfaces*, pages 81–88. ACM, 2010.

[61] J. McCrae, I. Mordatch, M. Glueck, and A. Khan. Multiscale 3d navigation. In *Proceedings of the 2009 symposium on Interactive 3D graphics and games*, pages 7–14. ACM, 2009.

[62] R. P. McMahan, D. A. Bowman, D. J. Zielinski, and R. B. Brady. Evaluating display fidelity and interaction fidelity in a virtual reality game. *IEEE Transactions on Visualization and Computer Graphics*, 18(4):626–633, 2012.

[63] E. A. McManus, B. Bodenheimer, S. Streuber, S. De La Rosa, H. H. Bülthoff, and B. J. Mohler. The influence of avatar (self and character) animations on distance estimation, object interaction and locomotion in immersive virtual environments. In *Proceedings of the ACM SIGGRAPH Symposium on applied perception in graphics and visualization*, pages 37–44. ACM, 2011.

[64] D. Medeiros, E. Cordeiro, D. Mendes, M. Sousa, A. Raposo, A. Ferreira, and J. Jorge. Effects of speed and transitions on target-based travel techniques. In *Proceedings of the 22Nd ACM Conference on Virtual Reality Software and Technology*, VRST '16, pages 327–328, New York, NY, USA, 2016. ACM.

[65] D. Medeiros, L. Teixeira, F. Carvalho, I. Santos, and A. Raposo. A tablet-based 3d interaction tool for virtual engineering environments. In *Proceedings of the 12th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*, VRCAI '13, pages 211–218, New York, NY, USA, 2013. ACM.

[66] M. Meehan, B. Insko, M. Whitton, and F. P. Brooks Jr. Physiological measures of

presence in stressful virtual environments. In *Acm transactions on graphics (tog)*, volume 21, pages 645–652. ACM, 2002.

[67] D. Mendes, F. Relvas, A. Ferreira, and J. Jorge. The benefits of dof separation in mid-air 3d object manipulation. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, pages 261–268. ACM, 2016.

[68] M. Mine et al. Virtual environment interaction techniques. *UNC Chapel Hill computer science technical report TR95-018*, pages 507248–2, 1995.

[69] C. Moerman, D. Marchal, and L. Grisoni. Drag'n go: Simple and fast navigation in virtual environment. In *3D User Interfaces (3DUI), 2012 IEEE Symposium on*, pages 15–18. IEEE, 2012.

[70] B. J. Mohler, H. H. Bülthoff, W. B. Thompson, and S. H. Creem-Regehr. A full-body avatar improves egocentric distance judgments in an immersive virtual environment. In *Proceedings of the 5th symposium on Applied perception in graphics and visualization*, page 194. ACM, 2008.

[71] B. J. Mohler, S. H. Creem-Regehr, W. B. Thompson, and H. H. Bülthoff. The effect of viewing a self-avatar on distance judgments in an hmd-based virtual environment. *Presence: Teleoperators and Virtual Environments*, 19(3):230–242, 2010.

[72] D. Monteiro, H.-N. Liang, W. Xu, M. Brucker, V. Nanjappan, and Y. Yue. Evaluating enjoyment, presence, and emulator sickness in vr games based on first-and third-person viewing perspectives. *Computer Animation and Virtual Worlds*, page e1830.

[73] M. Mori, K. F. MacDorman, and N. Kageki. The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine*, 19(2):98–100, 2012.

[74] M. Nabiyouni, B. Laha, and D. A. Bowman. Poster: Designing effective travel techniques with bare-hand interaction. In *3D User Interfaces (3DUI), 2014 IEEE Symposium on*, pages 139–140. IEEE, 2014.

[75] J.-M. Normand, E. Giannopoulos, B. Spanlang, and M. Slater. Multisensory stimulation can induce an illusion of larger belly size in immersive virtual reality. *PloS one*, 6(1):e16128, 2011.

[76] K. Perlin and D. Fox. Pad: an alternative approach to the computer interface. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 57–64. ACM, 1993.

[77] V. I. Petkova, M. Khoshnevis, and H. H. Ehrsson. The perspective matters! multisensory integration in ego-centric reference frames determines full-body ownership. *Frontiers in psychology*, 2:35, 2011.

[78] L. Piwek, L. S. McKay, and F. E. Pollick. Empirical evaluation of the uncanny valley hypothesis fails to confirm the predicted effect of motion. *Cognition*, 130(3):271–277, 2014.

[79] J. Plouzeau, D. Paillot, J.-R. Chardonnet, and F. Merienne. Effect of proprioceptive vibrations on simulator sickness during navigation task in virtual environment. 2015.

[80] E. D. Ragan, A. Wood, R. P. McMahan, and D. A. Bowman. Trade-offs related to travel techniques and level of display fidelity in virtual data-analysis environments. In *ICAT/EGVE/EuroVR*, pages 81–84, 2012.

[81] G. Ren, C. Li, E. O'Neill, and P. Willis. 3d freehand gestural navigation for interactive public displays. *Computer Graphics and Applications, IEEE*, 33(2):47–55, 2013.

[82] R. S. Renner, B. M. Velichkovsky, and J. R. Helmert. The perception of egocentric distances in virtual environments-a review. *ACM Computing Surveys (CSUR)*, 46(2):23, 2013.

[83] B. Rey, M. Alcañiz, J. Tembl, and V. Parkhutik. Brain activity and presence: a preliminary study in different immersive conditions using transcranial doppler monitoring. *Virtual Reality*, 14(1):55–65, 2010.

[84] M. Rheiner. Birdly an attempt to fly. In *ACM SIGGRAPH 2014 Emerging Technologies*, page 3. ACM, 2014.

[85] B. Ries, V. Interrante, M. Kaeding, and L. Anderson. The effect of self-embodiment on distance perception in immersive virtual environments. In *Proceedings of the 2008 ACM symposium on Virtual reality software and technology*, pages 167–170. ACM, 2008.

[86] B. Ries, V. Interrante, M. Kaeding, and L. Phillips. Analyzing the effect of a virtual avatar's geometric and motion fidelity on ego-centric spatial perception in immersive virtual environments. In *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology*, VRST '09, pages 59–66, New York, NY, USA, 2009. ACM.

[87] J. P. Rolland and H. Fuchs. Optical versus video see-through head-mounted displays in medical visualization. *Presence: Teleoperators and Virtual Environments*, 9(3), 2000.

[88] R. A. Ruddle, S. J. Payne, and D. M. Jones. Navigating large-scale virtual environments: What differences occur between helmet-mounted and desk-top displays? *Presence*, 8(2):157–168, April 1999.

[89] P. Salamin, T. Tadi, O. Blanke, F. Vexo, and D. Thalmann. Quantifying effects of exposure to the third and first-person perspectives in virtual-reality-based training. *IEEE Transactions on Learning Technologies*, 3(3):272–276, 2010.

[90] P. Salamin, D. Thalmann, and F. Vexo. Improved third-person perspective: a solution reducing occlusion of the 3pp? In *Proceedings of The 7th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*, page 30. ACM, 2008.

[91] M. V. Sanchez-Vives, B. Spanlang, A. Frisoli, M. Bergamasco, and M. Slater. Virtual hand illusion induced by visuomotor correlations. *PloS one*, 5(4):e10381, 2010.

[92] F. A. Sanz. Adaptive navigation for virtual environments. In *IEEE Symposium on 3D User Interfaces*, pages 91–94, 2014.

[93] E. L. Schuurink and A. Toet. Effects of third person perspective on affective appraisal and engagement: Findings from second life. *Simulation & Gaming*, 41(5):724–742, 2010.

[94] T. Seyed, A. Azazi, E. Chan, Y. Wang, and F. Maurer. Sod-toolkit: A toolkit for interactively prototyping and developing multi-sensor, multi-device environments. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces*, ITS '15, pages 171–180, New York, NY, USA, 2015. ACM.

[95] E. Sikström, A. de Götzen, and S. Serafin. Wings and flying in immersive vr #x2014; controller type, sound effects and experienced ownership and agency. In *2015 IEEE Virtual Reality (VR)*, pages 281–282, March 2015.

[96] M. Slater, D. Pérez Marcos, H. Ehrsson, and M. V. Sanchez-Vives. Towards a digital body: the virtual arm illusion. *Frontiers in human neuroscience*, 2:6, 2008.

[97] M. Slater, B. Spanlang, and D. Corominas. Simulating virtual environments within virtual environments as the basis for a psychophysics of presence. *ACM Transactions on Graphics (TOG)*, 29(4):92, 2010.

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

[99] M. Slater, M. Usoh, and A. Steed. Taking steps: the influence of a walking technique on presence in virtual reality. *ACM Trans. Comput.-Hum. Interact.*, 2(3):201–219, Sept. 1995.

[100] S. P. Smith and T. Marsh. Evaluating design guidelines for reducing user disorientation in a desktop virtual environment. *Virtual Reality*, 8(1):55–62, 2004.

[101] R. H. So, W. Lo, and A. T. Ho. Effects of navigation speed on motion sickness caused by an immersive virtual environment. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 43(3):452–461, 2001.

[102] L. Soares, L. Nomura, M. Cabral, L. Dulley, M. Guimarães, R. Lopes, and M. Zuffo.

Virtual hang-gliding over rio de janeiro. In *Proceedings of the 2004 IEEE Virtual Reality Workshop "Virtual Reality for Public Consumption*. Citeseer, 2004.

[103] M. Sousa, D. Mendes, R. K. D. Anjos, D. Medeiros, A. Ferreira, A. Raposo, J. a. M. Pereira, and J. Jorge. Creepy tracker toolkit for context-aware interfaces. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*, ISS '17, pages 191–200, New York, NY, USA, 2017. ACM.

[104] M. Sousa, D. Mendes, R. K. D. Anjos, D. Medeiros, A. Ferreira, A. Raposo, J. a. M. Pereira, and J. Jorge. Creepy tracker toolkit for context-aware interfaces. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces*, ISS '17, pages 191–200, New York, NY, USA, 2017. ACM.

[105] B. Spanlang, J.-M. Normand, D. Borland, K. Kilteni, E. Giannopoulos, A. Pomés, M. González-Franco, D. Perez-Marcos, J. Arroyo-Palacios, X. N. Muncunill, and M. Slater. How to build an embodiment lab: Achieving body representation illusions in virtual reality. *Frontiers in Robotics and AI*, 1:9, 2014.

[106] W. Steptoe, S. Julier, and A. Steed. Presence and discernability in conventional and non-photorealistic immersive augmented reality. In *Mixed and Augmented Reality (ISMAR), 2014 IEEE International Symposium on*, Sept 2014.

[107] R. Stoakley, M. J. Conway, and R. Pausch. Virtual reality on a wim: interactive worlds in miniature. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 265–272. ACM Press/Addison-Wesley Publishing Co., 1995.

[108] E. A. Suma, S. L. Finkelstein, M. Reid, S. V. Babu, A. C. Ulinski, and L. F. Hodges. Evaluation of the cognitive effects of travel technique in complex real and virtual environments. *Visualization and Computer Graphics, IEEE Transactions on*, 16(4):690–702, 2010.

[109] J. E. Swan II, A. Jones, E. Kolstad, M. A. Livingston, and H. S. Smallman. Egocentric depth judgments in optical, see-through augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 13(3), May 2007.

[110] D. Swapp, J. Williams, and A. Steed. The implementation of a novel walking interface within an immersive display. In *3D User Interfaces (3DUI), 2010 IEEE Symposium on*, pages 71–74. IEEE, 2010.

[111] H. Taunay, V. Rodrigues, R. Braga, P. Elias, L. Reis, and A. Raposo. A spatial partitioning heuristic for automatic adjustment of the 3d navigation speed in multiscale virtual environments. In *3D User Interfaces (3DUI), 2015 IEEE Symposium on*, pages 51–58. IEEE, 2015.

[112] X. Tong, A. Kitson, M. Salimi, D. Fracchia, D. Gromala, and B. E. Riecke. Exploring

embodied experience of flying in a virtual reality game with kinect. In *Mixed Reality Art (MRA), IEEE International Workshop on*, pages 5–6. IEEE, 2016.

[113] D. R. Trindade and A. B. Raposo. Improving 3d navigation techniques in multiscale environments: a cubemap-based approach. *Multimedia Tools and Applications*, 73(2):939–959, 2014.

[114] M. Usoh, K. Arthur, M. C. Whitton, R. Bastos, A. Steed, M. Slater, and F. P. Brooks Jr. Walking¿ walking-in-place¿ flying, in virtual environments. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pages 359–364. ACM Press/Addison-Wesley Publishing Co., 1999.

[115] J. Wang and R. Lindeman. Leaning-based travel interfaces revisited: Frontal versus sidewise stances for flying in 3d virtual spaces. In *Proceedings of the 18th ACM Symposium on Virtual Reality Software and Technology*, VRST '12, pages 121–128, New York, NY, USA, 2012. ACM.

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

[117] C.-J. Wu, S. Houben, and N. Marquardt. Eaglesense: Tracking people and devices in interactive spaces using real-time top-view depth-sensing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 3929–3942. ACM, 2017.

[118] S. D. Young, B. D. Adelstein, and S. R. Ellis. Demand characteristics of a questionnaire used to assess motion sickness in a virtual environment. In *IEEE Virtual Reality Conference (VR 2006)*, pages 97–102. IEEE, 2006.

[119] Y. Yuan and A. Steed. Is the rubber hand illusion induced by immersive virtual reality? In *2010 IEEE Virtual Reality Conference (VR)*, pages 95–102. IEEE, 2010.

[120] X. Zhang. Space-scale animation: Enhancing cross-scale understanding of multiscale structures in multiple views. In *Coordinated and Multiple Views in Exploratory Visualization (CMV'05)*, pages 109–120. IEEE, 2005.

[121] X. L. Zhang. Multiscale traveling: crossing the boundary between space and scale. *Virtual reality*, 13(2):101–115, 2009.

A
Questionnaires

A.1. Template of the Profile Questionnaire Used

Profile Magic Carpet

* Required

1. **Id**

2. **Age** *

3. **Gender** *

Mark only one oval.

- Male
- Female

4. **Professional Qualification** *

Experience with Virtual Environments

5. **What types of 3D applications do you use?** *

Check all that apply.

- [ ] 3D Visualization
- [ ] 3D Games
- [ ] 3D Modelling
- [ ] None of the Above
- [ ] Other:

6. **How often do you navigate in 3D environments:** *

Mark only one oval.

- At least once a day
- At least three times per week
- At least once a week
- At least once every 15 days
- Rarely
- Never

7. **How often do you select objects in 3D environments?** *
Mark only one oval.

- At least once a day
- At least three times per week
- At least once a week
- At least once every 15 days
- Rarely
- Never

8. **How often do you use HMDs? (Oculus Rift, Vive, etc)** *
Mark only one oval.

- At least once a day
- At least three times per week
- At least once a week
- At least once every 15 days
- Rarely
- Never

9. **How often do you play games with gamepad? (Xbox, Playstation 4, etc)** *
Mark only one oval.

- At least once a day
- At least three times per week
- At least once a week
- At least once every 15 days
- Rarely
- Never

10. **How often do you use games on Nintendo Wii, Playstation Move or Kinect?** *
Mark only one oval.

- At least once a day
- At least three times per week
- At least once a week
- At least once every 15 days
- Rarely
- Never

11. **How often do you use tracking systems (Optitrack, ART, etc.)?** *
Mark only one oval.

- At least once a day
- At least three times per week
- At least once a week
- At least once every 15 days
- Rarely
- Never

12. **How would you rate the level of your fear of heights?** *
Mark only one oval.

- Nonexistent
- Light
- Moderate
- Severe

A.2. Post-test Questionnaire of the Representation Fidelity Study

Abstract First Person Perspective

* Required

I felt that I was controlling the body that I was seeing *

	1	2	3	4	5	6	
Totally Disagree	○	○	○	○	○	○	Totally Agree

I felt that the virtual body was my own body *

	1	2	3	4	5	6	
Totally Disagree	○	○	○	○	○	○	Totally Agree

I felt that my body was located at the same place of my real body *

	1	2	3	4	5	6	
Totally Disagree	○	○	○	○	○	○	Totally Agree

I felt that I had more than one body *

	1	2	3	4	5	6	
Totally Disagree	○	○	○	○	○	○	Totally Agree

It was easy to avoid the obstacles in the virtual environment (task 1) *

	1	2	3	4	5	6	
Totally Disagree	○	○	○	○	○	○	Totally Agree

It was easy to go over the obstacles in the virtual environment (task 2) *

	1	2	3	4	5	6	
Totally Disagree	○	○	○	○	○	○	Totally Agree

It was easy to go under the tunnel in the virtual environment (task 3) *

	1	2	3	4	5	6	
Totally Disagree	○	○	○	○	○	○	Totally Agree

It was easy to go under the tunnel in the virtual environment (task 3) *

	1	2	3	4	5	6	
Totally Disagree	○	○	○	○	○	○	Totally Agree

SUBMIT

A.3. Post-test Questionnaire of the Interaction Fidelity Study (Magic Carpet)

Cybersickness Questionnaire (Joystick)

Indicate in each of the answers how you are feeling at the moment regarding ...

* Required

1. **General discomfort** *
Mark only one oval.
- None
- Slight
- Moderate
- Severe

2. **Fatigue** *
Mark only one oval.
- None
- Slight
- Moderate
- Severe

3. **Headache** *
Mark only one oval.
- None
- Slight
- Moderate
- Severe

4. **Eye strain** *
Mark only one oval.
- None
- Slight
- Moderate
- Severe

5. **Difficulty focusing** *
Mark only one oval.
- None
- Slight
- Moderate
- Severe

6. **Increased salivation** *

Mark only one oval.

- None
- Slight
- Moderate
- Severe

7. **Sweating** *

Mark only one oval.

- None
- Slight
- Moderate
- Severe

8. **Nausea** *

Mark only one oval.

- None
- Slight
- Moderate
- Severe

9. **Difficulty concentrating** *

Mark only one oval.

- None
- Slight
- Moderate
- Severe

10. **Fullness of head** *

Mark only one oval.

- None
- Slight
- Moderate
- Severe

11. **Blurred vision** *

Mark only one oval.

- None
- Slight
- Moderate
- Severe

12. **Dizzy (eyes open)** *

Mark only one oval.

- None
- Slight
- Moderate
- Severe

13. **Dizzy (eyes closed)** *

Mark only one oval.

- None
- Slight
- Moderate
- Severe

14. **Vertigo** *

Mark only one oval.

- None
- Slight
- Moderate
- Severe

15. **Stomach awareness** *

Mark only one oval.

- None
- Slight
- Moderate
- Severe

16. **Burping** *

Mark only one oval.

- None
- Slight
- Moderate
- Severe

17. **Other**

Technique Evaluation (Speed Control)

18. **It was easy to walk inside the blue circle.** *
Mark only one oval.

	1	2	3	4	5	6	
Totally Disagree	⬭	⬭	⬭	⬭	⬭	⬭	Totally Agree

19. **I felt safe inside the blue circle** *
Mark only one oval.

	1	2	3	4	5	6	
Totally Disagree	⬭	⬭	⬭	⬭	⬭	⬭	Totally Agree

20. **It was easy to indicate the direction of movement** *
Mark only one oval.

	1	2	3	4	5	6	
Totally Disagree	⬭	⬭	⬭	⬭	⬭	⬭	Totally Agree

21. **It was easy to control the speed** *
Mark only one oval.

	1	2	3	4	5	6	
Totally Disagree	⬭	⬭	⬭	⬭	⬭	⬭	Totally Agree

22. **It was easy to move around in the environment.** *
Mark only one oval.

	1	2	3	4	5	6	
Totally Disagree	⬭	⬭	⬭	⬭	⬭	⬭	Totally Agree

23. **It was easy to get to the rings** *
Mark only one oval.

	1	2	3	4	5	6	
Totally Disagree	⬭	⬭	⬭	⬭	⬭	⬭	Totally Agree

24. **It was easy to deflect obstacles** *
Mark only one oval.

	1	2	3	4	5	6	
Totally Disagree	⬭	⬭	⬭	⬭	⬭	⬭	Totally Agree

25. **I felt like I was controlling the body I was seeing** *
Mark only one oval.

	1	2	3	4	5	6	
Totally Disagree	◯	◯	◯	◯	◯	◯	Totally Agree

26. **I felt that the virtual body was my body** *
Mark only one oval.

	1	2	3	4	5	6	
Totally Disagree	◯	◯	◯	◯	◯	◯	Totally Agree

27. **I felt that the virtual body was located in the same place as my real body** *
Mark only one oval.

	1	2	3	4	5	6	
Totally Disagree	◯	◯	◯	◯	◯	◯	Totally Agree

28. **I was afraid of heights** *
Mark only one oval.

	1	2	3	4	5	6	
Totally Disagree	◯	◯	◯	◯	◯	◯	Totally Agree

29. **I felt tired after using** *
Mark only one oval.

	1	2	3	4	5	6	
Totally Disagree	◯	◯	◯	◯	◯	◯	Totally Agree

30. **Comments**