Non-visual Locomotion Techniques for Mobile Virtual Environments

João Pedro Matias Sequeira

Thesis to obtain the Master of Science Degree in Information Systems and Computer Engineering

Supervisor(s): Prof. Hugo Miguel Aleixo Albuquerque Nicolau

Examination Committee
Chairperson: Prof. Prof. José Carlos Martins Delgado
Supervisor: Prof. Hugo Miguel Aleixo Albuquerque Nicolau
Member of the Committee: Prof. Carlos Duarte

Outubro 2017
Dedicado á minha família...
Acknowledgments

Upon completion of this dissertation, I would like to thank everyone that made this work possible by contributing to it and by supporting me along the way.

Firstly, I would like to express my deepest and sincerest gratitude to my supervisor, Professor Hugo Nicolau, for entrusting me with this project and also for providing me with the guidance that allowed me to complete it. Thank you for the continuous support, motivation and patience throughout all the phases of this work. I could not have asked for a better supervisor. Thank you Professor Hugo.

I would also like to thank my family, for the unconditional love and support that allowed me to reach my goals up until now. Thank you for showing me the meaning of hard work, strength and love throughout all my years of study, including writing this dissertation.

A special thank you to everyone at the Raquel and Martin Sain Foundation who volunteered and participated in the evaluation sessions. Thank you for the enthusiasm and openness showed in the evaluation sessions. I would also like to thank Doctor Carlos Bastardo, for helping in recruiting volunteers and for always showing himself available to help for anything that came up.

Thank you to my friends and IST colleagues, in particular to Carlos Figueiredo, Filipe Almeida, Duarte Cruz, João Colaço, Francisco Dias & José Lino, for all the motivation, support, companionship and relaxed moments they provided me with which helped me get through the most difficult phases of this work.

In retrospect, this work would not have been possible without all these people and many more. There are no words that could describe all the gratitude and appreciation I feel for everyone that made this project possible.

Thank you everyone. For everything.
Resumo

Já foi estabelecido por numerosos estudos que indivíduos cegos podem beneficiar bastante de realidade virtual obtendo informação sensorial adicional para navegar e explorar ambientes desconhecidos e criar um mapeamento mental preciso desse mesmo ambiente. Graças à introdução de smartphones de baixo custo no mercado consumidor, a realidade virtual atingiu assim o potencial para a adesão global. No entanto, para uma experiência totalmente imersiva e benéfica na realidade virtual, técnicas de locomoção eficientes e naturais devem ser implementadas. Embora este campo tenha sido alvo de muitos estudos, há uma falta de pesquisa no tema de locomoção não visual em realidade virtual móvel.

Nesta tese, nós tentamos preencher essa lacuna desenhando e implementando três técnicas diferentes para a locomoção não visual, um D-Pad virtual, um D-Pad virtual com orientação fisicamente consciente e uma técnica tilt-to-walk fisicamente consciente. O nosso objetivo é determinar como a orientação física e diferentes métodos de input (gesto 3D versus input por toque) influenciam a imersão dos usuários, bem como a eficiência de navegação e exploração para a realidade virtual em contextos móveis. Para o efeito, realizamos uma avaliação de exploração e uma de navegação para cada técnica. Os nossos resultados não mostraram diferenças significativas nos desempenhos de cada técnica, apesar de uma quantidade significativa de participantes preferir o D-Pad com orientação fisicamente consciente.

Palavras-chave: Pessoas invisuais, Locomoção, Realidade Virtual, Smartphone
Abstract

It has already been established by numerous studies that blind individuals can greatly benefit from virtual reality for additional sensory information input for navigating and exploring unknown environments and creating an accurate mental mapping of that same environment. Due to introduction of low cost smartphone adapters, virtual reality has reached the potential for mainstream adoption. However, for a fully immersive and beneficial experience in virtual reality, efficient and natural locomotion techniques must be implemented. Although this field has been the target of many studies, there is a lack of research regarding non-visual locomotion in mobile virtual reality.

In this thesis, we attempt to fill this gap by designing and implementing three different techniques for non-visual locomotion, a virtual D-Pad, a virtual D-Pad with physically-aware orientation and a tilt-to-walk physically aware technique. Our goal is to determine how physical orientation and different inputs (3D gesture versus touch input) influence user immersion as well as navigational efficiency for virtual reality in mobile contexts. For this purpose, we conducted a exploration and navigation evaluation for each technique. Our results showed no significant differences in the performances of each technique despite a significant amount of participants preferring the Directional D-Pad with physically-aware orientation.

Keywords: Blind, Locomotion, Virtual Reality, Smartphone
Contents

Acknowledgments ........................................ v
Resumo ........................................ vii
Abstract ........................................ ix
List of Tables ........................................ xiii
List of Figures ......................................... xv
Nomenclature .......................................... xvii
Glossary ........................................ 1

1 Introduction 1
1.1 Problem ........................................ 1
1.2 Goal ........................................ 2
1.3 Contributions .................................... 2
1.4 Document Structure .............................. 2

2 Related Work 5
2.1 Cognitive Mapping by Blind People ............... 5
  2.1.1 Cross-modal plasticity in the occipital cortex .... 5
  2.1.2 Structural Reorganization of the Brain in Blindness .. 8
  2.1.3 Navigational Aids used by Blind People .......... 9
2.2 Non-visual Virtual Environments .................. 10
  2.2.1 Adapted Virtual Worlds ...................... 11
  2.2.2 Virtual Reality Approaches for Blind Navigation and Exploration .... 15
2.3 Mobile Virtual Locomotion ....................... 20
2.4 Discussion ..................................... 22

3 Non-visual Locomotion Techniques 25
  3.1 Techniques’ Design .............................. 25
  3.2 Multi Directional D-Pad ......................... 26
  3.3 Physically Aware 2-Directional D-Pad .......... 27
  3.4 2-Directional Tilt-to-Walk ...................... 28
# List of Tables

3.1 Our virtual locomotion techniques. A non-physically aware tilt-to-walk is not considered in this study due to its limitations regarding unnatural wrist rotations and virtual movement precision. ................................................................. 26

5.1 Participants profile. ................................................................. 42
5.2 Order of presentation for Techniques. ................................. 45
5.3 Exploration Order for each Technique. ............................... 47
List of Figures

2.1 Grand-average ERP’s to standard and deviant tones in sighted, early and late-blind individuals from frontal (Fz) and occipital (Oz) electrodes [14]. We can see that occipital cortex activity was largely increased in the attend condition, where participants were asked to focus on deviant tones in the binaural recording. .......................................................... 6

2.2 Results regarding Binaural localization tasks [31]. EBSP – Early Blind with Superior Performance. EBNP – Early Blind with Normal Performance (A) Sagittal and horizontal slices of activated cortices during the binaural task. (B) Graphical comparison between the cerebral blood flow to the occipital cortex of each groups during the task. (C) Behavioral comparison between the number of errors of each group across each iteration of the task. 7

2.3 Graphical comparison between early, late-blind and sighted individual’s hippocampal structures [12] .............................................................................................................. 9

2.4 Accessibility options in PowerUp [29]. ............................................................................................................. 12

2.5 Scene interaction using TextSL [11]. .................................................................................................................. 14

2.6 Controlled walk in PowerUp [29]. .................................................................................................................. 15

2.7 Photograph of a user using BlindAid [27]. The user hears spatialized audio while receiving haptic feedback from the Phantom haptic device. .................................................................................................................. 16

2.8 An object’s sound radiation in a sonified city map [13] ......................................................................................... 18

2.9 Navigation task in a study of VR-STEP [28]. Users had to walk around the grey obstacle in order to get to the blue target area. .................................................................................................................. 21

2.10 The Myo Armband [19] .................................................................................................................................. 22

3.1 Multi Directional D-Pad. .................................................................................................................................. 27

3.2 Physically Aware 2 Directional D-Pad. The user places his finger on the screen and slides forward or backwards to initiate virtual movement. .................................................................................................................. 28

3.3 Neutral position’s tilt angle (45°) and virtual locomotion thresholds (30° and 60°). ........................................... 29

3.4 Tilt angles for max locomotion speed. .................................................................................................................. 30

4.1 System architecture .......................................................................................................................................... 32

4.2 Orientations that produce Haptic feedback. ..................................................................................................... 36

4.3 Unity Standard Assets - Third Person Character model ..................................................................................... 37

4.4 Structures object’s composition ....................................................................................................................... 38
5.1 Huawei P10 smartphone running our prototype application .......................... 43
5.2 Sony ZX220BT Wireless Headphones ......................................................... 43
5.3 Tactile Maps for the Exploration Evaluation ................................................ 44
5.4 Virtual Training Room .................................................................................... 46
5.5 Exploration Scenes ....................................................................................... 47
5.6 Navigation Scenes ......................................................................................... 48
5.7 Average Duration of Exploration Tasks. ........................................................ 49
5.8 Average Distance Covered during Exploration Tasks per Implementation. .... 49
5.9 Average Number of Pauses of Exploration Tasks. ......................................... 50
5.10 Tactile Map Evaluation of Exploration Tasks. .............................................. 51
5.11 Duration of Navigational Tasks per Technique. ......................................... 52
5.12 Relation between Optimal Distance and Distance Covered. Values closer to 1 show better navigational performance. ................................................................. 53
5.13 Number of collisions per Technique During Navigation Tasks. ................. 54
5.14 Participants’ rating of evaluation tasks’ ease per technique. ......................... 55
Nomenclature

Greek symbols

χ  Chi square statistic
σx  Standard Deviation
Fx,y  F Statistic
p  p-value
Chapter 1

Introduction

Spatial orientation and navigation is the result of a combination of sensory and cognitive skills in which we form a mental mapping of our surrounding environment based on the information that our sensory channels provide us [18]. Most of the information of our surroundings comes primarily from the visual channel. People that are blind lack this information and therefore rely on compensatory channels to fill the gaps of information caused by the absence of sight. Unfortunately, even with compensatory channels, navigation in unknown environments by blind people can be as difficult as it is unsafe often leading to injury or loss of possessions.

For years, the primary aids for blind people have been the white cane and the guide dog but recent advances [33] and studies in virtual reality provide us with evidence that blind people can benefit from a virtual environment to gather sensory information and create an accurate mental mapping of an unknown environment in a safe and reliable way by using 3D sound as well as haptic information [17]. In recent years, research regarding virtual reality as well as blind accessibility of several virtual worlds that were initially designed for sighted individuals, has increased [29, 32, 11]. Allied with the introduction of low cost smartphones that can render virtual environments with 3D sound efficiently, we expect that mobile virtual reality shows potential for helping blind people in the navigation and exploration of unknown environments.

1.1 Problem

Virtual reality is able to supply blind individuals with enough additional sensory information for an accurate mental mapping of a physical environment using 3D sound and haptic feedback [17]. However, in mobile contexts virtual navigation and locomotion can become problematic. If the means of navigation and locomotion do not effectively transmit a natural sense of motion, then the exploration as well as the mental mapping of the virtual environment are compromised due to cognitive overload.

Presently, the lack of research into non-visual virtual locomotion techniques in mobile contexts, particularly in input techniques, makes it difficult to develop efficient virtual mobile navigational aids for blind people.
1.2 Goal

Recent research tells us that people that are blind outperform sighted individuals in sound source localization tasks, among other navigational tasks [12, 31]. Most solutions to the problem presented require customized hardware or offer many limitations, making it difficult for that solution to become widely used [13, 1, 10, 24, 4]. Taking this into account we intend to make blind navigation easier in mobile virtual environments by attempting to fill the gap of research into non-visual virtual locomotion in mobile contexts.

So, the main goal of this master thesis is to design, implement and compare non-visual locomotion techniques for virtual mobile navigation and study the influence of each technique’s characteristics in providing a beneficial virtual experience.

1.3 Contributions

This dissertation provides the following contributions:

- **State of the Art in virtual reality for blind people.** We present several works regarding blind accessibility in virtual worlds. We also present works into current navigation and exploration tools for blind people. To complete our research, we also provide insight into virtual locomotion techniques for sighted people.

- **Design and implementation of 3 non-visual locomotion techniques for mobile virtual reality.** We design each technique by taking into account the approaches presented in our state of the art research with the goal of providing an immersive virtual experience to blind users.

- **Evaluation of each of our blind locomotion techniques through user testing.** We provide blind users with a mobile application prototype where they can use each of our techniques with the goal of evaluating each technique in terms of navigational and explorational performance.

1.4 Document Structure

In the following chapter, we present related work regarding the approached problem. We start by differentiating cognitive and sensory skills between blind and sighted individuals. After which we also research into accessibility strategies in visually-focused virtual worlds, followed by works in different virtual reality approaches for exploration and navigation of unknown environments by blind people. We also dive into some interesting techniques used for sighted locomotion in virtual environments.

After discussing the key points of our related work research, we present our solutions for non-visual locomotion in mobile contexts in chapter 3 followed by a detailed description of the implementation of each technique and the technologies used in implementing them in chapter 4.

In chapter 5, we present our evaluation process for each of our solutions and offer our results while discussing their relevance through statistical analysis.
Finally, in chapter 6, we offer our conclusions and the objectives we have achieved before discussing about possible future work.
Chapter 2

Related Work

In this chapter, we start by presenting some insight on the cognitive mapping of spatial information by blind people and their hardships in achieving an accurate perception of their surrounding space based on previous studies. We then present some approaches about the state of the art in virtual environments for blind people and their approaches on giving an accurate representation of the space. Finally, describe previous work for locomotion in virtual environments in smartphones. In the end of this chapter we discuss the relevance of these works to the problem and solution presented in this thesis.

2.1 Cognitive Mapping by Blind People

Spatial perception of our surrounding environment is acquired through information provided by our sensory channels as well as our motor and cognitive skills [18]. However, most of this information is provided by the visual channel to which blind people do not have access to. Therefore, they need to rely on compensatory channels as well as other navigational aids and strategies to compensate for the lack of visual information and to be able to navigate and acquire an accurate spatial perception of their surroundings. To understand how blind people make use of compensatory channels such as the auditory and tactile channels it is important to first understand the differences between sighted and blind people’s brains.

2.1.1 Cross-modal plasticity in the occipital cortex

In traditional neuroscience, there is a common view that the human brain is divided into cortices per each sensory modality and that each cortex processes information regarding its corresponding modality, but recent studies suggest that this view might not be fully accurate and provide evidence for a cross modal plasticity reality for the brain, in which sensory deprivation leads to a reorganization of the brain’s neurons that compensates the sensory loss by recruiting its cortex for the processing of other sensory information. In some cases, this phenomenon enables individuals to acquire extraordinary abilities based on their remaining senses [3].

Following a cross modal plasticity hypothesis several studies have shown that when the visual cortex of blind people lacks sensory information to process, the brain adapts by plastically recruiting the
visual cortex of the brain for processing cognitive tasks involving sensory information provided by other cortices such as the auditory cortex and the somatosensory cortex [6].

In order to confirm if the occipital cortex was activated in tasks requiring auditory processing, Kujala et al. [14] performed an experiment in which scalp-recorded ERP’s (event-related potentials) were measured while the participants listened to binaural sounds. The first group’s task was to ignore the sounds while reading braille (ignore condition) while the second group’s task was to count the deviant tones in the binaural sounds (attend condition). The participants involved in the experiment were individuals who were either sighted, early or late-blinds. As we can see in Figure 2.1, the results showed that the attend condition group presented increased activity in occipital cortex as opposed to the ignore condition group. There was also little difference between the occipital cortex activity between early and late-blinds which leads to believe that, contrary to common belief, cross modal plasticity does not only occur in the early development phase.

In 2008, Voss et al. [31] confirmed these findings by finding evidence of occipital cortex recruitment in a study where the occipital cortex responses were measured in early and late blinds when performing sound source discrimination tasks. The study involved binaural and monoaural sound source localization tasks where the auditory stimuli was presented to the subject in a circular array of 9 speakers while the cerebral blood flow was measured to determine which cortices of the brain were active during the exercise. In the binaural localization task (Figure 2.2) of the experiment, early and late blinds as well

Figure 2.1: Grand-average ERP’s to standard and deviant tones in sighted, early and late-blind individuals from frontal (Fz) and occipital (Oz) electrodes [14]. We can see that occipital cortex activity was largely increased in the attend condition, where participants were asked to focus on deviant tones in the binaural recording.
as sighted participants performed equally well suggesting that blind people do not have trouble constructing a three-dimensional mapping of the surrounding space using sound. In cases where the sound source was located in more peripheral positions, it was noted that the blind individuals outperformed their sighted counterparts (Figure 2.2 C). In the monoaural localization task of the experiment the blind group outperformed the sighted group suggesting that blind people are capable of detecting subtle auditory cues that sighted people cannot due to the lack of processing power in their brains. The results also showed an increased blood flow to the occipital cortex in the groups composed of blind people during the exercise which lead to conclude that blindness does indeed cause plastic changes in the occipital cortex and that one of this changes is the recruitment of the occipital cortex for enhanced auditory spatial localization (Figure 2.2 B). There was also evidence that these changes are more pronounced in early blinds than in late blinds, since the number of occipital cortices recruited during the exercise was lower in late blinds than in early blinds.

**BINAURAL TASK**

Figure 2.2: Results regarding Binaural localization tasks [31]. EBSP – Early Blind with Superior Performance. EBNP – Early Blind with Normal Performance (A) Sagittal and horizontal slices of activated cortices during the binaural task. (B) Graphical comparison between the cerebral blood flow to the occipital cortex of each groups during the task. (C) Behavioral comparison between the number of errors of each group across each iteration of the task.

Although it’s clear that the occipital cortex is in fact recruited for the processing of other sensory in-
formation, this recruitment is very dependent on the compensatory behaviors of blind people as shown in an experiment by Kupers et al. [15], where repetitive transcranial magnetic stimulation (TMS) was applied to the occipital cortex of blind people while reading Braille, in order to disrupt the occipital cortices functions. The results showed that the disruption of the occipital cortices function impaired the performance of proficient Braille readers but that applying the same treatment to the primary somatosensory cortex had no effect on the performance of blind people while reading Braille. These findings also provide evidence for a relation between the recruitment of the occipital cortex and compensatory behaviors in the blind like Braille reading.

2.1.2 Structural Reorganization of the Brain in Blindness

The occipital cortex is not the only area of the brain that goes through changes with blindness, some regions of the brain that are responsible for somatosensory and auditory processing, among others, also go through several changes. In a study by Pascual-Leone et al. [23] when comparing somatosensory evoked potentials between right handed blind and sighted people it was found that, regarding tactile acuity, there was an expansion and reorganization of the right index finger representation in the somatosensory cortex of blind people which contrasted with the left index finger results. They concluded that this adaptation allowed the blind group to have more enhanced processing and skill when reading braille.

In terms of auditory cortical areas, Elbert et al. [9] recorded and compared tonotopic maps as well as tone burst responses of blind and sighted individuals, by measuring audio-evoked magnetic fields in the brain of blind and sighted individuals while providing auditory stimuli, and found that the blind brain exhibited an expansion in areas of the brain that respond to auditory stimuli. This finding leads us to confirm that blind people do not receive more auditory stimulation than their sighted counterparts but that it is this expansion that enables blind individuals to outperform sighted individuals in auditory localization as well as in auditory signal response latency.

Other structural differences found in the blind people’s brains when compared to the sighted were recorded by Fortin et al. [12] in a experiment where they investigated if blind individuals presented enhanced spatial navigation when compared to sighted ones and how their performance would relate to the size of their hippocampus, a structure that plays a crucial role in navigation and spatial memory. In this experiment both blind and sighted individuals were submitted to spatial navigation tasks within a maze in an attempt to verify if the blind individuals would commit a smaller number of errors as was expected. After performing the tasks, the participants had their hippocampal volume measured in order to establish some sort of correlation with the performance of each participant in the spatial tasks. The results of the spatial navigation tasks did in fact show a smaller number of errors by the blind participants when compared to their sighted counterparts. These results correlated with the hippocampal measurements (Figure 2.3) which showed a significant difference in the overall hippocampal volume between the participants, with the blind individuals showing an increased size of the hippocampus. They concluded that blindness contributes to an adaptive expansion of the hippocampus structure that results in
enhanced navigation skills by the blind, while also noting that these changes in the hippocampus are more pronounced in early blinds.

![Hippocampal volume graph](image)

**Figure 2.3**: Graphical comparison between early, late-blind and sighted individual's hippocampal structures [12]

### 2.1.3 Navigational Aids used by Blind People

In order to rehabilitate a recently blind person as a part of society, training in Orientation and Mobility (O&M) is usually provided/recommended to teach them strategies on how to navigate through new environments without additional technological assistance beyond the primary aids already used by the blind community. As was said before, the primary navigation aids used by the blind are the white cane and the guide dog. Although widely used, these two aids provide very different experiences to their users. A recent study by Williams et al. [33] interviewed several blind individuals while seeking to understand the differences between several blind aids, among which were the white cane and the guide dog.

The white cane is classified as an obstacle detection aid whose contact with obstacles provides the
user with more additional information about their surrounding environment which is why while using this aid the blind user's strategy is to seek out obstacles and walls in order to be able to form an accurate mapping of his surroundings. However, the white cane does have its disadvantages. Despite being able to detect obstacles in the immediate vicinities, it becomes very hard to navigate around smaller obstacles due to their reduced contact surface. Blind people also reported not using the white cane due to fear of striking people or fragile objects while scanning their surroundings and fear of colliding with objects situated at a greater height than the white cane's (in this case the guide dog was preferred). Some inconveniences also extend to social interactions where blind users report being offered help by strangers when trying to concentrate on traffic cues and when trying to scan their surroundings by bumping into obstacles.

On the other hand, the guide dog is classified as an obstacle avoidance aid where the dog guides the blind user around obstacles needing only a small correction from the user in the event of becoming distracted by its surroundings. Since the guide dog's strategy revolves around obstacle avoidance this means that the user might not get a good mental resolution of the surrounding environment but can navigate faster across a room than white cane user. Unlike white cane users, dog users do not report as much inconveniences regarding social interactions and fear of bumping into people. On the contrary, since dog users did not have to rely so much on constantly scanning their environment, they did not seem so clueless and therefore were not offered as much assistance from strangers. Thanks to the results of this study, we can conclude that these two primary aids both have advantages and disadvantages and can relate to the difficulty that blind individuals have when choosing their aid. Either they choose the guide dog and navigate faster across their surrounding environment but receive less information about said environment or choose the white cane and are able to form a more accurate mental mapping of the unknown environment but take longer to learn the space while being troubled by social inconveniences and smaller obstacles.

There was also investigation regarding the use of technological aids for navigation where it was noted by almost all participants that smartphone applications for navigation were more practical than other custom devices, with the only set back being that it was tiresome to switch between explorational and navigational applications since there was not a single smartphone application that could cater to all their navigational needs.

### 2.2 Non-visual Virtual Environments

In the following section, we present works in accessibility features and adaptations for virtual worlds that did not have blind people as their target users as well as some interesting approaches and solutions that attempt to combine virtual reality with other technologies in order to support blind people in collecting spatial information about unknown environments. While presenting these works we also point out the key advantages and disadvantages of each approach and how they can be used in our own solution.
2.2.1 Adapted Virtual Worlds

Nowadays virtual worlds are used in a wide variety of fields such as gaming, communication and education by allowing users to explore and navigate an emulated environment. The representation of this environment is done by conveying information to the user's senses, with visual rendering being the primary method to communicate the state of the virtual world. In nearly all approaches, haptic and audio communication are mostly used as a secondary method to complement the visual information provided by the virtual world.

Although virtual worlds were inaccessible to the blind community for a long time, recently developed virtual worlds are being released with a set of accessibility options that enable visually handicapped individuals to enjoy a virtual reality experience.

In the following subsections, we divide works on virtual world accessibility in accessible representation, where we focus on how the virtual world state is being conveyed to the blind user, and accessible navigation, where we focus on the approach’s solution for blind navigation and mobility in the virtual world.

Accessible Representation of Virtual Worlds.

One of the biggest challenges in making virtual worlds, that rely mostly on visual information, accessible to blind people is in the representation of the virtual space via other sensory channels in a comprehensive way while at the same time not overloading these same channels with too much information. There needs to be a balance such that the information provided to the blind user is enough for an accurate mental mapping of the surrounding environment but also not too much such that he cannot concentrate on what his objectives are.

An interesting approach is found in the work of Trewin & Cavender [29] in 2007 where they were tasked with improving the accessibility of a virtual world called PowerUp. In this work, they surveyed people with a variety of disabilities about their problems/preferences with current virtual world games accessibility options, where the blind participants reported that their biggest issue was with the poor representation of the surrounding virtual space which they found lacking in some textual and audio representation. With these reports in mind, the authors decided to implement self-voicing for all HUD’s that allowed all text on screen to be spoken aloud by using the Alt-A keys to have the text read aloud and the Tab key to navigate through dialog boxes.

Regarding features for the 3D World, several accessibility features for the handicapped people were developed but the one focused on accurately representing the surrounding virtual space to people who were blind was the translation of visual information into audio stimuli which included an option that allowed users to customize the amount of extra sound effects associated with objects in the virtual world (Figure 2.4).

Two commands, the find and look were implemented for providing the blind user with tools to explore their surroundings while trying not to overload the audio channel. The find command was implemented to help blind users scan their surroundings, find objects and stay locked on to them. This feature was
very helpful to blind users for tracking moving objects since they were able to stay locked on to their target without having to continuously scan the surrounding environment. With the find command and the speech option enabled a user is able to acquire audio feedback about surrounding objects such as their name, distance and orientation regarding the user. The look command was implemented in a way that provides the blind user with a more detailed description of the virtual scene in front of them with the L key providing a description of the user's current focus and the K and the ; key providing a summary of the visible objects to the left and right of the user.

In this approach the authors noted that one of the biggest obstacles was the designing of the labels for all 3D objects due to the object's orientation and the possibility of it being partially occluded by another object which difficulted the task of accurately representing the scene that was in front of the blind user.

Terraformers launched in 2003, developed and studied by Westin [32], is also a virtual game world which aimed at providing an interactive accessible experience to both sighted and blind users. In this approach the author focused on implementing 3D sound based and voice feedback features. In order to provide blind users with an accurate representation of the virtual environment in front of them, a 3D sound sonar is provided for a rough perception of the distance to objects in front of them, this tool used pitch to indicate how far away the objects were. A 3D sound compass is also provided for user orientation in the virtual space. In Terraformers, all objects in the virtual world have voiced feedback and 3D sound icons for a good cooperation with the above-mentioned tools. Regarding the accessibility of menus, the backpack, in which the user stores helpful objects for his quest, provides a voiced menu that divides objects in hierarchies. In a survey of the virtual world, blind users were reported to have found the tools provided for the representation of the virtual space helpful and easy to use and almost all the participants managed to complete the game by making use of them.

Both these approaches make use of sound output mechanisms such as 3D sound sonars and spatial sounds in order to distinguish a limited number of objects but in most virtual worlds these approaches do not scale well due to the large number of objects and avatars such as the case of the Second Life virtual world.

With this predicament in mind Folmer, et al. [11] designed TextSL, a command-based virtual world
interface that works with Second Life by using synthesized speech. When using TextSL, blind users interact with the virtual world solely using written text to query about the state of the virtual environment in front of them and to execute an action such as sitting or greeting another avatar. Although it might seem a tiresome solution for blind users to have to write on the command interface every action they wish to execute in the virtual environment, TextSL supports scripts for users to be able to execute regular queries and actions more efficiently like finding and opening a door. TextSL also differentiates itself from the other above-mentioned approaches by not just focusing on one object at a time, instead the user can interact with any avatar or object in sight by issuing commands through the command interface without having to search a list of visible objects, locking on to the desired object and then executing an action over it. Upon moving to another location or querying the virtual environment state, with the “describe” command, TextSL returns the number of avatars and objects in the virtual scene and users can then query on avatar and object sets to learn more about that virtual element and interact with it (Figure 2.5). In order to not overwhelm the user with information when the presented environment is densely populated the authors added a summarizer to TextSL, where all objects in a range around the user are retrieved and prioritized by proximity and clarity of description.

**Accessible Navigation of Virtual Worlds.**

Accessible navigation and movement in a virtual world is also a troublesome topic when working on making virtual worlds accessible to blind people. Most virtual worlds depend on visual cues for effective movement and orientation throughout the virtual environment such as looking at the user’s objectives at distance. Due to the absence of vision in blind individuals it becomes difficult to give them a sense of movement in a virtual world while also maintaining their orientation when navigating towards their objective.

In the work performed in PowerUp by Trewin & Cavender [29], the authors also worked on improving accessibility in virtual movement inside the platform. In their approach the navigation of the user’s avatar is controlled through the keyboard, with the arrow keys controlling the orientation and the WASD keys controlling the movement of the avatar. But for a more effective navigation simply controlling the avatar through the keyboard was not enough so a controlled walk function (Ctrl+W) was implemented in which the player’s avatar keeps its target objective in focus (focused through the find function) while walking towards it. In order to give a sense of movement to the blind user footstep sounds were added during the walk and audio feedback when an event causes the controlled walk to end such as arriving at the target objective or bumping into other players (Figure 2.6).

They also noted that the fast travel function was also essential for accessibility of a virtual world since it facilitates long scale navigation which was where they reported that blind people had most difficulties in.

In Terraformers [32], a simpler approach for movement and navigation was used with the ASWD keys being used for avatar movement and the numeric keyboard for avatar direct orientation in 8 directions with numeric key “2” being assigned to the north direction. Again, footsteps sounds were provided for giving a sense of movement to the users while navigating the virtual space. The users were also supplied
with a GPS tool which provided the exact positioning of objects in an area as well as the user’s avatar.

The simpler approach for navigation in Terraformers was well received and blind users were able to navigate the virtual world efficiently with some users even claiming that the game was too easy. The success with this approach was probably due to the several tools available to the users which focused on the representation of the virtual space that supplied enough information for efficient navigation of the virtual environment without having to provide additional features regarding avatar movement and navigation.

For TextSL, Folmer, et al. [11] focused on providing blind users with more freedom of choice regarding navigation and interaction by not focusing on one object at a time and being able to interact with any object in the virtual scene. Navigation using TextSL can be achieved by using the arrow keys but this approach comes with the inconvenience and time-wasting task of finding a path that is not obstructed, even with TextSL notifying the user every time a collision occurs. Although users can always circumnavigate obstacles using the arrow keys and the feedback provided, TextSL also provides a module that
enables collision free navigation where the blind user can just make use of the written “move” command, for example “move north 100”, for TextSL to plot a collision free route to the objective. In case the user does not arrive at the objective in the estimated time limit, they are teleported to the objective’s location.

Other navigation commands to facilitate blind user’s navigation within Second Life while using TextSL also include “fly”, “teleport” and “follow” commands.

### 2.2.2 Virtual Reality Approaches for Blind Navigation and Exploration

For an accurate mental mapping and efficient navigation of an unknown environment, blind people rely on spatial information provided by other sensory channels. Under the assumption that by supplying enough compensatory perceptual information about an environment could provide the means for an effective acquaintance of an unknown environment by people who are blind, virtual reality approaches for blind people’s navigation and exploration have been increasingly studied and explored.

These studies gave way to the implementation of a few navigational aids that range from preplanning aids, where the blind user is able to explore the unknown environment from a safe place, to in situ aids, where the aid provides the blind user with additional sensory information about his surroundings for a more efficient and accurate exploration and navigation.

#### Preplanning Virtual Reality Blind Aids.

Preplanning aids for unknown environment exploration and navigation offer the blind user the possibility of acquiring spatial information about an environment prior to the arrival at said environment. This approach consists of a safer way of exploration by the blind user although providing a lower spatial in-
formation resolution of the environment. Other issues regarding these approaches hang on the difficulty of obtaining updated spatial information about the explored environment. Some examples of this type of blinds aids are tactile maps, physical models and verbal descriptions of the environment [16].

In 2004, Lahav & Mioduser [17] built a multi-sensory virtual environment (MVE), which attempted to provide information about unknown environments via haptic and auditory stimuli to blind users. The MVE developed simulated a single room with non-spatialized sounds and haptic feedback provided by a Sidewinder Force Feedback Pro Joystick. This model was later improved and renamed to BlindAid in 2010 by Schloerb, et al. [27] where 3D spatialized sounds were introduced along with the Phantom haptic device substituting the Sidewinder haptic device.

The BlindAid system is able to run on personal computer providing virtual reality environments which blind users can explore with a haptic device and stereo headphones (Figure 2.7).

![Figure 2.7: Photograph of a user using BlindAid [27]. The user hears spatialized audio while receiving haptic feedback from the Phantom haptic device.](image)

In this environment, avatar motion is achieved by moving the Phantom’s stylus which is tracked by the BlindAid system and then mapped to the virtual environment. Besides avatar movement, the Phantom also serves to provide force feedback to the user in order to facilitate the spatial mapping of the environment in way that is similar to the white cane already used by the blind community. Due to the reduced range of motion of the Phantom the authors implemented Zoom features in the BlindAid for a better resolution and relative distance between objects when emulating a real-life workspace. The BlindAid is also able to simulate texture through the Phantom by providing horizontal force-feedback when the blind user interacts closely to a given surface. 3D spatialized sound was also added to BlindAid in order to give the user a sense of orientation and to be able to discern the distance and direction of objects as if he were standing in the location of his avatar.

In their study the authors evaluated blind participant’s spatial cognitive mapping while using the BlindAid system for exploring a virtual environment and then describing them verbally and through the use of a modelling kit. The participants were also evaluated on their ability to transfer the spatial informa-
tion learned through the BlindAid to a physical workspace. This was achieved by emulating a real-world environment that the participants would explore through the BlindAid and then evaluating their capacity to navigate through the physical environment.

The results showed promising results for the approach with blind users being able to, not only operate the system with their preferred settings but also being able to use learned spatial knowledge to navigate the physical environments. An interesting curiosity was that the participants, being blind, found the zoom concept completely foreign although they manage to use it efficiently once the feature was understood. Participants also reported that the spatialized sound was helpful in keeping them oriented in the virtual environment by providing spatial cues that they could use as landmarks. In conclusion, this approach attempted to provide blind users with means to explore an unknown environment by supplying compensatory haptic and audio information, and succeeded in producing a preplanning aid that could be used in rehabilitation of newly sighted individuals but is not scalable enough to make available to the whole blind community.

In the work of Heuten, et al. [13], the authors implemented a system that aimed at making city maps accessible to blind users for exploration and preplanned navigation. Their system received semantic information from vector maps or bitmaps of the environment, and used that information to make a 3D sonified model of the map that blind users could use to navigate and explore. In their approach, relevant geographic objects, such as parks or buildings, are associated with a distinct characteristic natural sound for that area. The sound associated with each object is louder the closer the user is to that object. Upon gathering all the semantic information and associating the sounds to each geographical object, the system provides a 3D virtual sound room with all geographical objects which can be explored by moving a virtual listener with a mouse or a digitizer tablet. During the exploration, the environmental sounds are being played continuously in order to allow the blind user to better perceive the spatial relations between geographical objects and to be able to navigate the environment more efficiently by hearing many objects at once which facilitates search and exploring tasks. To avoid auditory overload and for the user to be able to locate and interpret simultaneous sounds, the system assigns to the virtual sound objects a sound radiation which determines the area, centered in the object, in which the virtual user must be in to hear the object (Figure 2.8). When moving in the 3D virtual sound room created by the system, all sound source's position is updated accordingly in order to help the blind user develop an accurate spatial mapping while exploring the environment.

Evaluation of this system was performed with blind participants exploring, with no time limit, the virtual 3D sound room created by the system from a map that had 2 lakes and 6 parks using a digitizer tablet. The participants were then asked to draw a cognitive model of the map by using a foil which was put on the same tablet that they had used in the exploration task.

The results were very good with most of the blind users being able to find and differentiate the 2 lakes and 6 parks.

In this approach the objective of providing a way for blind users to make use of a city map was achieved but we cannot say that this approach is without limitations. For example, in the evaluation the blind participants reported that they had a hard time determining the extent and borders of each
geographical object. There was also an issue with avatar motion in which the authors recommended using a digitizer tablet over the computer mouse due to the difficulties that blind users have with relative pointer devices.

**In-Situ Navigational Blind Aids**

In-Situ virtual reality aids provide additional spatial information to the blind user while exploring the environment. Most examples of in-situ aids resort to obstacle detectors [24] and embedded sensors in the environment such as Bluetooth Low Energy beacons [1] or Radio Frequency Identification tags [10]. But these blind aids also have some limitations that make them less likely to be used by the blind community. One of the limitations, when using obstacle detectors, corresponds to the fact that while using these aids, the blind user must be the one to gather spatial information about the explored space by a trial and error method which does not improve the blind user's lifestyle since this is much like when he uses the white cane. Another problem with in-situ aids concerns safety of use, for most of these aids make use of sound to convey spatial information to the blind user via earphones which could distract the user from his surroundings giving way for accidents to happen. Nevertheless, some works take another approach by using virtual reality in an attempt to provide reliable in-situ blind aids.

In a recent study, Blessenohl, Simon, et al. [4], presented work for a new system to help facilitate blind mobility in indoor environments by using the input from a depth camera and mapping it to a spatialized sound environment. In this study, their prototype required that the blind user be equipped with a camera that was connected to a laptop which processed the spatial information and generated a 3D soundscape that the user could hear with stereo headphones. The system was capable of detecting the boundaries of the room as well as obstacles that were in the user's way. When the user got closer than 1.5 meters to an obstacle the system would respond by increasing the volume of the auditory cue associated with that...
obstacle. This approach was interesting since it used spatialized sound for indoor mobility, differentiating itself from the embedded sensors mainstream technology.

Blum, et al. [5] formulated that for a system to be deployable while at the same time be able to deliver a rich experience of an environment it has to run on a commodity device, rely on a preexisting worldwide point of interest database, and be able to render the environment that is superior to the mainstream spoken text playback method. With these 3 factors in mind, they developed the “In Situ Audio Services” (ISAS), an application that can run on a smartphone device using the Google Places API for supplying spatial information of the user’s surroundings. The environmental information received from Google Places is rendered into a spatialized audio scene that allows the user to hear his surroundings through the use of headphones. The authors suggest in their work that the user wear bone conduction earphones to not cause the sound nodes to distract the user from his current surroundings. The spatialized audio in ISAS is coupled with auditory icons for each point of interest in the environment in an attempt to give the user shorter and clear indications of multiple points of interest in the vicinity simultaneously. The ISAS was designed with blind exploration in mind and is equipped with 2 exploration modes, the Walking mode, and the Stop & Listen mode.

The Walking mode engages when the smartphone is in a vertical position, for example when the user has it in his pocket. This mode has 2 optional exploratory mechanisms, the radar and the shockwave mechanisms. During Walking mode, the radar mechanism plays sounds nodes to the front and the sides of the user in a clockwise manner. To avoid auditory overload in dense areas, the radar sweeping is slowed in order to give the user time to process each sound node played. The shockwave mechanism on the other hand, plays sound nodes by proximity-first. The user can touch the screen for detailed information about the currently played sound node and can also swipe left to play the last heard sound node.

The Stop & Listen mode engages when the smartphone is in a horizontal position. In this mode, the user can actively explore his surrounding by swiping his finger up and down the touch screen with the bottom of the screen being his immediate surroundings and the top being up to 150 meters forward. Upon crossing a sound node, while swiping through the screen, the auditory icon corresponding to that sound node is played. To compensate for the user’s difficulty in isolating a single location while hearing his surrounding’s auditory icons in this mode the user can tip the smartphone to the side with his finger still on screen to get detailed spoken information about the closest 4 locations.

This approach is interesting in the way that it focuses on blind exploration rather than navigation while at the same time circumventing the issue of most in situ aids, that have the user constantly gathering spatial information about the environment, by utilizing a geographical content database that provides the spatial information to the application. The authors noted that most limitations in the ISAS were due to sensor reliability that could affect user orientation in the case of the compass, and user positioning in the case of the GPS but these issues vary from smartphone to smartphone and therefore were outside of the scope of the system.

While ISAS utilizes a GPS signal to track the user’s position and use it to update the position of the virtual sound nodes, the system becomes unusable in the lack of a GPS signal and therefore making
it inadequate in most situations where the user is indoors. For navigation and exploration indoors, a GPS signal would not have the precision required for tracking the more refined movements of the user, therefore in such situations most approaches rely on embedded sensors in the environment.

Regarding in-situ blind aids, in our research we found that most state of the art solutions are based on embedded sensors technologies and not as much on virtual reality due to being more reliable in terms of localization precision which is essential for a safe and efficient in-situ navigation for people who are blind.

2.3 Mobile Virtual Locomotion

Until recent years, virtual reality was mostly associated with the use of expensive and bulky devices which were not normally available to the average consumer [2]. However, currently with the introduction of low-cost virtual reality head mounted displays (HMD), this notion is gradually dying and being replaced with a nomadic-like virtual reality concept where a user is able to use a mainstreamed smartphone adapter such as Google Cardboard or Samsung’s Gear VR to transform any smartphone into a head mounted virtual reality display.

Although being innovative, this new approach also comes with some new challenges such as input techniques for task execution in the virtual environment, which becomes harder since the smartphone is located inside the adapter and therefore its touchscreen is unavailable to the user [8]. Virtual locomotion is one of such challenges and it is considered a key aspect that influences virtual reality immersion [20].

The simplest and most common techniques for locomotion in virtual reality is through the use of a joystick controller. Improvements to this approach were made by adding control over locomotion speed to the joystick and enabling physical rotations of the user’s body for directing orientation, which showed better user performance in navigation tasks [26]. But while using a joystick or controller can be a simpler approach it does not translate to a better user immersion in the virtual environment when compared to other techniques. It has been shown that a reliable way to increase immersion and decrease motion sickness in virtual environments is natural locomotion (locomotion based on the walking metaphor) [20]. Therefore, several recent solutions were developed using natural locomotion such as real walking, where the user’s movements in the real space are optically tracked by a system which translates them to the virtual world. Although being the most effective and natural way for locomotion, this approach is limited by the fact that the virtual and physical environments must be of the same dimensions. Other approaches use omnidirectional treadmills but these devices are costly and too bulky to make sense in a mobile context [7]. Currently off-the-shelf smartphones come with several inertial sensors which can be used to provide a more natural way of virtual locomotion while wearing a HMD.

With the idea of natural locomotion as a basis and smartphone inertial sensors as conduit, Tregillus, Sam, and Eelke Folmer [28] developed VR-STEP, a walking-in-place implementation that uses the smartphone’s accelerometer for step detection. Whenever the accelerometer value passes a certain threshold it detects a possible step. In order to provide a greater sense of immersion and to decrease virtual stopping latency, the authors also implemented stride speed by measuring the time between steps. In their
evaluation of the implementation they compared user’s performance during navigation tasks (Figure 2.9) while using the VR-STEP implementation versus a Look-down-to-move implementation, in which the user looks down in order to toggle auto walk. For the evaluation of the navigation tools the locomotion speed was deactivated in order to not influence the time the users took to get to their objective.

Figure 2.9: Navigation task in a study of VR-STEP [28]. Users had to walk around the grey obstacle in order to get to the blue target area.

Results showed that users found the VR-STEP more intuitive and easier to use but found it difficult to stop due to the gliding effect, an issue that is solved by activating locomotion speed. In their work the authors point out that stopping and starting walk latency present great challenges for providing an immersive and fluid virtual experience.

In the work of McCullough, Morgan, et al. [19], we find another locomotion technique based on the natural walking metaphor. In his work, McCullough, Morgan, et al. [19] devised an arms-swinging method to enable virtual locomotion. The movement of the arms is equal to the one users already perform when walking in a physical environment. For arm movement detection, the authors made use of the Myo armband (Figure 2.10), a device which has four sensors, a EMG sensor, a three-axis gyroscope, a three-axis accelerometer, and a three-axis magnetometer. The authors suggested that the user place this device in the forearm, close to the elbow in order to for the implementation to get better readings. Locomotion speed was also implemented which only required the user to swing his arms faster, like in a running motion, to walk faster. The authors evaluated this method by comparing it to the joystick locomotion and real walking approaches in navigation tasks and found that their locomotion technique was superior to the joystick approach and presented similar results to the real walking. This approach also did not have the gliding problem that VR-Step did since the users only needed to freeze their arms in order to stop their virtual movement.

Being 2 of the most natural and scalable solutions for virtual reality locomotion, Wilson, et al. [34] conducted a study in order to compare the arm swinging and the walking in place approaches. The authors also found that it would be interesting to compare these 2 approaches with the approach that is considered the best for immersion and spatial awareness in virtual environments which is physical walking. For this study, walking in place and arm swinging was implemented using the Myo armband,
with the armbands being placed on the ankles of the users for the walking in place approach. In the evaluation, each user explored each virtual environment using the 3 techniques and after becoming comfortable with using each technique, they were asked to perform navigational tasks similar to those conducted in the study of McCullough, Morgan, et al. [19] The techniques were evaluated in regards to mean turning error, mean walking latency and median turning error.

As expected, physical walking returned the best results of the 3 techniques in every aspect evaluated since this technique offers a more immersive experience and therefore the users are naturally comfortable with it. In regards to the other 2 techniques, walking in place was shown to be better than arm swinging in terms of turning error and slightly better results in the other 2 criteria. Although the results suggest that walking in place might be better than arms swinging, the authors point out that this might be due to the lack of training in arms swinging technique which is not as natural as walking in place. Overall, in regards to virtual locomotion the current trend is set on techniques that mimic real world locomotion. The challenge is that these mimicking approaches must aim to be as immersive as possible or they will, very easily, present an inferior navigational performance when compared to low fidelity techniques, such as the joystick approach, or the real walking technique, even though this technique is predicted to always be better than any semi-natural approach [30, 21] . Nevertheless, new approaches should strive to be as immersive as real walking.

2.4 Discussion

In our related work, we talked about the cognitive differences and compensations between blind and sighted people and were able to conclude that blind people are, in fact, superior to sighted individuals in terms of spatial mapping and navigation when the visual channel is excluded. These extraordinary skills have their root in the neural reorganization of the brain (cross-modal plasticity) that takes place in the absence of visual stimuli caused by blindness. These findings motivate the notion that blind people can make use of the additional compensatory environmental information provided by virtual reality environments in order to make a more accurate mental mapping of a physical environment. Specialized virtual environments for the blind have been researched and developed in order to provide new navigational options for the blind community. These works divide themselves in navigational aids for in situ navigation
and preplanning aids. In our research, we found that the trend with most in situ solutions is using previously placed sensors in the explored environment to provide spatial information to the system used by the blind user. This makes sense since it circumvents the problem with some navigational aids like ISAS [5] that become unusable in the absence of a GPS signal. This does not happen with Bluetooth Low Energy networks which provide better resolution of the environment when compared to virtual reality approaches.

In regards to preplanning aids for blind navigation, which in our research we found to be very scarce in the mobile context, most solutions in this field resort to 3D sound maps [13], which do not offer an accurate resolution, or stationary virtual environments that provide good information but require custom made hardware that is either expensive or is simply not available to the average consumer[1, 10, 24]. Nevertheless, as we can see nowadays with the worldwide adoption of smartphones and their capacities to render virtual environments accurately we feel that these devices can be used to develop a preplanning aid with similar environmental resolution as previously developed stationary ones. However, we also believe that these approaches have not been explored as much as in situ ones due to the lack of research regarding non-visual locomotion for mobile virtual environments. In in situ approaches, real locomotion is translated into virtual locomotion by use of a GPS signal or a relative signal to sensors in the environment but in preplanning aids these approaches are unusable since the user is not actually moving in the physical environment. Therefore, more creative virtual locomotion approaches are required to convey the sense of motion while remaining stationary.

Works in blind accessibility of visual virtual worlds provide us with some ideas for accurate and easy to use virtual locomotion approaches, such as the PowerUp [29] controlled walk and the powerful but typing intensive TextSL interface [11].

Research on visual virtual locomotion also provides with some insight for good practices in the development of virtual locomotion techniques. Throughout our research, we were able to understand that locomotion approaches based on natural locomotion provide users with a greater degree of immersion in the virtual environment. But we were also able to conclude that some of these solutions which made use of semi-natural locomotion techniques performed worse than non-natural ones due to the poor immersion provided by the approach [21].

Therefore, in this work we intend to further the research of non-visual virtual locomotion techniques by comparing non-natural and semi-natural locomotion approaches in a nonvisual mobile virtual environment in the hopes of finding adequate results for a possible future work that combines virtual reality with a blind aid approach.
Chapter 3

Non-visual Locomotion Techniques

In the previous chapter, we presented several solutions for accessibility of virtual worlds and blind aids that make use of virtual reality to provide blind users with effective means for exploration and navigation of unknown environments. Each approach has its own solution for effective virtual movement. In-situ blind aids make use of GPS features to circumvent the problem of navigation and despite this method producing acceptable results, without optimizations and a good GPS signal these solutions produce low resolution results that make these approaches unusable. Preplanning approaches have more varied solutions that range from specialized haptic devices to the more simplistic movement with arrow keys.

But as we can see, most solutions that use virtual reality for blind user’s navigation and exploration of unknown environments are based on non-mobile hardware. Also, most virtual locomotion solutions are aimed at visually focused virtual worlds and therefore depend on the visual channel which blind people do not have access to. Consequently, non-visual virtual locomotion research in mobile devices has not been explored to its full extent.

In this chapter, we present our approach for non-visual virtual locomotion in mobile virtual reality. Our objective was to differentiate and evaluate different techniques that can be used for blind user’s navigation in mobile virtual reality, with the goal of finding a virtual locomotion method that blind users can easily use without leading to cognitive overload while providing a good navigation and exploration performance.

3.1 Techniques’ Design

In our work, we have designed 3 techniques for virtual locomotion in mobile contexts, 2 D-pad based techniques that leverage the device’s screen (with and without physically aware orientation) and a tilt-to-walk technique with physically aware orientation (Table 3.1).

We make this differentiation in order to study if using the user’s physical orientation translates into a more immersive and natural virtual locomotion and if it influences navigational performance. Also, we aimed to compare touch and 3D gesture (tilting gesture) input for virtual locomotion regarding virtual reality immersion and navigational efficiency.
We excluded a non-physically aware orientation approach for the tilt-to-walk technique due to limited wrist rotation [25], which makes it unfeasible to turn the smartphone while providing little precision for the virtual movement.

For the D-Pad technique, we transformed the smartphone’s touchscreen into a virtual D-Pad that the user can use to move himself in the virtual environment. For the tilt-to-walk technique we used accelerometer values to translate the forward and backward inclination of the smartphone into virtual movement. Note that virtual locomotion speed was added to the application as was suggested by various related works [26, 22]. Virtual locomotion speed in the D-Pad approach is achieved by sliding the finger further in the direction one is moving in order to increase movement speed. In the tilt-to-walk technique, locomotion speed is calculated by the magnitude of the inclination of the smartphone.

Table 3.1: Our virtual locomotion techniques. A non-physically aware tilt-to-walk is not considered in this study due to its limitations regarding unnatural wrist rotations and virtual movement precision.

<table>
<thead>
<tr>
<th>Input / Virtual Orientation</th>
<th>Physically aware orientation</th>
<th>Non-Physically aware orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virtual Input</td>
<td>2-direction D-Pad</td>
<td>Multi-directional D-Pad</td>
</tr>
<tr>
<td>3D Gesture Input</td>
<td>2-direction tilt-to-walk</td>
<td>No implementation considered due to limited wrist rotation</td>
</tr>
</tbody>
</table>

3.2 Multi Directional D-Pad

In our non-physically aware orientation design of the D-Pad, the virtual avatar’s movement and orientation are fully controlled with a multi directional virtual D-Pad (Left side in Figure 3.1). The user is able to initiate the D-Pad by placing his finger anywhere on the smartphone’s screen which will then be the center of the D-Pad. Then the user needs only to slide his finger in the direction he wants move in to initiate the virtual movement (Right side of Figure 3.1). To increase the virtual movement speed the user only needs to slide his finger further in the direction he is moving. In order to provide users with feedback regarding their current movement speed, synchronized footstep 2D sounds are used, which play whenever the user’s avatar’s feet touch the ground. To stop moving, the user needs to slide his finger back to the neutral position or remove his finger from the smartphone’s screen. Note that the user’s orientation is always aligned with his movement direction, for example, if the user slides his finger 90 degrees to the right from the direction he was moving in, we update his virtual avatar’s orientation to 90 degrees in the right direction as we can see in Figure 3.1.

In earlier versions of the technique, the virtual movement of both D-Pad techniques was achieved by the user placing a finger in the center of the smartphone’s screen (which was the neutral position) and sliding in the direction desired. However, after some preliminary testing, we found that the task of finding
the center of the screen, without visual feedback, was not straightforward.

Therefore, in our work, we improved both D-Pad techniques by moving the center of the D-Pad to the user’s first touch position and keeping it there until he lifts the finger from the screen. This led to a more dynamic experience in which the user didn’t have the concern of trying to find the center of the screen.

3.3 Physically Aware 2-Directional D-Pad

In order to evaluate the importance of physical orientation in navigation and exploration performance of blind people in virtual environments we designed a physically aware D-Pad.

For the physically aware orientation implementation of the D-Pad, we take into account the physical user’s orientation (which is measured using the smartphone’s gyroscope) and update it in the virtual scene. In this implementation, the virtual avatar’s orientation only changes when the smartphone’s physical orientation does also. With orientation being physically managed, the D-Pad has only 2 directions, forward and backward movement (Figure 3.2). Similarly to the non physically aware implementation, to initiate the virtual movement the needs to place his finger on the smartphone's screen and slide up or down to move forward or backwards respectively with locomotion speed being determined by how far the user has slid his finger in a given direction. In order to halt his virtual movement, the user has to slide his finger to the neutral position or remove his finger from the smartphone’s screen.
3.4 2-Directional Tilt-to-Walk

One of our objectives was to also compare the influence that different inputs provided in regards to virtual reality immersion and navigation efficiency. Therefore, we designed a Tilt-To-Walk technique in order to compare touch input with 3D gesture input.

The tilt-to-walk technique consists on a solely physically aware orientation implementation. Similarly to the physically aware implementation of the D-Pad technique, in our Tilt-to-Walk implementation, the virtual avatar’s orientation is determined by the smartphone’s physical orientation. However, instead of using the screen to obtain input for virtual movement, we use the smartphone’s accelerometer to determine the angle of inclination between the smartphone and the floor and we use those values to determine the direction and magnitude of the virtual movement.

Initially, the neutral position was with the smartphone parallel to the floor, with virtual movement occurring at -5 degrees or less (for forward movement) or at 5 degrees or greater (for backward movement). However, after testing we realized that a deviation from the neutral position as small as 5 degrees al-
lowed for accidental virtual movement. Therefore we increased the deviation to 15 degrees to decrease the chances of accidental virtual movement. We also changed the neutral position’s angle with the floor from parallel (0 degrees) to an inclination angle of 45 degrees with the floor which allowed for a more natural and comfortable hand position.

Therefore, in this technique, in order to stay stationary the user needs to keep the smartphone in an angle between 30 and 60 degrees with the floor (Figure 3.3). To initiate the virtual movement the user needs only to tilt the device in an angle of less than 30 degrees (for forward movement) or in an angle of more than 60 degrees (for backward movement). Increasing the tilting angle will also increase the virtual movement’s speed. The max speed is reached at 0 degrees for forward movement and at 90 degrees for backward movement (Figure 3.4).

Figure 3.3: Neutral position’s tilt angle (45°) and virtual locomotion thresholds (30° and 60°).
Figure 3.4: Tilt angles for max locomotion speed.
Chapter 4

Implementation

In order to conduct user tests with each technique we implemented a mobile application using the Unity Game Engine. Our mobile application renders 6 virtual scenes and provides spatialized audio and haptic feedback to the user while using each technique. We make use of the 3D sound engine as well as the cross-platform libraries included in Unity to provide the spatial audio and haptic feedback.

We make use of Unity’s standard assets package of which we import a Third Person Character prefab, audio clips for footstep sounds and Mobile Single Stick Controller prefab. The Third Person Character prefab serves as our user’s avatar when using the mobile application and the Cross Platform Mobile Single Stick Controller serves an interface to provide our character and environment scripts with touch input from the user. We use cross-platform libraries to register input provided by the user via smartphone such as finger touch positions and number of fingers on screen as well as the smartphone’s gyroscope and accelerometer values to pass onto each of our techniques scripts (Figure 4.1).

In this chapter, we go into detail about the implementation of our non visual locomotion techniques and environment. We start by providing details regarding our implementation of each technique. We proceed to describe our implementation of the environment, such as collision mechanics and sensory feedback.

4.1 Techniques

In this section, we describe the implementation of the non-visual locomotion techniques under study. We go into detail about the implementation of the functionalities described for each technique in Chapter 3 as well as design changes that were made during the implementation.

4.1.1 Multi Directional D-Pad

For the Multi Directional D-Pad we used the Mobile Single Stick Controller prefab from the Unity Standard Assets package which provided us with a canvas that had simple D-pad in the center of the screen. Upon entering the scene, this controller creates and registers 2 cross platform virtual axes, one for horizontal movement and another for vertical movement.
While the scene is running, this prefab provides an image in the center of the screen which serves as the D-pad’s joystick which can be dragged in order to move the user’s avatar. Upon dragging the joystick, the script calculates the difference between the center position and the new position of the joystick in order to make a Vector3 object which translates into the direction vector the user intends to move in. To pass the direction vector onto the avatar’s animation and control scripts, we update each of the virtual axes for the horizontal and vertical movement with the X and Y components of the direction vector respectively. In order to control movement speed, the direction vector is clamped with the joystick’s drag range value (in order to ignore dragging out of the intended D-Pad range) and then divided by that same value in order to normalize it. This allows us to update the virtual axis with a scaled and normalized direction vector. Upon releasing the joystick, the 2 virtual axis are updated with a null vector in order to stop movement.

With each of the 2 axes updated every frame by the Mobile Single Stick Controller, the Third Person Controller uses the values provided by each of the virtual axes to reproduce the direction vector and use it to update it’s position as well as to determine which animation to use to reproduce that movement.

However, using this prefab implied that movement could only occur by dragging the joystick in the center of the screen which was troublesome for blind users. Therefore, we decided to improve upon this prefab by creating a new script that allowed us to register the direction vector without having to use the center of the screen as our main reference. For that, we maintained the registering and updating of the 2 axes in our new script but we changed the way our direction vector is formed. In order to do that, each frame we check to see if a touch phase has began (if a finger was placed on screen) and we register that touch position as the center of our D-Pad in the form of a Vector3 object (StartPos). For each frame we check the touch position and subtract it to the StartPos in order to get the direction vector. This direction vector is then clamped at a movement range value to ignore dragging out of the intended D-Pad range and then divided by the same value in order to normalize it and update the virtual axes with the X and Y components.
Y components of the resulting vector (Equations 4.1 and 4.2). When the user lifts the finger from the screen and the touch phase ends, the virtual axes are updated with a null vector in order to stop the virtual movement.

\[
\text{HorizontalAxisUpdateValue} = \frac{\max(\text{DirectionVector}.x, \text{DPadRange})}{\text{DPadRange}}. \tag{4.1}
\]

\[
\text{VerticalAxisUpdateValue} = \frac{\max(\text{DirectionVector}.y, \text{DPadRange})}{\text{DPadRange}}. \tag{4.2}
\]

### 4.1.2 Physically Aware 2-Directional D-Pad

For the Physically Aware implementation of the D-Pad, we used the Mobile Single Stick Controller prefab from the Unity Standard Assets again. But for this implementation we changed the script to only register the vertical virtual axis for forward and backward movement. We used the same script as with the Multi Directional D-pad in order to allow the user to be able to initiate virtual movement without having to locate the center of the screen. For this, we added a global variable to the script, which indicates the technique that’s being used, and we use it as a flag when registering and updating the virtual axes. When we are using this technique we only take into account the Y component of the direction vector (Equation 4.2).

In order to make this technique physically aware, we implemented and attached a script to a Structures object containing all of the structures in our scene. In this script we enable the gyroscope input upon scene loading in order to receive the rotation rate of the smartphone. On each frame update, we rotate the Structures object around the the Z axis of the user’s avatar with the rotation rate value in the Z axis provided by the smartphone’s gyroscope. We use the rotation rate instead of the attitude indicator of the gyroscope in order to have our own local referential and simply rotate the Structures object each time the gyroscope registers a change in the rotation on the Z axis. This allows us to start each scene with the Third Person Character aligned with our own referential instead of having to align the participant with the virtual scene. Also, the reason why we chose to rotate the Structures object instead of the Third Person Character was due to the fact that our virtual axes record movement on the X and Y components and rotating those axes more than 90 degrees (where the Horizontal axis would be at the Vertical Axis previous position and vice-versa) led to several bugs where the virtual movement didn’t correspond with the orientation of the smartphone. Therefore, due to the simplicity of our virtual scenes, we decided to simply rotate the structures in our scene around the Third Person Character model.

### 4.1.3 2-Directional Tilt-to-Walk

For the Tilt-To-Walk implementation, we use the same script for the rotation of the Structures object to produce the physically aware behaviour of the technique. We disable the script that handles touch input for virtual movement and enable a script that retrieves accelerometer values from the smartphone in order to use those values for the Tilt-To-Walk behaviour. We use the accelerometer instead of the gyroscope since it is able to provide us with a simpler measurement in the change of orientation of the smartphone.
As with the Physically Aware D-Pad, on scene load, we register only the vertical virtual axis for forward and backward movement. While the virtual scene is running, for every frame, we check the Y component of the linear acceleration of the smartphone in order to obtain the angle between the smartphone and the floor. If the angle is between 30 degrees and 60 degrees we update the virtual axis with a zero value in order to stop virtual movement. If the angle obtained is lesser than 30 degrees, then we update the virtual axis with the following equation’s return value, which determines the locomotion speed of the forward movement:

\[(\text{Acceleration.y} + 0.3) \times 10/3.\]  

(4.3)

Note that when the tilt angle is at threshold for forward movement (30 degrees) the value of Y component of the linear acceleration (Acceleration.y) is approximately -0.3. Therefore, the locomotion speed is zero when at threshold and increases linearly with the tilt angle.

When the angle obtained between the smartphone and the floor is greater than 60 degrees the Y component of the linear acceleration is approximately -0.7. And so, we use the following equation for the locomotion speed:

\[(\text{Acceleration.y} + 0.7) \times 10/3.\]  

(4.4)

In order to maintain the same max speed as the other techniques, both the values obtained by these equations are clamped between zero and one, for forward movement, and minus one and zero for backward movement. After being clamped, we update the virtual axis with the resulting value. Note that for forward movement, the max speed is reached when the Y component of the linear acceleration is zero which corresponds to the smartphone being parallel to the floor. For backward movement, the max speed is reached when the Y component of the linear acceleration is -1 which corresponds to the smartphone being in a perpendicular position to the floor.

\[4.1.4 \text{ Haptic Feedback}\]

To provide the user with a better sense of direction and orientation we added a simple compass with haptic feedback to our techniques similar to the Terraformers approach [10]. We implemented the compass so that when the user is aligned with one of the cardinal points (North, South, East and West) the smartphone will produce a short vibration to inform the user of his orientation. Note that these cardinal points are calculated locally and therefore correspond to the cardinal points of the virtual scene, which are the X-Y axis, and not the real world’s (Figure 4.2).

For the Multi Directional D-Pad, due to it’s non-physically aware behaviour, we implemented the compass by producing the vibration whenever the Third Person Character is moving in one of the cardinal directions. We implement this feature by storing the direction vector of the previous frame and the current one. We then check the quadrant of each direction vector and emit the vibration only if their quadrants differ which would mean that, during the frame update, the direction vector was at some point
aligned with one of the cardinal points. We compare the previous and current direction vector instead of just checking the current frame's direction vector due to the possibility of the frame rate being too low to detect an alignment with the cardinal points during a quick change of direction.

In order to prevent cognitive overload from successive vibrations when aligned with one of the cardinal points, we created a flag that requires the direction vector to make 35 degrees deviation from the cardinal point before allowing another vibration to occur upon alignment with that same cardinal point.

For the Physically Aware techniques, instead of using the direction vector to determine the smartphone's orientation, we use the gyroscope values to find the rotation angle between the smartphone's orientation and the Y axis. For each frame, we store the previous and current rotation angle and use the Cosine and Sin functions to determine the quadrants of each angle. In the case that the angle's quadrants differ, we emit a vibration to inform the user that his rotation is aligned with one of the cardinal points (Figure 4.2).

As with the non-physically aware implementation, we use a flag to prevent the cognitive overload due to multiple vibrations from short deviations and alignments with the cardinal points. Therefore, we again use the Cosine and Sin functions to make sure the smartphone's orientation makes a 35 degrees deviation from the cardinal point before allowing another vibration.

### 4.1.5 Audio Feedback

Audio feedback is a key component for providing spatial information to blind people. However, if too much audio feedback is provided, the cognitive load increases and so it becomes increasingly difficult to form an accurate spatial map. For these reasons, we opted for a simplistic approach for the audio feedback provided by our implementations.

In order to effectively transmit a natural sense of motion, we added 2D footstep sounds to each of our techniques. However, instead of simply playing the sounds at a continuous rate, we synchronized the footstep sounds with the Third Person Character of Unity’s standard assets using animation events (further details in Section 4.2.1).

We also provide 3D spatial sounds for our collisions to provide users with audio cues about the position of the obstacle they collided with. For this, we attached AudioSource objects with a simple knock AudioClip to the wall objects of our virtual scenes. We also added a script to our Third Person Character which checks for a collision every frame and plays those sounds whenever a collision occurs between the Third Person Character and a wall object (further details in Section 4.2.2).
4.2 Virtual Environment

In this section, we describe the virtual environment rendered by our test mobile application. We go into detail regarding our virtual scenes composition and each virtual element’s functionality.

Each virtual scene is composed of 3 elements, a Structures game object which contains all of the wall objects of the scene, a Third Person Character which serves as the user’s avatar in the virtual scene and a Canvas object that is used as an interface to retrieve touch input from the device.

4.2.1 Avatar

For the avatar, which represents the user in the virtual scene, we chose to use a Third Person Character prefab available in the Unity Standard Assets package (Figure 4.3). To this object we added an AudioSource component with an array of 4 AudioClip objects for the footsteps sounds. In order to provide a more realistic feeling to the virtual walk, we synchronized the audio of the footsteps with the walking animation using animation events as opposed to simply playing a footstep sound at a continuous rate which didn’t fully convey the speed of the virtual walk to the user. We also added a script which
chooses the footstep’s AudioClip randomly from the array of footstep sounds for each step for a more immersive experience.

For collision mechanics, we created and attached a script to this model that checks for collisions every frame. When a collision with one of Structures’s child objects occurs, the object plays its AudioSource and we store the direction in which the collision occurred in the form of a Vector3 object. While the avatar is colliding with an object, all virtual movement is stopped until the angle between the direction vector, provided by one of the locomotion techniques, and the collision vector is greater than 90 degrees, making sure that the user can only move in a direction where there is no obstacle. In case the user tries move in a direction where an obstacle exists, that same obstacle will play its collision sound to inform him of so.

We also added a logging script to this model in order to record each of the tasks performed by the user in the form of a log file. In this file we record the position of the user at each point in time as well as the number and length of user pauses, restarts and collisions that occur during each task. We also record the distance travelled and the elapsed time in our log files.

![Image: Unity Standard Assets - Third Person Character model](image3.png)

**Figure 4.3: Unity Standard Assets - Third Person Character model**

### 4.2.2 Structures Object

In order to implement the physically aware behaviour of our techniques we chose to rotate the virtual scene around the avatar’s position. For that, we grouped all of the scene’s wall objects into a parent object named Structures (Figure 4.4). We attached the script for physically aware rotation to the Structures object and enable it on scene load if the technique we are using requires that functionality.

We attached an AudioSource object with a simple knock AudioClip to the wall objects in our virtual scene. We defined the AudioSource component with a 3D spatial blend and play the AudioClip whenever a collision occurs with the user’s avatar. However, in preliminary testing we noted that, upon collision, the AudioSource was too close to the user’s avatar to provide the user with an accurate positioning of the
obstacle he collided with. This is due to Unity’s AudioSource components being centered in the object they belong to, which makes the sound seem almost 2D at close distances. Therefore, we extended the walls in order to create a larger distance between the user’s avatar and the center of each wall object which helped in providing the users with a spatial cue of the positioning of the walls upon collision. For this same reason, we also divided each wall object into several smaller wall objects to make sure that the audio feedback originated in the direction of the collision (Figure 4.4).

4.2.3 Canvas Object

The canvas object serves as an interface to retrieve touch input from the user. For the D-Pad techniques we use a modified version of the Mobile Single Stick prefab, available in the Unity Standard Assets package, in order to obtain the direction vector with which the virtual axes are updated with. However, for the Tilt-To-Walk technique, we disable the script for updating the virtual axes due to the input no longer being supplied by touch.

During preliminary tests, we noted that participants would often lose their whereabouts and requested to return to the starting position in order to explore the virtual scene again. To that end, we created and attached a script to the Canvas object which checks the number fingers on screen for every
frame. If at least 4 fingers are on screen, the user’s avatar position is updated to the starting position and the Structures object’s orientation is reset. We also play a beeping sound to let the user know that the reset has occurred.
Chapter 5

Evaluation

In order to assert if our non-visual locomotion techniques were successful in providing an immersive virtual reality experience, it was important to evaluate the exploration and navigation performance provided by each technique when used by blind people in virtual reality.

Therefore, for evaluating our implementations, we conducted user tests using our own mobile application. Therefore, our user testing was comprised of 2 phases, an exploration phase and a navigation phase.

For the exploratory phase, we wanted to evaluate the mental mapping formed by the participants after exploring a virtual scene with each locomotion technique. Therefore, we asked participants to explore a simple virtual scenario composed of 3 straight corridors with 2 turn points and after asked them to identify the explored virtual space among several tactile maps we provided. We decided upon this setting in order to have 3 virtual scenes with an adequate complexity so that the evaluation of the participant’s mental mapping was not subject to chance.

For the navigational phase, our goal was to record the navigational performance for each technique by having participants perform simple navigational tasks in a simple virtual scene. Since navigation evaluations usually rely on straight paths [28, 21], our virtual scenes for the navigational phase were composed of one straight corridor, one with a clear path and another with an obstacle in the center of the path.

5.1 Participants Profile

In order to evaluate our implementations through user testing we required blind participants. Our selection criteria required participants to be totally blind, older than 16 years old, with an onset of blindness at least 5 years and with no auditory problems.

And so, our study was comprised with a total of 10 participants (Table 5.1), 6 males and 4 females, with ages between 25 years old and 58 years old ($\bar{x} = 50.1 ; \sigma = 10.5$). Of the 10 participants, 7 were introduced to us by the Raquel and Martin Sain Foundation. All 10 participants reported using mobile phones on a daily basis, with 5 participants (4 males, 1 female) owning a smartphone.
All of the 5 participants that reported having a smartphone, used them for calling purposes. Of those 5 participants, 4 used them for messaging purposes with 3 of them using other convenience applications such as the calculator or the alarm clock. Only 1 of the participants that owned a smartphone used social media.

Of all 10 participants, only 2 participants reported to understanding and having had prior experience with virtual reality.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Age</th>
<th>Total Blindness Onset</th>
<th>Smartphone</th>
<th>Smartphone Onset</th>
<th>Frequency of Use</th>
<th>Uses</th>
<th>Virtual Reality Experience</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>25</td>
<td>6 Years</td>
<td>Yes</td>
<td>7 Years</td>
<td>Daily</td>
<td>Calls, Messaging, Apps, Social Media</td>
<td>Yes</td>
</tr>
<tr>
<td>P2</td>
<td>45</td>
<td>17 Years</td>
<td>Yes</td>
<td>5 Months</td>
<td>Daily</td>
<td>Calls</td>
<td>No</td>
</tr>
<tr>
<td>P3</td>
<td>43</td>
<td>20 Years</td>
<td>No</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>No</td>
</tr>
<tr>
<td>P4</td>
<td>57</td>
<td>29 Years</td>
<td>Yes</td>
<td>2 Years</td>
<td>Daily</td>
<td>Calls, Messaging, Apps</td>
<td>No</td>
</tr>
<tr>
<td>P5</td>
<td>56</td>
<td>53 Years</td>
<td>Yes</td>
<td>7 Months</td>
<td>Daily</td>
<td>Calls, Messaging, Apps</td>
<td>No</td>
</tr>
<tr>
<td>P6</td>
<td>51</td>
<td>48 Years</td>
<td>Yes</td>
<td>1 Year</td>
<td>Daily</td>
<td>Calls, Messaging</td>
<td>No</td>
</tr>
<tr>
<td>P7</td>
<td>56</td>
<td>55 Years</td>
<td>No</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Yes</td>
</tr>
<tr>
<td>P8</td>
<td>53</td>
<td>5 Years</td>
<td>No</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>No</td>
</tr>
<tr>
<td>P9</td>
<td>57</td>
<td>23 Years</td>
<td>No</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>No</td>
</tr>
<tr>
<td>P10</td>
<td>54</td>
<td>47 Years</td>
<td>No</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 5.1: Participants profile.

5.2 Apparatus

For the evaluation process, we used a Huawei P10 smartphone (Figure 5.1) as the running device for our mobile testing application with our techniques’ implementations. We used the Huawei P10 smartphone since it was the most recent smartphone available to us. As an audio listener we provided users with a Sony ZX220BT Wireless Headphones (Figure 5.2). We decided on these Bluetooth headphones for commodity of our participants, allowing them to focus on the virtual task while having a better sound feedback from our application and a better isolation from possible outside distractions.

Also, in order to evaluate each exploration task, we provided participants with different tactile maps. The tactile maps 1, 5 and 7 in Figure 5.3, correspond to the 3 virtual exploration scenes and one of them was presented to each participant, along with 3 other wrong tactile maps, upon completion of each exploration task to evaluate the participant’s mental mapping of the area explored.

5.3 Measures

For each phase of our evaluation we gathered several metrics from the log files provided by our mobile application prototype. For the exploration phase, since we aimed at evaluating spatial mental mapping accuracy of each participant, we recorded the following metrics for each task:

- Total Task Duration
- Total Distance Covered
• Number of short pauses (4-10 seconds) for technical purposes such as physical repositioning [17]
• Number of long pauses (more than 10 seconds) supposedly for cognitive processing [17]
• Evaluation of mental mapping through tactile map selection

We recorded the number of long and short pauses in order to determine the ease of use as well as the cognitive load provided by each technique. Also, the total task duration and distance covered helped us determine how efficient each technique is in providing a spatial mapping of the area. Ultimately, the evaluation through tactile map selection was the deciding measure of the correctness of the mental spatial map provided to the user by each technique.

For the navigational phase, our aim was to evaluate navigational performance and ease of participants when conducting simple navigation tasks with each technique. Therefore we recorded the following navigational performance metrics:

• Total Task Duration
In this phase, the total task duration and distance metrics helped us determine the efficiency of each technique when performing simple navigational tasks. With the total distance covered metric we aimed to record the relation between optimal distance and the distance covered by each participant:

\[
\frac{\text{Distance Covered}}{\text{Optimal Distance}}. \tag{5.1}
\]

We also recorded the number of collisions to determine the navigational ease of each technique in avoiding obstacles and reaching target destinations.
5.4 Procedure

For our evaluation procedure we conducted 7 of our evaluation sessions at the Raquel and Martin Sain Foundation, whilst the other 3 were conducted in another controlled environment with little background noise to avoid interfering with the results.

Each session began with a simple introduction followed by a few profiling questions for the participant. After which, we talked to the participant about the goal of the session and explained the virtual environment and feedback that would be provided in our prototype application (audio and haptic feedback). Before moving any further, we had a brief period for clearing up any questions that the participants might have had.

We then presented the first technique to the participant and explained its functionalities before beginning a 5 minute training session, in a simple rectangular virtual room, to allow the user to familiarize himself with the locomotion technique (Section 5.4.1). In order not to favour any technique by the order they were presented, we established a semi-randomized presentation order for each participant (Table 5.2) which was decided prior to the beginning of the session. After concluding the 5 minute training session, we then proceeded to conduct an exploration evaluation (Section 5.4.2) followed by a navigational evaluation (Section 5.4.3). Upon completion of the 2 evaluations we moved on to the next technique’s training and evaluations as according to the predefined order in Table 5.2.

At the end of the session, we inquired participants regarding their technique preference as well as to classify each technique in relation to the difficulty of completing the tasks using a 7 point Likert scale. We also asked each participant about any critic or suggestion regarding the techniques as well as the virtual environment provided.

<table>
<thead>
<tr>
<th>Multi Directional D.Pad</th>
<th>Presentation Order 1</th>
<th>Presentation Order 2</th>
<th>Presentation Order 3</th>
<th>Presentation Order 4</th>
<th>Presentation Order 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 Directional D.Pad</td>
<td>2°</td>
<td>3°</td>
<td>3°</td>
<td>1°</td>
<td>1°</td>
</tr>
<tr>
<td>Tilt-To-Walk</td>
<td>3°</td>
<td>2°</td>
<td>1°</td>
<td>3°</td>
<td>2°</td>
</tr>
</tbody>
</table>

Table 5.2: Order of presentation for Techniques.

5.4.1 Training

In order to allow the participant to familiarize himself with each locomotion technique, a 5 minute training session was conducted. The virtual scenario for this training session was comprised of a simple rectangular room (Figure 5.4). For the first 4 minutes of the training session we asked participants to perform simple training tasks to better understand the locomotion mechanism of each technique as well as the sensory feedback provided by our prototype application. Some of the tasks we asked of our participants were as follows:

- Moving forward until colliding with the North Wall
- Rotating towards East wall for haptic feedback (for physically aware techniques)
• Move towards East wall

• Rotate and move towards Southwest Corner

• Use 3 finger restart to return to starting position

Upon completion of each of these tasks, we would ask participants where they thought they were at the time of completion. In case they did not know where they were, we would provide them with feedback to help them understand their position in the virtual scene. Each time the participant collided with a wall we would also ask where the wall was (to the right/left/front) to make sure they understood the spatial audio feedback. For the last minute of the training session we would let participants familiarize themselves with the technique while clearing up any questions that came up. At the 5 minute mark, we would stop the training session and move on to the exploration evaluation.

Figure 5.4: Virtual Training Room.

5.4.2 Exploration Evaluation

For the exploration evaluation, we explained to the participant that the exploration session would be comprised of a 7 minute free exploration of a virtual scene composed of a varying number of straight corridors (with the same length) and that any possible turn point would be in fact a 90 degree turn. After this explanation, we began the exploration session with one of our exploration scenes (Figure 5.5).

In order not to compromise results due to having the same scene explored by the same technique in every session, we chose the exploration scene for each technique prior to the beginning of the evaluation session as according to Table 5.3.

At the end of exploration session we presented 4 tactile maps to the participant and asked if any one of the presented tactile maps corresponded to the virtual scene explored (the option "None of the 4" was
also allowed). Note that for each exploration scene, all the participants were presented with the same options in order not to compromise results.

<table>
<thead>
<tr>
<th>Scene / Techniques</th>
<th>Explore Order 1</th>
<th>Explore Order 2</th>
<th>Explore Order 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Multi Directional</td>
<td>2 Directional</td>
<td>Multi Directional</td>
</tr>
<tr>
<td>1</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>2</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>3</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Table 5.3: Exploration Order for each Technique.

Figure 5.5: Exploration Scenes

### 5.4.3 Navigation Evaluation

The navigation evaluation took place after the exploration evaluation. We explained to the participant that the virtual scene for this evaluation would be composed a wide corridor, similar to the training session’s virtual scene, and that there would be a varying number of obstacles. This evaluation was composed by 2 virtual scenes (Figure 5.6), one containing an obstacle in the center of the corridor and another with a unobstructed straight corridor.

For this evaluation, we asked the participant to try and reach the end of the corridor in front of him (Yellow zone at the end of the virtual scenes in Figure 5.6), circumnavigating around any possible obstacles. For this task we set a time limit of 5 minutes. For the unobstructed virtual scene, our aim was to determine if the participants were able to walk a straight path with little amount of veering which would translate to a better locomotion precision. For the obstructed path, we aimed to evaluate the navigational efficiency of each technique when avoiding and circumnavigating obstacles.

Again, to avoid compromising results due to familiarization with the virtual scene, we switched the order of both scenarios for each participant.
5.5 Results

Using the information gathered from our log files we analyzed the results produced from both the navigation and exploration evaluations. Our goal was to determine if there were significant differences regarding navigation and exploration performance by blind people when using each locomotion technique.

We divide the analysis of these results into exploration tasks, navigation task and user feedback analysis.

5.5.1 Exploration Tasks

In order to determine if there were significant differences between technique implementations regarding exploration time and distance travelled needed to form a spatial mapping of a virtual scene, we analyzed the total task duration and distance covered with each implementation during the exploration tasks. Our results showed that our participants had a mean average of 303.5 seconds ($\sigma = 79.9$) when using the Multi directional D-Pad implementation, 262.4 seconds ($\sigma = 121.5$) when using the 2 directional D-Pad implementation and 319.9 seconds ($\sigma = 147.9$) when using the Tilt-To-Walk implementation (Figure 5.7). We conducted a Shappiro-Wilk normality test, which showed us that our sample had a normal distribution. After which, we conducted a Repeated-Measures ANOVA test to determine if there were any significant differences between each implementation. Results of the ANOVA test showed that differences between each implementation’s mean average exploration time were not significant ($F(2,18) = 0.689$, $p = 0.515$).

Regarding distance covered, our results showed a mean average of 854.5 meters ($\sigma = 381.9$), 933.5 meters ($\sigma = 606.7$) and 958.7 meters ($\sigma = 561.8$) for participants using the Multi Directional D-Pad implementation, the 2 Directional D-Pad implementation and the Tilt-To-Walk implementation, respectively (Figure 5.8). We again performed a Shappiro-Wilk normality test which classified our sample with a
normal distribution. To determine the statistical significance of the differences between each implementation’s results we conducted an ANOVA repeated measures test. The ANOVA test results lead us again to conclude that no statistical significant differences were found between the mean average distance travelled between each implementation ($F(2,18) = 0.097, p = 0.908$). These results were expected due to the large time window given to the participants to explore the scene, however we noted that participants using the physically aware implementations showed more spatial coverage of the virtual scene than with the non-physically aware implementation.

To evaluate the cognitive overload of each implementation we recorded and analyzed the number of pauses that participants made during the exploration tasks. We divided pauses into 2 categories, short
pauses (4-10 seconds) for technical adjustments, and long pauses (more than 10 seconds) supposedly for cognitive processing [17]. Regarding short pauses, results showed that participants made a mean average of 14.6 short pauses ($\sigma = 4.9$) when using the Multi directional D-Pad implementation, 14.3 ($\sigma = 12.9$) when using the 2 directional D-Pad implementation and 15.2 ($\sigma = 10.8$) when using the Tilt-To-Walk implementation (Figure 5.9). Upon conducting a Shappiro-Wilk normality test, we learned that our sample for short pauses had a normal distribution. After conducting a Repeated-Measures ANOVA test, we determined that there were no significant differences between the mean average of short pauses for each implementation ($F(2,18) = 0.034, p = 0.966$). For our sample of long pauses, our results showed an average of 5.5 long pauses ($\sigma = 3.3$) when using the Multi directional D-Pad implementation, 3.7 ($\sigma = 1.6$) when using the 2 directional D-Pad implementation and 5.3 ($\sigma = 3.4$) when using the Tilt-To-Walk implementation (Figure 5.9). A Shappiro-Wilk test followed by a Repeated-Measures ANOVA test, also determined that there were no significant differences between the mean average of long pauses across all implementations ($F(2,18) = 1.454, p = 0.260$).

These results lead us to assume that there were no significant differences regarding the cognitive load of each implementation.

![Average Number Of Pauses](image)

Figure 5.9: Average Number of Pauses of Exploration Tasks.

To determine the accuracy of each implementation in providing a correct mental mapping of the virtual scene upon exploration, we evaluated participant’s using tactile maps. We recorded the participants’ selection of tactile maps after virtual exploration with each technique. Our evaluation showed that 30% of participants choose the correct tactile map, while using Multi Directional D-Pad implementation, 70% while using the 2 Directional D-Pad implementation and 50% while using the Tilt-to-Walk implementation. A Cochran’s Q test was used to determine if there was a statistically significant difference between each implementation’s results. The results of the test showed no statistically significant difference in the proportion of participants that chose the tactile map correctly while using each implementation ($\chi^2(2) = 3.000, p = 0.223$), however our results show a clear tendency for the 2 Directional D-Pad producing
better results in exploration tasks.

These results lead us to conclude that, despite the 2 Directional D-Pad implementation having better mean averages across the exploration metrics, we cannot statistically differentiate the 3 implementations regarding exploration performance.

![Figure 5.10: Tactile Map Evaluation Results.](image)

### 5.5.2 Navigation Tasks

The navigation evaluation was comprised of 2 tasks which involved reaching the end of the corridor (yellow zone) of both virtual scenes in Figure 5.6. For each of the navigation tasks, we recorded the total task duration, the total distance covered and the number of collisions made by each participant when using each implementation. In order to evaluate the differences in navigational performance between each implementation we performed a statistical analysis of the results while trying to find statistically significant differences.

For the total task duration, our results (Figure 5.11) showed that the mean average time for participants to complete the navigational tasks were of 17.3 seconds ($\sigma = 17.6$) when using the Multi directional D-Pad implementation, 10.8 seconds ($\sigma = 9.5$) when using the 2 directional D-Pad implementation and 7.3 seconds ($\sigma = 4.7$) when using the Tilt-To-Walk implementation for the first navigation task (Left scene in Figure 5.6). To determine if these differences were statistically significant we conducted a Shappiro-Wilk test to learn if our sample did not have a normal distribution. Upon learning this, we conducted a Friedman test, which concluded that there were no statistically significant differences between implementations regarding the mean average completion time of the first navigational task ($\chi^2(2) = 3.935$, $p = 0.140$).

For the second navigation task (Right scene in Figure 5.6), our results showed a mean average completion time of 69.7 seconds ($\sigma = 43.5$) when using the Multi directional D-Pad implementation,
97.9 seconds (σ = 63.4) when using the 2 directional D-Pad implementation and 101.2 seconds (σ = 75.6) when using the Tilt-To-Walk implementation (Figure 5.11). Using a Shappiro-Wilk test we learned our sample for completion times for the second navigational task had a normal distribution. Therefore, we conducted an ANOVA repeated-measures test which found no statistically significant differences between each implementation’s mean average completion time for the second navigational task ($F(2,18) = 0.840, p = 0.448$).

![Figure 5.11: Duration of Navigational Tasks per Technique.](image)

Regarding the distance covered by each of the participants during both navigation tasks, in order to better evaluate the navigational efficiency of each technique, we decided to compare the relation between the optimal distance and the distance covered by participants upon completion of navigational tasks with each implementation. Therefore we gathered the total distance covered for each participant and used the equation 5.1 to determine the relation with the optimal distance for each navigation scene.

Since the first navigation task consists of an open corridor (left scene in Figure 5.6), we aimed to evaluate the steering performance of each implementation when moving forward. For the first navigation task, our results showed a mean average of 1.28 (σ = 0.5) for the Multi directional D-Pad implementation, 1.05 (σ = 0.1) when using the 2 directional D-Pad implementation and 1.03 (σ = 0.1) when using the Tilt-To-Walk implementation (Figure 5.12). With these results we see that, on average, the physically aware implementations provided a better steering than the Multi Directional D-Pad due to their mean averages being closer to 1. To determine if these results were statistically significant, we conducted a Shappiro-Wilk test which concluded that this sample did not have a normal distribution. After conducting a Friedman test we learned that there were no statistically significant differences between the mean averages between implementations ($\chi^2(2) = 1.600, p = 0.449$).

For the second navigation task, the relation between optimal distance and the distance covered helped us to evaluate each implementation regarding the navigational ease in circumnavigating obsta-
cles. Our results showed a mean average of $2.61 (\sigma = 1.7)$ for the Multi directional D-Pad implementation, $3.02 (\sigma = 1.6)$ for the 2 directional D-Pad implementation and $2.99 (\sigma = 2.8)$ for the Tilt-To-Walk implementation (Figure 5.12). As we can see, in the task of circumventing obstacles, the physically aware implementations results show that, on average, participants covered 3 times more distance than the optimal to reach the objective. These results show a tendency for non-physically aware implementations to be more efficient in circumventing obstacles. In order to determine if these results were significant, we conducted a statistical analysis of the sample. Due to the sample not having a normal distribution, we conducted a Friedman test, which concluded that the differences between the averages of each implementations where not statistically significant ($\chi^2(2) = 0.200, p = 0.905$).

Figure 5.12: Relation between Optimal Distance and Distance Covered. Values closer to 1 show better navigational performance.

Finally, we also measured the number of collisions that participants made during each navigation task (Figure 5.13) in order to determine if there were significant differences between techniques in regards to the avoidance of obstacles.

For the first navigation task, our results showed an average of 1.6 collisions ($\sigma = 2.7$) for the Multi directional D-Pad, 0.2 collisions ($\sigma = 0.6$) for the 2 directional D-Pad and 0.7 ($\sigma = 1.6$) for the Tilt-To-Walk implementation. A statistical analysis conducted using Shappiro-Wilk test followed by a Friedman test concluded that there were no statistically significant differences between the mean collision average for each technique in the first navigational task ($\chi^2(2) = 2.941, p = 0.230$).

In the second navigation task, involving the circumnavigation of an obstacle, results showed an average of 13.6 collisions ($\sigma = 8.5$) for the Multi directional D-Pad, 11.7 collisions ($\sigma = 7.2$) for the 2 directional D-Pad and 11.8 ($\sigma = 10.9$) for the Tilt-To-Walk implementation. However, results from a Friedman test determined that there was again no statistically significant difference between the mean collision average for each technique in this navigational task either ($\chi^2(2) = 0.800, p = 0.670$).
After analyzing all the data recorded from the navigation tasks (task duration, distance covered and number of collisions), we were unable to statistically prove if any implementation surpasses the other 2 in terms of navigation performance.

5.5.3 Participant Feedback

We asked users to rate each implementation on 7 point Likert scale based on the difficulty or ease that each technique provided when conducting both exploration and navigation tasks. In this Likert Scale we considered 1 point as the task being near impossible to complete and 7 points as the task being trivial. As we can see in Figure 5.14, most participants agreed that the 2 Directional D-Pad implementation (Median = 6, IQR = 3.25) helped them perform the navigation and exploration tasks more efficiently than the Multi Directional D-Pad implementation (Median = 4, IQR = 2.25) and the Tilt-To-Walk (Median = 5, IQR = 2.25). To determine the statistical relevance of these results we applied a Friedman test to the data which determined that there was no statistically significant difference between implementations in regard to the difficulty perceived by the participants when using each implementation ($\chi^2(2) = 3.879, p = 0.144$).

However, when asked which technique they would prefer to use in a mobile virtual environment, 70% of participants chose the 2 Directional D-Pad implementation, 20% chose the Tilt-to-Walk implementation and 10% chose the Multi Directional D-Pad implementation. Upon measuring the confidence intervals using a 95% Adjusted Wald method, we were able to find statistically significant differences between the 2 Directional D-Pad implementation and the other 2 implementations (Lower Interval = 39%, Upper Interval = 89%). We were then able to conclude that the 2 Directional D-Pad technique was significantly preferred over the other 2 implementations.

We also asked participants if they had any criticism or suggestions regarding the techniques’ or the virtual environment’s implementations. In regards to physically aware characteristics, 8 participants said
that they found the physically aware implementations interesting and helpful to maintain an accurate sense of orientation, 2 participants preferred the non physically aware implementation due to commodity of use (same efficiency when used standing and sitting down). Regarding locomotion speed input, opinions were more diversified with 5 participants preferring the tilt gesture to walk and the other 5 participants preferring the touch input. Tilt gesture advocates reported that, after adapting to the tilting thresholds, they felt more in control of the locomotion speed. Touch input advocates, on the other hand, reported that they felt more control in stopping their virtual movement since it only involved lifting the finger from the screen as opposed to tilting the smartphone back to a neutral position (between 30 and 60 degrees). In regards to the spatial sound provided by the virtual environment in our prototype application, 4 participants found that the spatial collision sounds were crucial to understanding the position of obstacles and finding a way to circumnavigate them. However, 2 participants that found the spatial sounds crucial, also pointed out that there was not enough audio feedback during the time that they were stationary. Both suggested that having a stationary spatial sound source (for example, in their starting position) could have facilitated both evaluation tasks by allowing them to have a better perception of how far from their starting position they had deviated from. Regarding haptic feedback provided from alignment with each cardinal point, 3 participants reported that this feature helped them maintain their sense of orientation mostly in physically aware implementations and suggested that this feature be extended to provide haptic feedback upon alignment with intercardinal points (Northeast, Southeast, Southwest, and Northwest).

5.6 Discussion

After analyzing our results, we were unable to differentiate the 3 techniques in terms of navigation and exploration performance. Although, the 2 Directional D-Pad technique had better averages across most metrics, our statistical analysis was unable to find significant differences between performances with
other techniques. In terms of user feedback, we found a significant preference for the 2 Directional D-Pad across our participants, even though no significant differences were found when participants were asked to rate each technique regarding the ease they provided when completing the tasks. Therefore, with our results, we can only conclude that the 2 Directional D-pad was the most preferred technique among our participants.

However, when participants were asked about any suggestions or criticism regarding the techniques, most of their feedback focused mainly on the environment provided by our prototype application. Most suggestions focused on adding static spatial sounds in order to provide more spatial information about the environment when remaining stationary. Remarks were also made for improvements in the haptic feedback feature in which participants reported that provisioning haptic feedback when turning 45 degrees would be helpful. This could indicate that the lack of significant differences between each technique might have been due to the basic virtual environment provided during the evaluations.

We also noted that the size of our test sample was somewhat reduced which lead to the confidence intervals for our statistical results being quite large, making it difficult to find significant differences across all our metrics. After crosschecking our evaluation’s results with our participants profiles we noted that participants characteristics could have had an influence on our results. We noticed that the 5 participants who owned a smartphone obtained better results than the 5 that didn’t own a smartphone. Also, 8 participants had never learned the concept of virtual reality and that this evaluation was their first contact with it, meaning that more familiarization time might have been required in order to obtain better results. These 2 factors (smartphone ownage and virtual reality experience) could have been relevant to the participants’ interaction with our prototype application and therefore have influenced navigation and exploration results.

Finally, we can only conclude that there are no significant differences in navigational and exploration performance between our 3 implementations when applied in a simplified virtual environment such as the one provided. However, user feedback results suggest that a significant number of participants (8 participants) considered physically aware characteristics crucial to virtual movement. As such, the 2 Directional D-Pad implementation was significantly preferred among all 3 implementations by blind people.
Chapter 6

Conclusion

In this chapter, we reflect over our work’s achieved objectives as well as possible future work for further research in non-visual locomotion techniques for mobile virtual environments.

6.1 Achieved Objectives

The main goal of this master thesis was to design, implement and compare non-visual locomotion techniques for virtual mobile navigation. We also aimed to determine which locomotion characteristics were more important in providing an immersive virtual experience for blind people.

For our first goal we conducted a research into the State of the Art virtual reality for blind people. We identified virtual worlds with accessibility features for blind people which ranged from typing intensive interfaces to simple keyboard shortcuts for virtual locomotion. We also presented works in specialized virtual environments for blind navigation in the real world. These works were divided into in-situ and preplanning approaches, with in-situ approaches focusing more on previously placed sensors in the explored environment. Preplanning approaches provided a safer approach for navigation but most solutions had limitations or relied on custom hardware which was not available to the average consumer. We also verified that preplanning approaches within the mobile context were very scarce. Research was also made into State of the Art virtual locomotion, which led us to various studies into non-natural, semi-natural and natural locomotion comparisons. We verified that natural locomotion techniques provided better immersion in most cases. However, there were cases when non-natural approaches performed better than semi-natural ones due to the approach’s poor design.

As our second goal, we designed and implemented 3 non-visual virtual locomotion techniques. Our aim was to evaluate the influence of physically aware orientation and input method on navigation and exploration performance of blind people in virtual environments. Therefore, our 3 techniques consisted of a Multi Directional D-Pad (Non Physically Aware with Virtual Input), a 2 Directional D-Pad (Physically Aware with Virtual Input) and a Tilt-To-Walk implementation (Physically Aware with Gesture input). In order to evaluate each technique’s navigation and exploration performance we implemented a mobile application that rendered 6 simple explorable virtual scenes with spatialized audio for user testing.
For our third goal, we conducted navigation and exploration evaluations for each technique through user testing. We conducted a statistical analysis of the results gathered from the user tests for each technique, which showed no significant differences between each technique’s navigation and exploration performance. However, a significant number of participants preferred the 2 Directional D-Pad technique. Upon crosschecking of results, participant’s profile and participant’s feedback, we concluded that these results are not decisive due to a variety of factors such as possibly inadequate metrics for technique comparison, simplicity of the virtual environment provided by our prototype application, sample size, participants’ experience with smartphone usage and virtual reality.

6.2 Future Work

Upon reflection of the evaluation phase of this project, we were able to identify some ideas for future research into non-visual locomotion techniques.

In our evaluation, we did not find any statistically significant differences across the metrics recorded for each locomotion technique. However, these results might have been influenced by a number of factors that we noted upon a crosschecking of results, participant’s profile and participant’s feedback.

When analyzing user feedback, a significant number of participants preferred the 2 Directional D-Pad technique, which had better average across most metrics, despite results showing no significant differences from the other techniques. After performing evaluations, 2 out of the 10 participants said that there was a lack of sensory feedback in the virtual environment when they were stationary and suggested that a static spatial sound in one point of the virtual scene could have helped them perform both tasks more easily. Also, 1 participant suggested that the haptic feedback feature be extended to provide feedback upon alignment with intercardinal points also. These suggestions and criticism focus more on the virtual environment provided and less on the actual virtual locomotion techniques. Therefore, we believe that further work must be done with these techniques on a richer virtual environment in order to better differentiate each technique.

We also believe that the selection criteria for future work should be refined. In our work we required participants to be totally blind, older than 16 years old, with an onset of blindness at least 5 years and with no auditory problems. When crosschecking evaluation results with participants’ profiles we found that participants who owned a smartphone performed better than the ones that did not own a smartphone. Virtual reality experience also seems to be an important factor. In our evaluation, 8 participants did not have virtual reality experience and therefore performed worse than the other 2 participants that did have experience. So, for future work, we believe that refining the selection criteria, to require participants to be smartphone users, would produce more precise results when comparing locomotion techniques.

Regarding the techniques, participants were divided between the tilt gesture and touch input for locomotion speed input. This was mostly due to the control offered by each technique when halting the virtual movement. According to participants, the Tilt-To-Walk technique provided more control over the speed of the locomotion but offered little commodity for halting the virtual movement since it required users to tilt the device to an angle between 30 and 60 degrees. Therefore, further work could be
done to review and improve the Tilt-To-Walk technique's thresholds for virtual movement and stationary deviation.
Bibliography


Appendix A

Scripts for User Testing
A.1 Script for User tests

Guião para Testes de Utilizador

Boa Tarde, o meu nome é João Sequeira e sou um aluno do IST a fazer uma tese cujo o objetivo é identificar técnicas de movimento em realidade virtual para pessoas cegas que ofereçam uma melhor percepção do espaço e facilidade de navegação.

Para tal, gostaria da sua ajuda na utilização de uma aplicação que criei onde poderá usar 3 técnicas de movimento diferentes. O objetivo será avaliar as técnicas, portanto quaisquer sugestões ou críticas são bem-vindas. A sessão terá uma duração de cerca de 15/20 minutos sendo que poderá interromper ou desistir a qualquer momento.

Antes de mais, gostaria de lhe fazer algumas perguntas:

- Poderia dizer os seu nome e idade?
- A quanto tempo é que o Sr/Sra (Nome da pessoa) é cego?
- Usa smartphone?
- Se sim, com que frequência?
- Conhece o conceito de realidade virtual? Já experimentou alguma vez?

(Explicar as tarefas)

Hoje iremos fazer 2 tipos de tarefas. Exploração e Navegação

Nas tarefas de exploração poderá explorar um espaço sem tempo limite. Quando se sentir confortável com a sua percepção do espaço irei lhe fornecer 4 mapas tácteis e pedirei lhe para escolher aquele que corresponde ao espaço que acabou de explorar. Tem também a opção de dizer que o espaço que explorou não se encontrar entre os mapas fornecidos.

Nas tarefas de navegação, o objetivo é avaliar a facilidade de navegação com cada técnica. Neste caso teremos dois espaços virtuais que lhe irei explicar a priori a sua constituição (fornecer mapa táctil do espaço?) e terá de se movimentar até ao final da sala. Para estas tarefas irá navegar pelas duas salas usando cada técnica sendo que no final irei perguntar lhe qual a que a preferiu.
A.2 Script for User tests

(Explicar técnica a usar – explicação no final do documento)

(Pré Tarefas Exploratórias)
Antes de iniciar a tarefa exploratória teremos um pequeno treino com cada técnica onde poderá navegar um espaço virtual composto apenas por uma sala retangular durante 5 minutos para se ambientar a cada técnica.

Durante o movimento irá ouvir os seus passos na aplicação de acordo com a velocidade em que se move. Caso embata numa parede irá ouvir um som de knock que irá soar mais alto de um lado dos headphones de acordo com o lado em que embateu na parede. Quando embater numa parede só se conseguirá mover se a sua direção for uma em que não haja um obstáculo. Caso sinta que se perdeu ou quiser voltar ao inicio, necessita apenas clicar no ecrã com 4 dedos e ouvirá um bleep e voltará a sua posição inicial. Para facilitar a navegação, cada vez que se estiver a movimentar numa das direções Norte, Sul, Este ou Oeste o smartphone irá vibrar para o informar de tal. Alguma dúvida? (Pausa)

(Iniciar cenário com sala de teste)
Muito bem, agora iremos passar à tarefa de exploração.

As salas para as tarefas exploratórias são compostas apenas por alguns corredores retos e caso haja alguma curva esta será sempre uma curva à esquerda ou à direita. Terá 7 minutos de exploração livre sendo que no final dos 7 minutos ou quando se sentir confiante no seu conhecimento do espaço, irei fornecer-lhe 4 mapas tácteis e irei pedir lhe para me dizer se algum dos mapas corresponde ao espaço que acabou de explorar. Alguma dúvida? (Pausa)

(Efetuar Tarefa Exploratória)
(Efetuar Avaliação cognitiva usando os Mapas Tácteis)

(Tarefas Navegacionais)
Muito bem, agora iremos passar à tarefa de navegação

Nestas tarefas de navegação, a sala em que se encontra é uma sala semelhante a sala de teste, um corredor retangular, em que o objetivo será chegar ao fim do corredor sendo que quando lá chegar irá ouvir um som de sucesso. Iremos fazer 2 tarefas sendo que em cada uma delas pode ou pode não haver um obstáculo que terá de contornar.

(Efetuar Tarefas de Navegação)

(Efetuar o mesmo procedimento para cada Técnica)
Acabamos assim a nossa sessão, espero que tenha gostado. Gostaria de lhe fazer apenas algumas perguntas:

- Com base nestas avaliações, gostaria de pedir para classificar cada técnica numa escala 1 a 7 de acordo com a dificuldade/facilidade ao efetuar as tarefas.
- Houve alguma técnica que tenha preferido acima das outras?
- Tem alguma sugestão ou criticas a fazer às técnicas que acabou de utilizar?
A.3 Script for User tests

(Técnica de D-Pad)
Nesta técnica, para se movimentar necessita apenas de colocar um dedo no ecrã do smartphone e deslizar e manter o dedo na direção que deseja ir. Poderá parar o movimento voltando ao centro ou simplesmente tirando o dedo do ecrã. Pode controlar a velocidade do movimento deslizando mais ou menos o dedo no ecrã. Novamente o smartphone irá emitir uma vibração sempre que estiver alinhado com um dos pontos cardeais.

(Técnica de D-Pad com Rotação Corporal)
Nesta técnica, poderá movimentar se para a frente e para trás deslizando mais ou menos o seu dedo no ecrã conforme a velocidade que queira andar. Para virar tem apenas de rolar o seu corpo/smartphone na direção que deseja ir. Novamente o smartphone irá emitir uma vibração sempre que estiver alinhado com um dos pontos cardeais.

(Tecnica de Tilt-To-Move com Rotação Corporal)
Nesta técnica, para se movimentar necessita apenas de inclinar o telemóvel para a frente ou para trás. Quanto mais inclinar o telemóvel maior será a velocidade a que andará. A inclinação do smartphone servirá apenas para determinar a sua velocidade de andamento sendo que para virar necessita apenas de rolar o seu corpo/smartphone na direção em que deseja ir. Novamente o smartphone irá emitir uma vibração sempre que estiver alinhado com um dos pontos cardeais.