Summer 7-26-2016

Towards Understanding and Developing Virtual Environments to Increase Accessibilities for People with Visual Impairments

Miao Dong

TOWARDS UNDERSTANDING AND DEVELOPING VIRTUAL ENVIRONMENTS TO INCREASE ACCESSIBILITIES FOR PEOPLE WITH VISUAL IMPAIRMENTS

Approved:

Rongkai Guo
Thesis Advisor, Committee Chair

Dan Lo
Professor, Committee Member

Yong Shi
Associate Professor, Committee Member

TOWARDS UNDERSTANDING AND DEVELOPING VIRTUAL ENVIRONMENTS TO INCREASE ACCESSIBILITIES FOR PEOPLE WITH VISUAL IMPAIRMENTS

A Thesis Presented to

The Faculty of the Computer Science Department

by

Miao Dong

In Partial Fulfillment

of Requirements for the Degree

Master in Computer Science

Kennesaw State University

July, 2016

In presenting this thesis as a partial fulfillment of the requirements for an advanced degree from Kennesaw State University, I agree that the university library shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish, this thesis may be granted by the professor under whose direction it was written, or, in his absence, by the dean of the appropriate school when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from or publication of, this thesis which involves potential financial gain will not be allowed without written permission.

Miao Dong

Miao Dong

Notice To Borrowers

Unpublished theses deposited in the Library of Kennesaw State University must be used only in accordance with the stipulations prescribed by the author in the preceding statement.

The author of this thesis is:

Miao Dong

1100 South Marietta Pkwy
Marietta, GA 30060

The director of this thesis is:

Rongkai Guo

1100 South Marietta Pkwy
Marietta, GA 30060

Users of this thesis not regularly enrolled as students at Kennesaw State University are required to attest acceptance of the preceding stipulations by signing below. Libraries borrowing this thesis for the use of their patrons are required to see that each user records here the information requested.

TOWARDS UNDERSTANDING AND DEVELOPING VIRTUAL ENVIRONMENTS TO INCREASE ACCESSIBILITIES FOR PEOPLE WITH VISUAL IMPAIRMENTS

An Abstract of

A Thesis Presented to

The Faculty of the Computer Science Department

by

Miao Dong

Bachelor of Laws, South-Central University for Nationalities, 2013

In Partial Fulfillment

of Requirements for the Degree

Master in Computer Science

Kennesaw State University

July, 2016

Abstract

The primary goal of this research is to investigate the possibilities of utilizing audio feedback to support effective Human-Computer Interaction Virtual Environments (VEs) without visual feedback for people with Visual Impairments. Efforts have been made to apply virtual reality (VR) technology for training and educational applications for diverse population groups, such as children and stroke patients. Those applications had already shown effects of increasing motivations, providing safer training environments and more training opportunities. However, they are all based on visual feedback. With the head related transfer functions (HRTFs), it is possible to design and develop considerably safer, but diversified training environments that might greatly benefit individuals with VI. In order to explore this, I ran three studies sequentially: 1) if/how users could navigate themselves with different types of 3D auditory feedback in the same VE; 2) if users could recognize the distance and direction of a virtual sound source in the virtual environment (VE) effectively; 3) if users could recognize the positions and distinguish the moving directions of 3D sound sources in the VE between the participants with and without VI.

The results showed some possibilities of designing effective Human-Computer Interaction methods and some understandings of how the participants with VI experienced the scenarios differently than the participants without VI. Therefore, this research contributed new knowledge on how a visually impaired person interacts with computer interfaces, which can be used to derive guidelines for the design of effective VEs for rehabilitation and exercise.

TOWARDS UNDERSTANDING AND DEVELOPING VIRTUAL ENVIRONMENTS TO INCREASE ACCESSIBILITIES FOR PEOPLE WITH VISUAL IMPAIRMENTS

A Thesis Presented to

The Faculty of the Computer Science Department

by

Miao Dong

In Partial Fulfillment

of Requirements for the Degree

Master in Computer Science

Rongkai Guo

Kennesaw State University

July, 2016

DEDICATION

This thesis is dedicated to my parents,
I appreciate everyone who supported me throughout my study.

ACKNOWLEDGEMENTS

I would like to thank my thesis director Dr. Rongkai, Guo, for his enlightenment and guidance.

I would also like to thank Haiqi Wang, who compiled data analysis for study 3.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE: INTRODUCTION

Numerous surveys reveal that staggering numbers of people are suffering from visual impairment all over the world. Data from the National Center for Health Statistics has shown that over 20 million American Adults age 18 and older reported experiencing vision loss in 2012 [1]. The World Health Organization released data which indicated that about 285 million people are visually impaired worldwide: 39 million are blind and 246 million have low vision (severe or moderate visually impairment) [2]. An analysis of the 1999 Survey of Income and Program Participation (CDC, 2001) revealed blindness or vision difficulties to be among the top 10 disabilities among adults aged 18 years and older [3]. Statistics from the National Federation of the Blind (2014) reported that there are more than 60 thousand legally blind children (through age 21) enrolled in elementary and high schools in the U.S. [4].

In an effect to ameliorate the condition of visual impairment, Virtual Reality (VR) technique has progressed in recent years. This technology is used to simulate reality, including sight, touch, hearing, and smell. VR applications have been developed for numerous diverse fields, such as physical rehabilitation, education, and healthcare. Beneficiaries of VR technology include children, the elderly, and persons with physical and mobility impairments [5-16]. One of VR technology's most important contributions is its use to ensure safety [7, 12, 13, 17-19], such as pilot training simulators or mining conveyer belt safety training. A gap remains, however, as the majority of research associated with VR applications have been based on visual feedback, thus excluding people with visual impairment. Such a dearth of studies of VR for people with visual impairments [20-22] warrants exploration. I believe that although people with visual impairment cannot enjoy the advantage of VR in visual aspects, they might be able to utilize an auditory function of VR. Therefore, I explored the possibility that VI individuals are able to utilize audio feedback to support effective Human-Computer Interaction by using Head Relative Transfer Functions (HRTFs) in a virtual environment (VE). Furthermore, integrating this VR technique into training could improve safety and raise confidence to VI individuals in activities outside the home. Outside activity plays an important role in daily life for VI people and it has been proven that such physical activity has a beneficial effect [23].

Specifically, I presented the results of two phase studies. In the first phase, I recruited experienced gamers as participants in the first phase. There were two studies in the first phase: 1) Investigating if/how gamers could navigate themselves with different types of 3D auditory feedback in the same VE. 2) Investigating if gamers could recognize the distance and direction of a virtual sound source in the virtual environment (VE) effectively. The results of the first phase could suggest the design of VEs for general purposes that might increase the user experiences of VEs with 3D auditory feedback. In the second phase, I recruited VI people and sighted people as participants who are all non-experienced gamers. Therefore, the third study aimed at investigating the differences of recognizing the positions and distinguishing the moving directions of 3D sound sources in the VE between the participants with and without VI. The results of this study would help us understand more if and how we could build a 3D audio based VE for visually impaired users.

CHAPTER TWO: BACKGROUND AND RELATED WORK

2.1 Head Related Transfer Function

People were able to locate sounds in space by three primary methods: distinguishing the difference in pressure between the two ears, determining the difference in arrival time between the two ears, and filtering sound differently based on its location. The sound could be diffracted and reflected by a unique and consistent way on its path to the eardrums. The brain uses this filtering function to localize sounds. The head-related transfer function (HRTF) is a digital model which simulates this physical filtering process. It describes acoustic transmission effects that a sound experiences from a certain sound source to the listener`s eardrum in the free field.

A large quantity of research investigated if and how the HRTFs help listeners to localize position [24-26]. HRTFs contain all the spatial cues used to decode spatial information encoded in binaural signals by the hearing mechanism. Binaural audio allows listeners to use both ears because it has been recorded by placing two microphones close to the ears as is practical. These two microphones collect the location information including frequency, amplitude, and phase response from left and right channels. This characteristic of binaural audio allows listeners to perceive genuine 360 degree sound when they listen to binaural audio through headphones. Therefore, the binaural audio has been represented customarily by 3D audio [24]. Martin et al., utilized the open source VoIP software and Session Initiation Protocol to construct telephone conference server software [25]. This software enables users to receive stereo signals through their headphones by the central conference server that extends all participants` signals by binaural sounds using HRTFs.

German et al., proposed there are several problems when designing the real-time binaural rendering systems [26]. Thus, they designed an efficient method based on a low-order infinite impulse response filter. Utilizing the parameter of estimated second order sections, this modeling filer provided a simple way to perform HRTF interpolation. Although utilizing the HRTF has slight deviations, much other research on HRTFs found similar results and proposed a number of methods to improve the performance of this function.

According to amount of studies on HRFTs, it seems that HRTFs could work by 3D audio signal and I believed that VI people would performance well with HRTFs by hearing 3D audio in VE.

2.2 Phonon 3D

As VR gradually develops, in order to enhance the sense of presence in VR, developers should consider utilizing sound design. By accurately portraying the location and moving direction of the sound in VR, gamers feel that they are truly in another world we constructed. Therefore, Impulsonic has released their VR audio development tool, Phonon 3D, which uses HRTFs to render high-quality 3D audio.

Phonon 3D is the fastest 3D audio plugin available, and provides a simple method for developers to integrate binaural 3D audio into their VR projects [27]. Automatic performance scaling ensures a smooth, immersive player experience, which allows users to pinpoint the most important sounds, even with hundreds of sources playing. It integrates various aspects, such as the position of the player's head, to help produce accurate sounds. Now, Phonon 3D includes full support for head tracking, including rotational and positional tracking. It could be applied in Oculus Rift, Samsung GearVR, or OSVR, and adds a new dimension to VR audio. It supports a range of game engines and audio middleware tools, and seamlessly integrates into existing workflow, with built-in support for Unity, Unreal Engine, Wwise, and FMOD Studio.

2.3 Feedback in VR

2.3.1 Visual Feedback

Most VR research work and applications use visual feedback to present virtual environments to users. For example, research work that compared different navigation methods in virtual environments [28-31], investigated how avatars could affect behaviours [32-38] and how to use VR technology for education [39-44]. Presence is a theoretical concept describing the user's level of involvement in the VE, the suspension of disbelief, and refers to an experience of actually 'being there' in a VE. In visual feedback of VR research, most researchers focus on studying the presence for people in VE. Jennifer M et al. ran a study to investigate the influence of specific VR application characteristics in a navigation task, because they aimed at exploring how to structure a VR application to enhance the sense of presence in a VE [31]. They claimed that it was unexpected that the display type presenting the VE did not significantly influence performance in navigation tasks. Researchers are creating more techniques for virtual environments to enable better experiences for their subjects. The use of a Head-Mounted Display (HMD), a CAVE system, or 3D displays to give users more immersive display are a few of the new techniques. The use of a motion capture system to provide a real-walking interaction method to users will enable a more natural navigational process, and the use of see-through HMD and augmented reality technology will allow users to combine both the real world and virtual world. While these are cutting edge innovations, most of these areas focus on visual feedback and cannot be used for people with visual impairments.

2.3.2 Audio Feedback

Audio feedback is the second most popular response in VE, and one of the essential aspects that improve presence in VE. However, most of research has investigated in developing three-dimensional viewing simulations in VR, so audio was always in addition to visual feedback. Three-Dimension audio is involved in the field of VR and is an essential part in VR [45-50]. Francesco et al. combined virtual reality technology and audio rendering techniques to experiment a new approach for environmental noise assessment [45]. Participants were exposed to 3D sighted reconstructions of an actual landscape, while noise was provided, and their reactions were recorded. We should recognize the significance of 3D audio in VR and utilize it to enhance user experience with VR. In 3D space, when a sound source is simulated at a certain location, the listener would locate the position by filtering the audio stream with the left and right ear HRTF corresponding to the sound source location [51].

HRTFs allow the developer to create a low-cost 3D sound system. To describe 3D audio feedback of a sound source in a VE, we actually need to synthesize two HRTFs, one for each ear, to simulate the sound source from a position. The research of Doerr et al. 2007 provided a 3D sound simulation for participants to localize the sound even in 3D space. Their sound engine was based on OpenAL and EAX. The three types of sound signals they offered were impulse signal (0.5 seconds), prolonged signal (2 seconds) and speech signal (2 seconds). In the experiments, participants were placed in the center of a sphere divided in eight sectors. The sounds were placed on the surface of the sphere at different positions and the task for the participants was to estimate the sound location. Their results showed that users were able to identify the location of sounds simulated by using HRTFs, with the exception of the top position which was not well recognized. Furthermore, they found that the best signal to provide a localizable 3D sound event was the impulse signal [52]. In 2010, Haraszy, et al., presented an Acoustic Virtual Reality (AVR) implemented by an improved HRTF [53]. They utilized two artificial neural networks (ANN) to generate a pair of HRTFs, one ANN for each ear. Their work demonstrated that with the aid of the ANN, the generation of the HRTFs for people with visual impairments is possible. If there is at least one subject included in the Listen HRTF or any other available database, their method could avoid the duration of the complex measurement process required to acquire the HRTFs for one certain subject. Lokki et al. ran an experiment which required participants to navigate in an immersive 3D environment by using either audio, visual, or audio-visual cues [54]. They found that the audio cues were more effective than visual counterparts during the initial stages of locating their target.

Nevertheless, there have been no usability studies investigating user experience of people with visual impairments.

2.3.3 Haptic Feedback

The Haptic system has enabled users' touch sense in VR since the 1990s [55-57]. In 1999, Jansson et al., utilized force feedback devices to investigate the perception of virtual textures, the identification of virtual objects, and the perception of an object`s size and angles [58]. They recruited blind people and sighted people as participants. Based on their results, they found that the performance of blind people was obviously different from sighted people`s. However, in their studies, the objects were so simple that it would not provide enough spatial information for participants when the textures and 3D object were more complex. In 2012, Romano et al., presented new methods to create much more detailed and realistic haptic virtual textures [59]. Their system employed a handheld tool to capture the feel of a texture, recording 3D tool acceleration, position and contact force over time. Furthermore, they proposed a method to reduce the recorded 3D acceleration information down to 1D signal, and detailed a process for converting each 1D time-domain signal into a frequency-domain. This research only proved haptic feedback can be used as one example of feedback in virtual environments to recognize virtual objects, which is not an effective interaction method for other types of tasks in virtual environments. Since haptic sense is one essential perception for VI people and based on the haptic system development in VR, it would be an innovative subject for future research.

2.3.4 Olfactory Feedback

Olfaction is different from the visual, auditory and haptic sensation that is activated by physical stimuli, and it is activated by chemical stimuli. This characteristic of olfaction brings difficulty for the research; therefore, little research focuses on this area. In 2004, Yanagida and et al. used a camera to track users noses and send olfactory information to users by an air cannon [60]. They believed the VR system could achieve a high level of presence by incorporating olfactory interfaces. However, they found there were still several problems remaining in their system. The density of the odor was insufficient for distinction, and users should detect the odor in very short duration after a scented vortex ring was detectable. In 2013, Matsukura et al., discovered the popular olfactory display system simply ejected odor directly towards users. In order to resolve this potential limitation, they developed a smelling screen, which distributed an odor vapor to a specific position on a two-dimensional screen [61]. Their olfactory display system allowed users to move their head to follow the change of odor intensity by sniffing location in front of the screen. It covered the shortage of Yanagida`s system as users did not need to immediately notice the odor distribution, which enhanced the reality of the users` experience. Although their system advanced olfactory display systems, several problems remain, such as inability to control the release rate of odor vapor precisely. Generating the kind of virtual odor source was simple, and the distance at which the user could perceive the smell was limited. According to these studies, utilizing olfaction to localize position still did not provide efficient performance. Although olfactory feedback does not require sight, and can be also a possible feedback for people with visual impairments, yet we still do not know how we can use olfactory feedback to develop virtual environments for people with visual impairments.

2.4 VR for People with Visual Impairments

Unlike sighted persons, people with visual impairments use only their sense of hearing, touch and smell. The existing virtual environments that developers have created could merely provide an extremely minimal sense of hearing, touch or smell, but no major sense of vision. Therefore, it is still unknown whether or how effectively one could simulate, or create, an environment for people with visual impairment. Gareth et al., interviewed eight visually impaired expert users in 2008 concerning an increase in the accessibility of 3D virtual environments for the blind and visually impaired [62]. They proposed their resolutions for VI users in the Second Life (SL) around the use of audio and haptic technologies. Their VI user demonstrated the issues involved in using the conventional screen-readers, thus it seems an efficient technique for sighted users, but not VI users. The technique in the audio game BAFTA enables users to infer the size, shape and volume of an environment and objects within it. Also the further technique in various audio games is a form of sonar, which informed users the distance to virtual objects by changing the pitch or volume of the audio in real time. Participants confirmed this to be a powerful technique. Furthermore, their VI participants expressed that they are more likely to want to build up the routes in their head before visiting new areas than sighted people. Therefore, they felt the sonified maps could allow them to develop a mental model in their mind, which efficiently helps them in navigation. Based on Gareth`s research, it seems possible for VI users to gain enough spatial information by 3D audio and haptic device in VE to navigate effectively.

The suggested approaches for 3D navigation used audio and haptics. David et al., used haptic and 3D audio to develop the BlindAid system, which provides a means for blind users to safely explore and learn about new areas in advance on their own [63]. To build this system, they combined the haptic feedback provided by a Sidewinder Force Feedback Pro Joystick with a low-cost spatialized sound system to build this system This system is also a research and learning tool for improving orientation and mobility training. Orly et al., integrated BlindAid system, which is equipped with a haptic device and stereo headphones in an orientation and mobility rehabilitation program named O&M as a training aid for blind people [64]. Their system provided blind people training in a safe learning environment without the stress associated with exploration of real spaces. Their results demonstrated the participants who were newly blind were able to acquire spatial information when they explored the unfamiliar areas, and apply it in real space orientation tasks. Iglesias et al., developed the GRAB system to investigate the interaction process of visually impaired persons with haptic environments [65]. Their GRAB system allows user to feel the shape of virtual 3D objects with their fingers. When users move one finger over the virtual object, they feel contact forces at the fingertip and recognize an object`s geometric features by distinguishing the sizes and distances and understanding spatial relationships between elements. Maria et al., designed an Audio-Haptic learning environment to enable people to use their fingers to explore the shapes of small scale 3D objects [66]. In their environment, users were able to explore 3D objects by shape and texture. Nevertheless, the above four research efforts were primarily based on haptic feedback. Audio feedback was an addition, which only provided extremely limited information.

Lorenzo et al., allowed the use of spatial auditory feedback to assist blind people while learning an unknown environment [67], but it was not for interaction. Oana et al., developed a navigational 3D audio-based game [68]. Their research showed that the game could help users manipulate the location, and thus create a spatial cognitive map for imaginary representation. The results also demonstrated the physical characteristics of sound, just as loudness and pitch can convey relevant information. Their research showed that spatial audio feedback might provide a suitable amount of information to enable people with visual impairments to interact with 3D virtual environments. Still, the findings of the study only demonstrated that sighted people were able to use 3D binaural sounds as the only means of navigation.

2.5 Auditory assistance for People with Visual Impairments

As we know, audition is one of the essential senses for VI people, allowing design training or exercise routines. These training or exercise routines are used to accommodate VI individuals seeking to enhance self-reliance, and improve skills in specific sports. Beep baseball is an example of one such game using auditory assistance, and is highly competitive and skill-oriented. Implanting a sound module into a 16-inch softball creates a beeping sound, allowing blind and visually impaired persons to participate. In addition, the bases contain sounding units that give off a buzzing sound when activated. Therefore, VI players can distinguish the moving direction of the ball and localize the base by hearing the sound from the ball and the base. Although VI players have this auditory assistance, they also risk an occasional injury. Therefore, we could consider utilizing VR techniques to resolve this safety issue for the training of beep baseball. By training, VI players improve their skills for beep baseball, which also offers protection from injuries.

2.6 A Gap in Knowledge

A review of the VR literature above reveals that research has been conducted to investigate VR effectiveness for navigation tasks on visual function, but little research prescribes how a VR application could be presented to enhance the auditory sense in a VE and provide for the most effective training. One of the reasons why visual feedback is necessary is to provide continuous feedback to users, which has remained unresolved for the visually impaired. To our knowledge, how to use spatial audio techniques to develop continuous feedback in virtual environments for people with VI has been minimally explored. It is critical to understand the effects of using spatial audio feedback in virtual environments without visual feedback, especially considering safety for training purposes. Based on the previous research, it is likely that spatial audio feedback might help people with VI, but the ability of a stand-alone spatial audio feedback in virtual environment is unknown.

2.7 Contributions

In the real world, people navigate by sound, and often utilize echolocation. However, Ludwig Wallmeier and Lutz Wiegrebe have since demonstrated that simulating echolocation in VEs did not bring the same perception as in the real world (Wallmeier and Wiegrebe, 2014). How, then, could we build a virtual world that people with VI can use? I conducted studies to investigate depth, location, and movement recognition by using object-generated sounds to provide information in VE. Additionally, I compared the differences of recognizing and locating 3D sound sources in the VE between people with and without VI.

CHAPTER THREE: SYSTEM DESCRIPTION

I utilized Unity to develop all of the VEs for the three studies. In the first phase, for study 1) and 2), participants wore an eye cover to block all visual feedback and an over-ear headphone during the study. The headphone blocked out all ambient sounds from the real world to keep the participant focused on the VE. All of the 3D sounds in the VE were generated by headphones which could be recognized by generic Head Related Transfer Function (HRTF). Participants were asked to use an Xbox 360 controller to move in the VE. I designed two valid thumb-sticks, the left thumb-stick which allowed participants to move forward, backward or sideways and the right thumb-stick which allowed participants to make rotations. The task of study 1 was to navigate through a virtual gallery by hearing different auditory feedback and using the 360 Xbox controller. The task of study 2 was asking participants to estimate the distance between a sound source and themselves and the direction of sound source.

Based on the first phase, I used in-ear earphones for the third study to improve audio presence. I also changed the generic HRTF to Phonon 3D to more efficiently provide 3D audio in VE. A universal VR Glasses Headset for smartphones was used to hold a Samsung S6 to run the VE. The gyroscope sensor in S6 android phone was used for head rotation. The visual feedback of the VR Glasses Headset was blocked by foam paper. The VE running on the smartphone was synchronized with a laptop connected to the same network. The experimenter could use the laptop to monitor, control, and log the study. The tasks of the third study were locating fixed 3D sound sources around the participants and recognizing the moving directions of moving 3D sound sources in the VE.

All of the studies were conducted in a quiet, air-conditioned environment. Only the participant and experimenter were present.

CHAPTER FOUR: STUDY ONE

4.1 The Study 1: Audio Reminder VS Continuous Audio Feedback (ARCAF)

In this study, the participants were asked to navigate in a simple virtual gallery (Figure 1 Left) by using an Xbox 360 controller. The task was to use audio reminders or feedback to walk through the entire virtual gallery from the start point to the end point. The path of this virtual gallery was the red line in Figure 1 Left. There were three sessions provided randomly and each session had different audio reminders or feedback for participants. I also provided beeps with four different pitches at 0 degrees, 90 degrees, 180 degrees, and 270 degrees. Thus, participants could use these beeps to locate the direction they were facing.

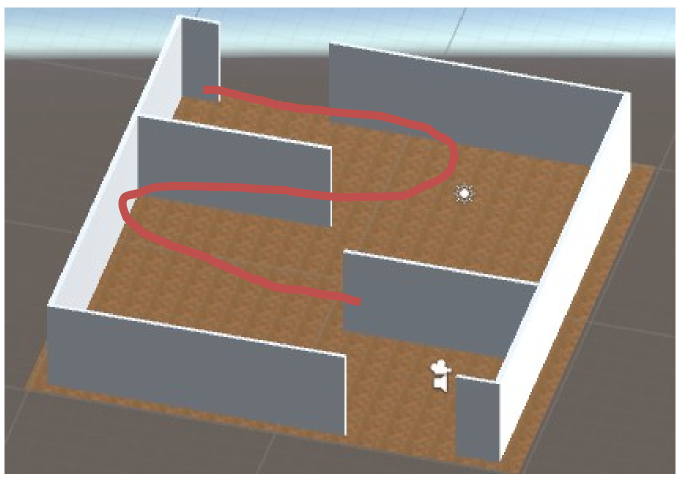


Figure 1：The virtual gallery

4.2 Population

I recruited 28 college students from the Computer Game Design & Development program for participation in this study. I discarded two participants' data, as one of them had a hearing problem, and the other was disrupted by a fire alarm test while participating in the study. The amount of participants ensures that all possible orders of three sessions had been averagely provided. During the entire study, participants` eyes were covered to simulate being visually impaired. Each participant would earn one extra credit in their final grade as compensation.

4.3 Conditions

Usually, people with visual impairments have two different types of audio feedback: passive audio feedback, such as echoes, and active audio feedback, such as moving vehicles. In this study, I used object-generated sounds to simulate both types of audio feedback, including three conditions for this study. Each condition was a separate session.

Condition 1 – Wall Alert Audio Reminder (WAAR): I provided an alert coming from the wall only when the participants were close to the wall and this safety distance was beyond 1.5 meters. The alert was generated on the wall side, which meant the alert served as 3D audio feedback. For example, if they were close to a wall on the left side, they should be able to hear the alert from the left side as well.

Condition 2 – Continuous Path Sound Audio Feedback (CPSAF): I generated one bell sound at each turning point. These sound sources were presented one by one consecutively, and the participants could walk through the virtual gallery by following these sound sources. When participants arrived at the location of the sound source, the sound source was deactivated and the next sound source became activated until the participants reached the final sound source.

Condition 3 – Continuous Wall Alert & Path Sound Audio Feedback (CWAPSAF): In this condition, participants could ALWAYS hear continuous alerts from the wall that participants were facing, NOT only when the participant was close to the wall. The alert had a different pitch and frequency. The closer wall had higher pitch and frequency alert. When the participants knew where to go, they used the Xbox 360 controller to generate a continuous path sound source by pressing the left thumb-stick. When the participants arrived at the location of the path sound source, the path sound source became deactivated.

4.4 Procedure

This study had three sessions for three conditions. The order of the three sessions was randomized and counter-balanced. Each session included two scenes: 1) Training Scene and 2) Experiment Scene. Before the Experiment Scene, the participants were trained how to use the controller and how to interact with the virtual environments with different audio reminders or feedback. The experimenter also guided the participants to walk through the virtual gallery in the training session to make sure the participants were familiar with the navigation path. The entire procedure is as shown in Figure 2.

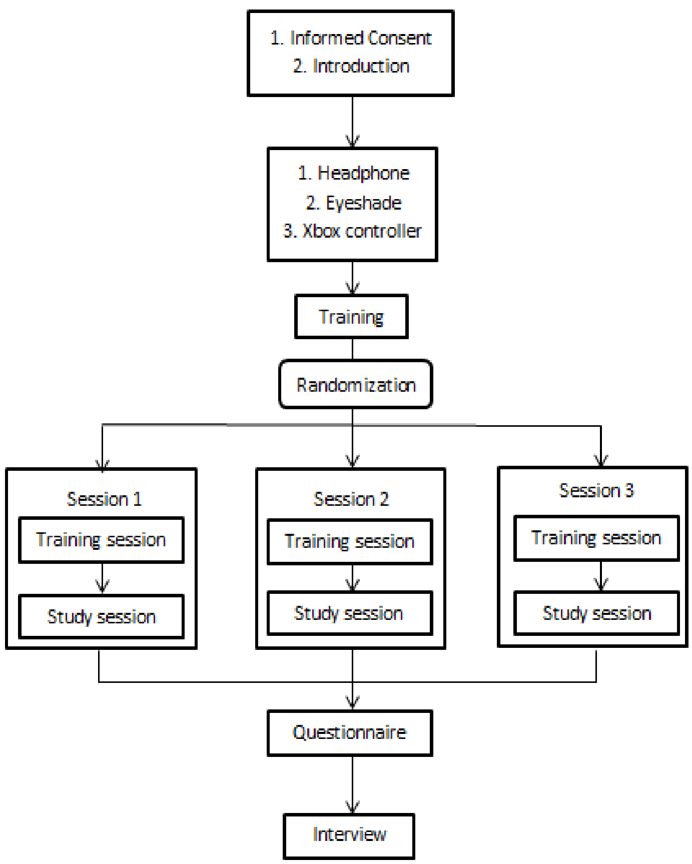


Figure 2: The flowchart of the study 2 procedure

I *Informed Consent and Introduction* – When the participants arrived, they were asked to read the informed consent and sign it if they did not have questions. I also briefly introduced the study to them.

II *Headphone, Eye Cover, and Initialization* – The participants wore the headphones and the eyeshades before beginning the study. Then they were requested to hold the Xbox controller and try to use it. Furthermore, I verified if they could hear the sound from the headphones and adjusted a suitable volume for them.

III *Training for Path of Gallery* – This training was designed to help the participants understand the walking path of the virtual gallery. I used the same VE (Figure 1) for participants in this training to allow the participants to become familiar with the path.

IV *Training for Orientating in the VE* – I designed a user interface to assist the participants in identifying the orientations in the VE. This training aided the participants in understanding this user interface. I used a virtual room which had walls (Figure 1 Right) as the training scene for participants.

V *Training of Session 1* – This training helped the participants to learn the audio reminders and the Xbox 360 controller they would use in the VE. I used the same training scene (Figure 3) for participants in this training.

VI *Session 1* – At the beginning of the game, participants were at the starting point and faced 0 degrees. The participants only had 5 minutes to finish the task. The sound of applause indicated the end of the gallery. They had a 3-5 minute break before the next session.

VII *Training of Session 2* – I used the same training scene for session 2, allowing the participants to learn the audio feedback and controller in the VE.

VIII *Session 2* – In this session, the participants also had 5 minutes to finish the task. The sound of applause signaled finishing the game. They had a 3-5 minute break before the next session.

IX *Training of Session 3* – Session 1 and session 2 scenarios were replicates for study 3 training.

X Session 3 – Previous scenarios were replicates for session 3.

XI Questionnaire

XII Post Study Interview

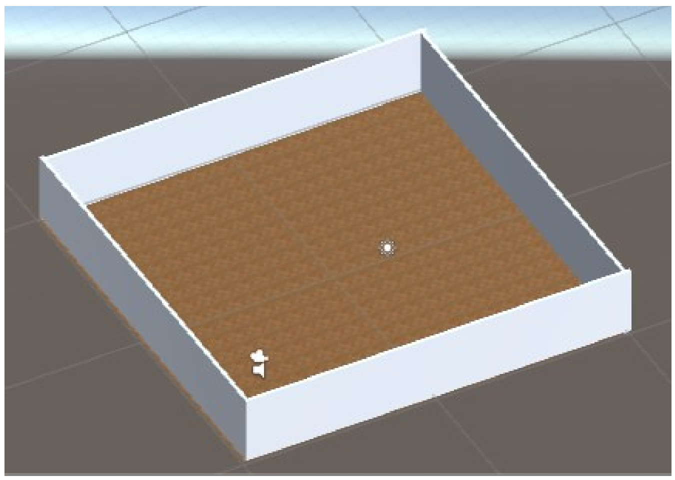


Figure 3: The training scene

4.5 Metrics

Total time: Total time spent for participants to finish the task.

Questionnaire: I used a modified Presence, Involvement, Flow, Framework 2 (PIFF2) questionnaire to validate uses' experiences and enjoyment in games after finishing all of the conditions. Other subjective questions related to 1) ease of use, 2) different experiences between distinct interfaces, 3) confidence of using different interfaces, and 4) comments about the virtual environment.

4.6 Hypotheses

I expected that participants to walk through this virtual gallery successfully by using different 3D audio.

ARCAF-H1: Participants will finish the task significantly faster with continuous audio feedback (Condition 2 and 3) than without any continuous audio feedback (Condition1).

ARCAF-H2: Participants will finish the task significantly faster in Condition 3 CWAPSAF than in Condition 2 CPSAF.

ARCAF-H3: Participants will prefer continuous audio feedback (Condition 2 and 3) to no continuous audio feedback (Condition1).

4.7 Result and Discussion

I ran the one way ANOVA and paired T tests to complete the data analysis.

4.7.1 Total Time

I found significant differences of Total Time between the Condition 1 WAAR and Condition 2 CPSAF (Table 2). The ARCAF-H1 could only be half accepted. The Total Time of CPSAF was significantly shorter than the Total Time of WAAR ($p = 0.05$). However, the Total Time of CWAPSAF was almost the same as the Total Time of WAAR (Table 1). Although both of the CPSAF and CWAPSAF provided continuous audio feedback, the evidence was weak to reject the ARCAF-H2. Normally, if users could have more information, it might help them to finish the task more effectively. Therefore, it did not help in this case. After the study, I asked the participants to compare and contrast their experience of each of the three conditions of this study. More than half of the participants mentioned the CPSAF was simple to follow. Some said the interaction method of CWAPSAF was confusing, and some felt even the similar continuous path sound source was not as simple as the CPSAF condition. CWAPSAF could provide more information but not better performance. The reasons for this are not clear. I assume the two most plausible reasons may be: 1) unfamiliar/complicated interaction methods, 2) too much information to be processed at the same time. I will investigate this area in future work. The

ARCAF-H2 was completely rejected.

Table 1: Descriptive of Total Time for each condition (s)

Condition	Mean	Std. Dev	Std. Error	Lower Bound	Upper Bound
WAAR	153.15	94.13	18.46	115.13	191.17
CPSAF	102.86	86.22	16.91	68.03	137.68
CWAPSAF	145.18	76.39	14.98	114.33	176.04

Table 2: Significant differences of Total Time between each condition (s)

Conditions	p Value	Z
WAAR - CPSAF	**0.050**	4.037
WAAR - CWAPSAF	0.739	0.112
CPSAF - CWAPSAF	***0.067***	3.510

I also ran the same data analysis based on the different sessions by time order. No significant differences were found (Table 3 and Table 4).

Table 3: Descriptive of Total Time for each session (s)

Condition	Mean	Std. Dev	Std. Error	Lower Bound	Upper Bound
1st Session	127.44	95.79	18.79	88.75	166.13
2nd Session	127.19	69.58	13.64	99.09	155.30
3rd Session	146.55	96.83	18.99	107.44	185.66

Table 4: Significant *differences of Total Time between each condition (s)*

Conditions	p Value	Z
1st Session - 2nd Session	0.992	0
1st Session - 3rd Session	0.478	0.512
2nd Session - 3rd Session	0.412	0.685

I found the Total Time between WAAR and CPSAF were highly correlated ($p = 0.027$). However, the Total Time between CWAPSAF and the other two conditions were not (Table 5). The reason causing this is not known. Similar to the significant difference of the Total Time, the only significant correlation I found was between the two conditions with simple audio feedback or reminders. There was no significant correlation found between the different sessions by time order (Table 6).

Table 5: Total Time correlations between each condition

Conditions	p Value	t
WAAR - CPSAF	**0.027**	2.344
WAAR - CWAPSAF	0.733	-1.885
CPSAF - CWAPSAF	0.071	-0.345

Table 6: Total Time correlations between each condition

Conditions	p Value	t
1^{st} Session - 2^{nd} Session	0.992	0.010
1^{st} Session - 3^{rd} Session	0.444	0.777
2^{nd} Session - 3^{rd} Session	0.371	-0.911

4.7.2 Questionnaire

As I could not find any significant differences from the questionnaire, the ARCAF-H3 was rejected. According to some participants` feedback information, they found that as long as they could memorise the path of the virtual gallery and apply expertly 4 direction beeps, they would rapidly and successfully find the exit. Therefore, I assume the most plausible reasons may be that some participants have a keener sense of direction so that they could master the task in a shorter time.

CHAPTER FIVE: STUDY TWO

5.1 The study 2: Distance and Direction Recognition (DDR)

In this study, the participants were asked to estimate the distance between the location of sound, their own orientation and the direction of the sound in a simple virtual space (Figure 2). They needed to walk to the location or face the direction by using an Xbox 360 controller. I took advantage of sound features to help participants, and presented the beep sound as audio reminders for them. To indicate the origin of sound in front of participants, I designed the beep sound with four different pitches at 0 degrees, 90 degrees, 180 degrees and 270 degrees. Thus participants could use these beeps to determine which direction they were facing.

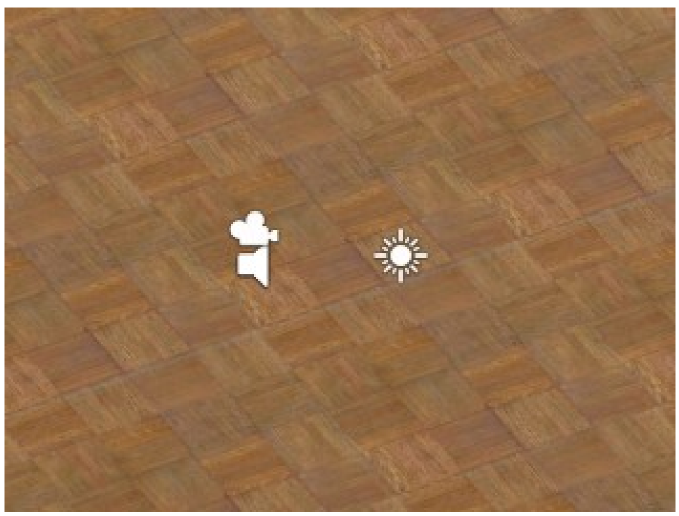


Figure 4: The virtual space

5.2 Population

I recruited 13 college students as participants from the same program in this study. None of the recruited participants had participated in Study 1. Their eyes were covered during the entire study, beginning with the training session. I eliminated one participant's data because he has a hearing problem. Each participant would earn 1 extra credit in their final grade as compensation.

5.3 Procedure

In this study, the participants needed to finish two sessions: 1) Estimating the distance by hearing the 3D audio; 2) Estimating the direction by hearing the 3D audio. I provided the sequence of these sessions randomly for each participant. Two scenes for each session included: 1) the training scene and 2) the experiment scene. The two sessions of this study shared the same scenes; the directions of the sound sources and interaction methods were different. The whole procedure is as shown in Figure 5.

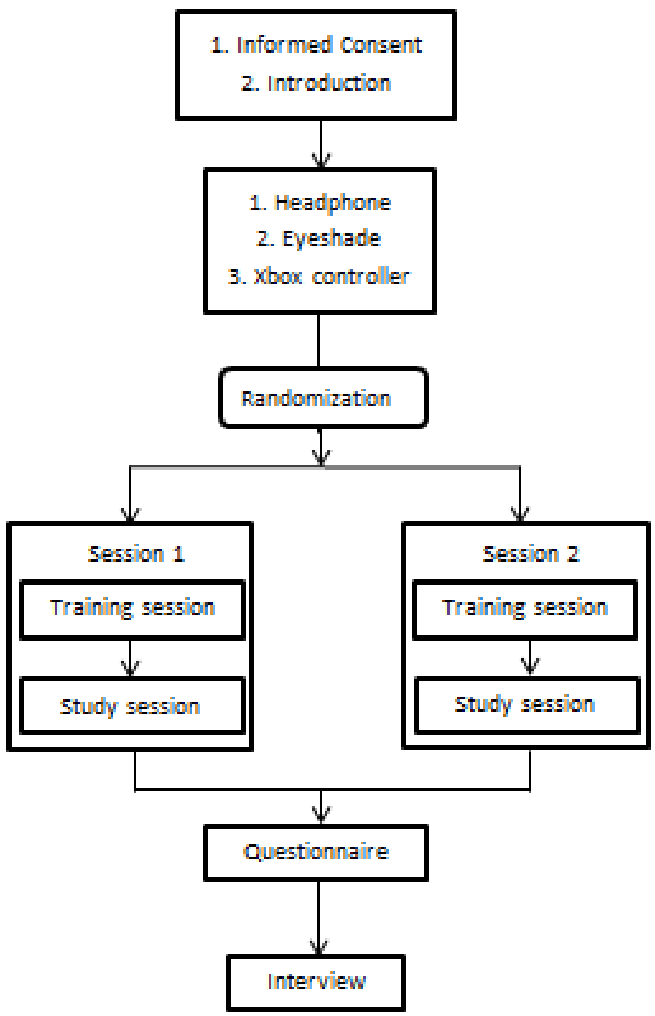


Figure 5: The flowchart of the study 1 procedure

I *Informed Consent and Introduction* – When the participants arrived, they were asked to read the informed consent and sign it if they did not have questions. I also briefly introduced the study to them.

II *Headphone, Eye Cover and Initialization* – The participants wore the headphones and the eye cover before starting the training session. Then they were requested to hold the Xbox 360 controller and attempt to use it. The experimenter checked if they could hear the sound from the headphones and adjusted a suitable volume for them.

III *Training of Session 1* – Subjects were presented 5 sound sources one by one, which were at 10 feet (3.048 meter), 15 feet (4.572 meters), 20 feet (6.096 meters) and 25 feet (7.620 meters). Next the participants were asked to walk to the sound location with the sound playing. When they arrived at the location, the sound would stop. I used the same virtual space (Figure 4) as a training scene for participants in this training.

IV *Session 1* – I randomly presented 5 different sound sources at 5 different distances in front of participants one by one, which were at 7.5 feet (2.286 meters), 12.5 feet (3.810 meters), 17.5 feet (5.334 meters), 22.5 feet (6.858 meters) and 27.5 feet (8.382 meters). I used different distances in session 1 than the training session to avoid double exposure, yet retained similar ranges as the training system. The pitch and the frequency of the sound provided the distance information between the sound source and the participants. Each sound source at different distances had a unique pitch and frequency. The higher pitch and frequency of the sound indicated further distance between participants and the sound source. Only one of the sound sources was presented for 5 seconds at a time, which means the participants could only hear the sound source for 5 seconds without any movements. Then the participants needed to estimate the location by moving their own position in the VE to the estimated sound source location by using the Xbox 360 controller. In this session, participants were only allowed to use the left thumb-stick moving forward and backward. Left and right movements were disabled. After the participants confirmed they arrived at the location of the sound source they heard, the experimenter would reset the participants' location in the VE and be ready to present the next sound source.

V *Training of Session 2* – There were four beep sounds with four different pitches at 0 degrees, 90 degrees, 180 degrees and 270 degrees. The participant could make clockwise rotations to hear 4 sound beeps. Additionally, I presented 4 sound sources clockwise at four positions which surrounded the participant in a circle with a 10 foot (3.048 meters) radius. These four sound sources were at 4 different degrees: 0, 90 180, 270 and existed for 10 seconds. Within these 10 seconds, the participant could make a rotation to face to the sound by using the Xbox 360 controller. I used the same virtual space (Figure 2) as Session 1 for participants in this training.

VI *Session 2* – I randomly generated ten sound sources one by one from 16 positions, which were around the participant on a circle with a 10 foot (3.048 meters) radius. These sound sources were at 10 different directions. 5 were well distributed in front and another 5 were well distributed behind participants. When the sound stopped, participants estimated the direction by facing the direction of the sound source using the Xbox 360 controller. In this session, participants were only allowed to use right thumb-stick to make rotations left or right.

VII *Questionnaire*

VIII *Post Study Interview*

5.4 Metrics

Errors of the Distance Estimation: The difference between the participants' estimated locations and the locations of the sound source.

Errors of the Direction Estimation: The degrees of the angle between the vectors of the participants are facing the direction and the vector from the participant to the sound source.

5.5 Hypotheses

DDR-H1: The errors between the participants' estimated distances and exact distances of the sound sources would not have significant differences.

DDR-H2: The errors between the participants' estimated directions and exact directions of the sound sources would not have significant differences.

5.6 Result and Discussion

I recorded the errors between the participants' estimations and exact locations and directions of each sound source. I used one-way ANOVA for the data analysis. There were no significant differences of the errors between each time of the participants' estimation. Therefore DDR-H1 and DDR-H2 were accepted.

After the participants had finished the entire study, I asked them, "Which session do you think is easier for you?" All answered estimating the direction was easier because those four beeps at 0 degrees, 90 degrees, 180 degrees and 270 degrees were very helpful. They complained that they did not exactly know the pace in the virtual environment, so they could not ensure they arrived at the location they estimated. Therefore, they thought if they could hear reminders at fixed distances in the estimated distance session, they would do better.

CHAPTER SIX: STUDY THREE

6.1 The Study 3: Fixed and Moving Sound Sources Recognition (FMSSR)

From the results of the first two studies, using 3D auditory feedback as user interfaces for people with VI in VEs seems possible. Therefore, I conducted this study to explore the differences of recognizing fixed and moving 3D sound sources in the VE between people with and without VI. Localizing the fixed sound and distinguishing the moving sound are two essential abilities for VI individuals when they utilize auditory assistance. Therefore, I designed two sessions in this study, one providing fixed sound and another one providing moving sound. Each session includes a training session and an experiment session. The order of the two sessions was counterbalanced randomized.

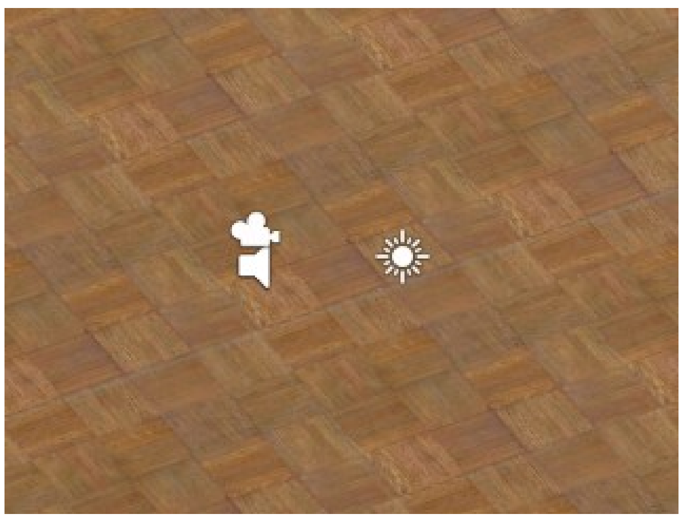


Figure 6: Virtual space

6.1.1 Fixed Sound Session

In this session, the participant should sit on the swivel chair. After the participant put on the headgear, I would generate a sound somewhere around the participant in the virtual environment. The participants should rotate the chair and try to pinpoint the sound's location. Participants might also rotate their head side to side, or up and down as needed to locate the sound. When participants have fixed on the sound, they should be facing the sound with their head and body, not just head or just body. The sound source would be randomly generated and participants would not necessarily turn at right angles to locate the sound.

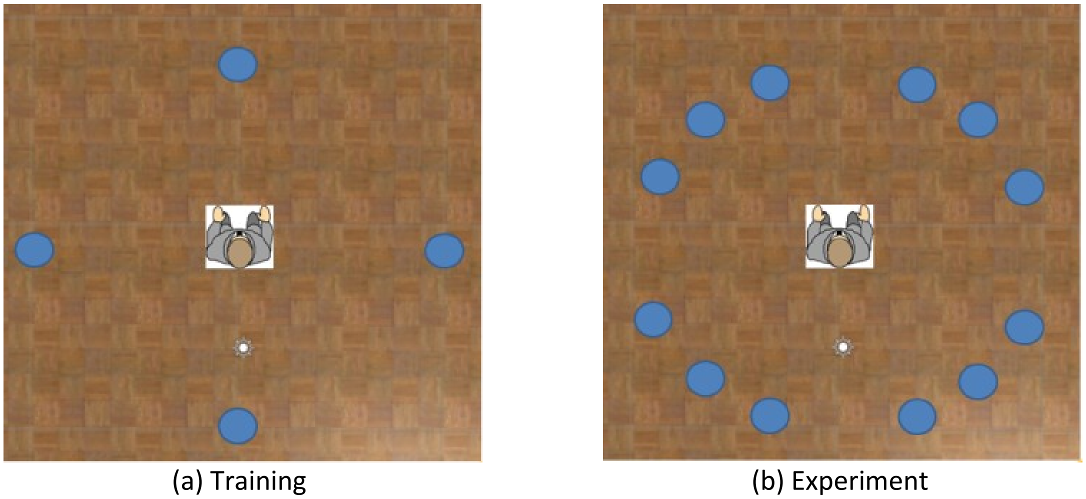


Figure 7: Sound position in Fixed Sound Session

6.1.1.1 Training

I presented 4 examples at 0 degree, 90 degrees, 270 degrees and 180 degrees which were in front of the participant, right hand side, behind, and on the left hand side (Figure 2, (a)). In addition, these examples surrounded the participant in a circle with a 10 foot (3.048 meters) radius. These four fixed sound examples were used to familiarize the fixed sounds for participants. Each sound existed for 20 seconds. Within these 20 seconds, the participant should rotate the chair and turn their body and head to face the sound. According to participant`s demand, I would repeat these examples for them and help them understand their task in this session. I used the same virtual space (Figure 7) as Fixed Sound Session for participants in this training.

6.1.1.2 Experiment

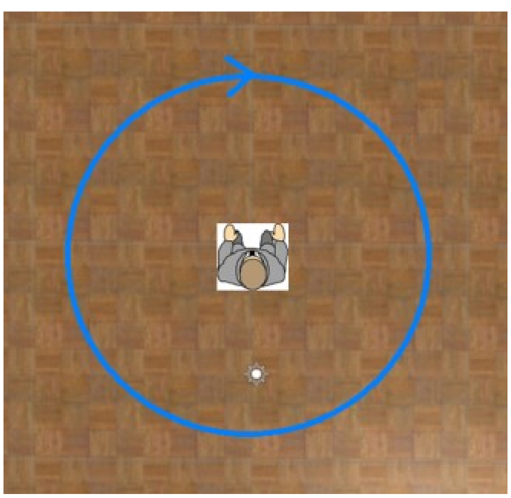
I randomly generated 10 sound sources one by one from 12 positions, which were around the participant on a circle with 10 feet (3.048 meters) radius. These 10 sound sources were at 10 different directions (Figure 2, (b)). 5 were well distributed in the front and another 5 were well distributed behind participants. Each sound lasted 20 seconds, and participants located the fixed sound by facing the direction of the sound source within 20 seconds.

6.1.2 Moving Sound Session

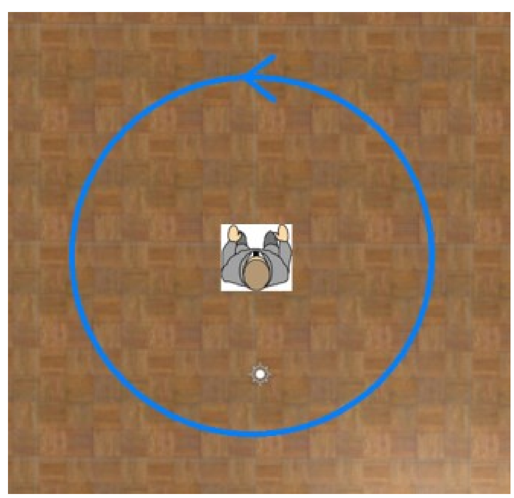
In the second session, the participant would sit in a stationary chair and hear a sound that was moving. Firstly, the task was to locate the quadrant of the sound, and then determine the direction of the sound.. In this session, participants might rotate their head to hear the sound source.

6.1.2.1 Training

In the training session, I would provide two examples for participants. The first sound was clockwise and the other would be counter-clockwise. These two examples help participants become familiar with hearing moving sound in VE. Each moving sound would last 20 seconds. When the sound stopped, they should confirm the movement of this sound. According to participants' demand, I would repeat these two examples and help them understand their task in this session. I used the same virtual space (Figure 1) as Moving Sound Session for participants in this training.



(a) Clockwise

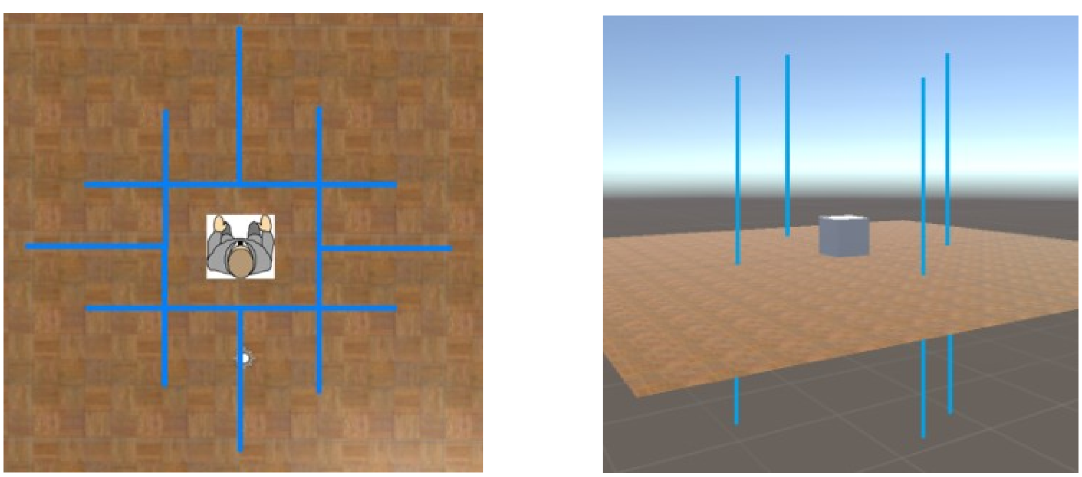


(b) Counter Clockwise

Figure 8: Sound movement in Training

6.1.2.2 Experiment

In this session, the space was divided into four quadrants, the front quadrant, which was in front of participants, the "back" quadrant which would be behind participants, the "left" quadrant which was on left side, and "right" quadrant which was on right side. I would randomly generate six moving sounds one by one for participants in each quadrant. Therefore, the total amount of the moving sounds was 24 (Figure 4). Also, those moving sounds would be linear movement in the experiment. I also disordered the sequence of four quadrants for each participant and ensured these sequences to be used averagely. In the front quadrant, a random moving sound would be from right to left, left to right, would go from bottom to top or top to bottom. The sound might also begin at a distance and move towards participants. In the back quadrant, the movements were the same as the front quadrant. The right and left quadrants might have a sound that started from behind participants and moved in front of them. Or, the sound may start in front of them and move behind them. The sound might also come from a distance and move towards them, or go from bottom to top or top to bottom. Each sound would last for only 10 seconds. Participants should determine the direction of the sound and tell experimenters the moving direction of the sound.



(a) Horizontal (b) Vertical

Figure 9: Sound movement trajectory (each blue line implies one trajectory has two directions)

6.2 Population

The total amount of participants was 39, 25 females and 14 males. Each participant would be paid $50 at the end of the study. I eliminated one participant's data because of cognitive disabilities. 20 of 58 eligible participants were people with VI, 18 were without VI. Each potential participant was interviewed by phone to verify qualification for this study to ensure homogeneity.. They were asked several questions, such as their name, contact information, age, vision condition and other sense conditions (to eliminate the possibility of hearing impaired participants). By asking these questions, the mental faculty of each potential participant was verified. I selected people who could understand the questions and answer them fluently. After the phone interview, the VI candidates finally selected as our study participants were only legally blind or blind, and did not have any mental/cognitive impairment or hearing problem. In addition, the age range of the selected VI participants was from 25 to 68, the mean value was 44.62, and the standard deviation was 10.64. The age range of the non-visually impaired participants was from 21 to 68, the mean value was 38.17, and the standard deviation was 14.34.

6.3 Procedure

The study procedure is shown in Figure 5. It took about one hour for each participant.

This study had two sessions, and I would disorder the sequence of these sessions for participants. Between these two sessions, there were several minutes for participants to have a break.

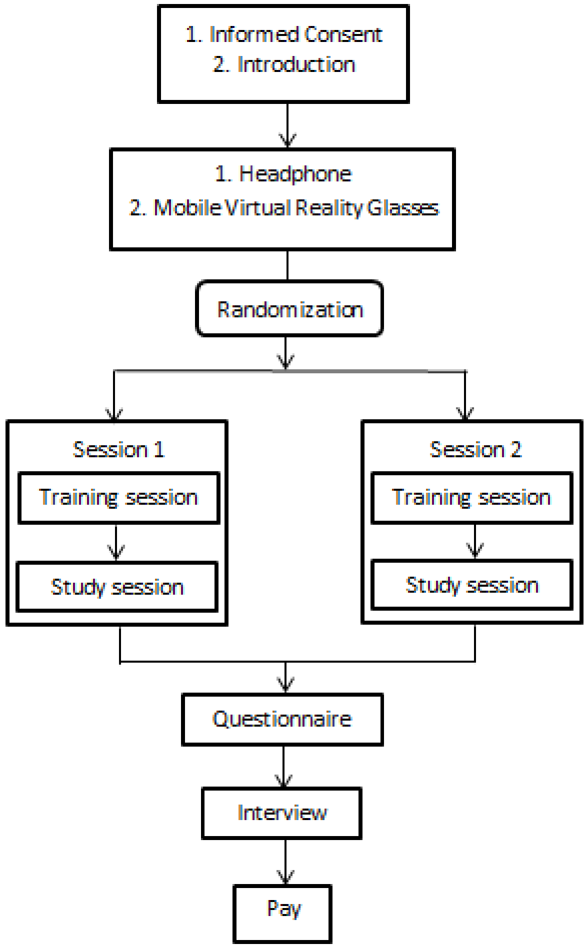


Figure 10: Sound movement trajectory (each blue line implies one trajectory has two directions)

I *Informed Consent and Introduction* – When the participants arrived, they were asked to read or listen to the informed consent. If they preferred to read it by themselves, then they could sign it if they did not have questions. If they preferred to listen to the informed consent, then I would ask them to give a verbal confirmation for us and sign it. I also briefly introduced the study to them.

II *Headphone, Mobile Phone, Headgear and Initialization* – The participants wore the headphones and the headgear before starting the training session. Then they were requested to sit on a swivel chair or a stationary chair, depending on the session they had. The experimenter checked if they could hear the sound from the headphones and adjusted a suitable volume for them. During each session, the experimenter also had to ensure headgear was even.

III *Session 1* – Before each study session, I would give participants a training session to become familiar with the VE. After participants finished the training session, I would begin the formal experiment for them.

IV *Session 2* – Participants would have a break for several minutes. As in previous sessions, there was a training session for participants to adjust to the VE. After participants finished the training session, I would begin the formal experiment for them.

V *Questionnaire* – There were two questionnaires for each participant, one of them for session 1 and another for session 2.

VI *Post Study Interview* – After they finished the study task sessions, the participants were asked to answer a final interview questionnaire about their experiences in the virtual environment compared to the real world.

VII *Pay* – Participants would each be paid $ 50 for their time and effort.

6.4 Metrics

Errors of the direction estimation (Degrees): The degrees of the angle between the vector the participants are facing and the vector from the participant to the sound source.

Errors of the moving direction estimation (Correct or Incorrect): The correctness of the participants` estimated moving directions of the sound sources compared to the actual moving direction of the sound source.

Total time (Seconds): The total time of finishing each direction or moving direction recognition.

Questionnaire: I used a modified Presence, Involvement, Flow, Framework 2 (PIFF2) questionnaire to validate uses' experiences and enjoyment in games after finishing all of the sessions. Other subjective questions related to 1) ease of use, 2) differences in experiences between the different interfaces, 3) confidence of using different interfaces, and 4) comments about the virtual environment.

6.5 Hypotheses

FMSSR-H1: The participants with VI will have more accurate direction recognition results than the participants without VI.

FMSSR-H2: The participants with VI will have a higher correct ratio of distinguishing the moving directions of the 3D sound sources than the participants without VI.

FMSSR-H3: The participants with VI will spend more time to finish each task than the participants without VI.

FMSSR-H4: The participants with VI will have a different experience in the same VE with only the 3D auditory feedback than the participants without VI.

6.6 Result and Discussion

Firstly, we used a Q-Q plot to determine the normality of the data. Since most of the data sets are ordinal, nonparametric methods can be employed as quite general assumptions regarding the population are used in these tests. We consider three methods: Mann-Whitney U-Test, and Fisher exact test. The Mann-Whitney U test is used to compare differences between two independent groups when the dependent variable is either ordinal or continuous, but not normally distributed. Similar with Mann-Whitney U-Test, the Fisher exact test is used to compare differences between two independent groups when the dependent variable is categorical data. It assesses how well the relationship between two variables can be described using a monotonic function.

6.6.1 Errors of the Direction Estimation

Among the 10 times performing tasks, I only found one significant difference between the participants with and without VI, which is the fifth time ($p = 0.042$). The FMSSR-H1 cannot be accepted.

6.6.2 Errors of the Moving Direction Estimation

There is no significant difference found between the participants with and without VI. The FMSSR-H2 is rejected.

6.6.3 Total Time

No significant difference was found here as well between the two population groups. The *FMSSR-H3* is rejected.

The *FMSSR-H1* to *FMSSR-H3* were rejected, which was not completely unanticipated. The reasons may vary. It was the first time all of the participants were exposed to such kind of VE with only 3D auditory feedback, and they might not perform the similar tasks before.

6.6.4 Questionnaire

Although I did not have many significant differences from other metrics, I found significant differences from more than half of the questions for both sessions.

1) In the session of recognizing the directions of static 3D sound sources around the participant in the VE only, I found significant differences between participants with and without VI on questions 1 ($p = 0.028$) and 10 ($p = 0.015$).

Question 1: “I felt that I was in the virtual environment.”

Question 10: “I would recommend it to my friends.”

2) In the session for distinguishing the moving directions of moving 3D sound sources around the participant in the VE only, I found significant differences between participants with

and without VI on questions 8 (p = 0.021).

Question 8: "The environment responded quickly to my actions."

3) I found significant differences between participants with and without VI on questions 6, 9, 12, 13, 14, 15, 16, 17, 18, 19 and 20 in both sessions.

Table 7: Questions of Significant Differences Between-Subjects

Questions	Recognizing Directions p-value	Distinguishing Moving Directions p-value
6. The virtual environment mattered to me.	< 0.01	< 0.01
9. I felt confident when I was in the environment.	< 0.01	0.02
12. I was in control of the action.	0.01	0.02
13. I would like to integrate this environment into my training or exercise routine.	< 0.01	< 0.01
14. Please rate how well you think you were performing actions WITH the sound feedback in the virtual environment.	0.01	0.01
15. Please rate how well you think you were performing actions in the virtual environment comparing to real world.	< 0.01	0.02
16. Do you think the sound feedback in the virtual environment helped you finish the task?	< 0.01	0.03
17. Do you think the sound feedback in the virtual environment may help you in the real world?	0.02	< 0.01
18. Do you feel comfortable when you CAN hear such sound feedback in the virtual environment?	< 0.01	0.03
19. Please rate how well you think you can map the task in the virtual world to the real world tasks.	< 0.01	< 0.01
20. Please rate how well you think you can feel the virtual world and the objects in the virtual world.	< 0.01	< 0.01

From the results of the questions, I can tell the participants with VI thought the VE was significantly different than the participants without VI. FMSSR-H4 is accepted.

CHAPTER SEVEN: GENERAL DISCUSSION

7.1 Phase 1

After participants had completed two studies in first phase, I asked them "Do you think the audio feedback is helpful?" Most thought the self-generated audio feedback was more helpful than the automatically generated because they felt that it was challenging to find the correct direction when they found the automatically generated feedback, while it was easier for them to find feedback generated by themselves because it was in front of them. This part of the survey showed that the participants thought following a sound source in front of them might be easier than finding a sound source beside or behind them. However, from the results, I found the participants finished the task in significantly shorter time with the automatically generated audio sources. It seems that the participants were efficient with self-generated sound sources and multiple audio feedbacks. The reasons why participants finished the tasks with a shorter time in Condition CPSAF might be: 1) They could focus more on the single audio feedback in the VE. 2) They could feel safer and do better in Condition CWAPSAF, so they persevered instead of finishing the task as soon as possible. 3) The Condition WAAR provided less information that forced participants to walk slower.

Why did participants feel it was difficult to find the correct direction? The headphones were not sensitive enough, which leads to the slight deviation. Also, the over ear headphones were not the best choice for using HRTF. When the distance was very large, a tiny deviation of direction could lead to the wrong direction so that users could not arrive at the correct location.

From the walking trajectory graphs, most of the participants could follow the CPSAF well. Thus, their walking trajectories were similar. To simplify the figure, I picked two participants' walking trajectories. For example, Figure 6 shows two participants' walking trajectories under Condition CPSAF.

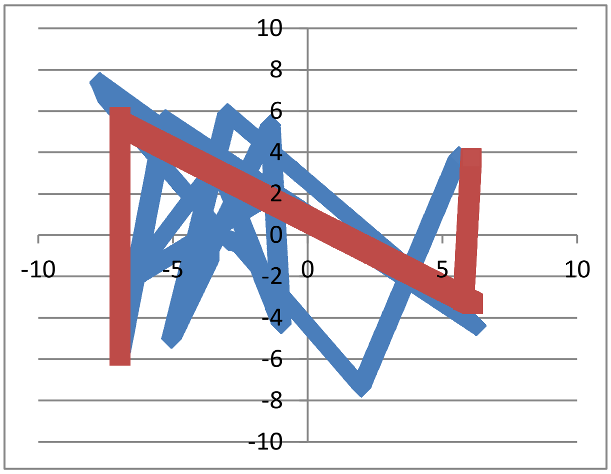


Figure 10: The walking trajectories of the same two participants (blue and red) as shown in Figure 3 under Condition WAAP (meter)

The walking trajectories of Condition WAAR and CWAPSAF were varied. In Figure 7 (left and right), there are two walking trajectories from the same two participants. Obviously, the two trajectories in both left and right figures in Figure 7 are different from each other. The walking trajectories demonstrated that different participants might have different understandings of using WAAR or CWAPSAF without any nonvisual feedback training. However, different participants should have a similar understanding of using CPSAF. We should consider CPSAF when designing interaction methods for non-vision VEs.

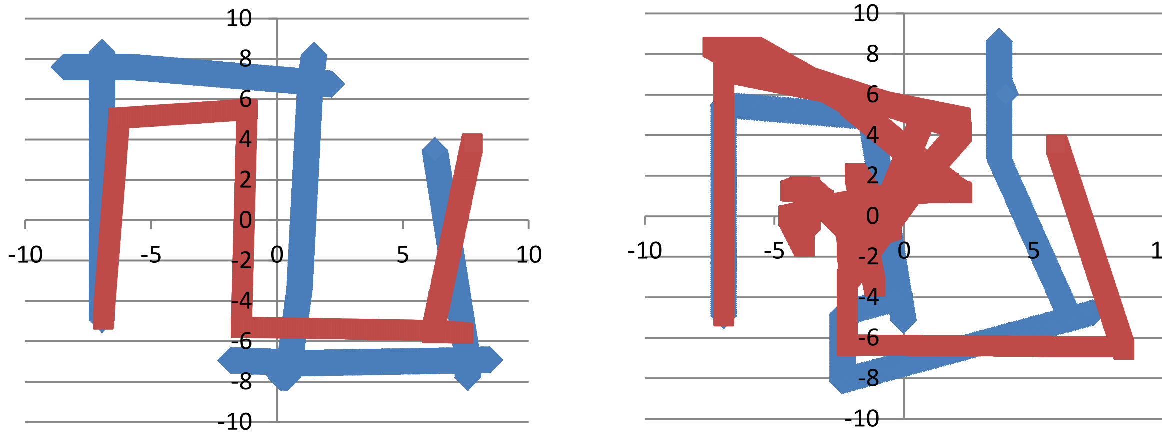


Figure 11: The walking trajectories of the same two participants (blue and red) as shown in Figure 6 under Condition (left) and WAAP (left) CWAPSAF (right) (meter)

7.2 Phase 2

After participants finished the third study, most of them felt that the Moving Sound session was more difficult for them than the Fixed Sound session. Furthermore, they explained it was necessary to pay more attention to the sound because they should follow the sound sustainability while the sound was moving. This might be due to lack of 3D audio training in VE, because this study was the first time they interacted in a non-visual VE.

In the post-Interview session, some VI persons informed me that for some examples they used auditory feedback to help them in their daily life. They usually utilize sound to localize position; for example, one VI participant mentioned she would localize the washing machine downstairs by hearing its noise. Furthermore, they claimed that this study inspired them to use their hearing to navigate and believed VR would play an essential role in their exercise routine. One VI participant mentioned he expected this technique could be implemented into his running exercise to help navigate the route.

According to data analysis, I found much more significant differences from the questionnaire other than the task performance. Some of the questions are related to presence.

Some are related to interactions. It is very interesting the participants with and without VI experienced the VE differently but finished the task with similar results. It may be the metrics we used could not represent the differences existing between the two populations or a long-term study is needed. However, considering other factors we examined, there is another "significant" difference between the participants with and without VI. A VE with only 3D auditory feedback means more informative feedback for a visually impaired user, but may be less informative feedback for a non-visually impaired user. Therefore, when answering those questions, the participants without VI may compare to those VEs with visual feedback, which could also cause differences. In the future, we may need to design the study to have all of the participants restricted comparing the auditory feedback only tasks.

CHAPTER EIGHT: CONCLUSION

From the Study 1 Audio Reminder vs. Continuous Audio Feedback: I learned that a continuous spatial audio feedback could significantly improve navigation performance in a VE without vision feedback. However, multiple spatial audio feedback or complicated interaction methods might slow down the navigation procedure. The developers should be able to use Head Related Transfer Function to design interaction methods for non-vision VEs.

From the Study 2 Distance and Direction Recognition: I noted that although the generic Head Related Transfer Function was not working perfectly, it allowed the users to recognize the directions of the sound sources in VEs. To be able to estimate the depth better, I may need to involve different feedback, such as using pitch or beeping frequency to help users better understand the depth. From the post-study interview, to help the users recognize the directions of the sound sources, allowing them to be able to know the direction they are facing is important.

From the Study 3 Fixed and Moving Sound Sources Recognition: I explored the differences of having the participants with and without VI in the same VE. Although they finished tasks with similar results, they still had significantly different feelings about the VE. As a designer without VI, we may not be able to use our previous experiences to design or develop any VE for people with VI.

REFERENCES

[1] National Center for Health Statistics. Available: http://www.cdc.gov/nchs/

[2] IAPB. Available: http://www.iapb.org/vision-2020/global-facts

[3] CDC. Available: www.cdc.gov/visionhealth/basic_information/vision_loss_burden.htm

[4] National Federation of the Blind. Available: https://nfb.org/blindness-statistics

[5] M. Griffiths, "Video games and clinical practice: issues, uses and treatments," British Journal of Clinical Psychology, vol. 36, pp. 639-642, 1997.

[6] G. Alankus, A. Lazar, M. May, and C. Kelleher, "Towards customizable games for stroke rehabilitation," in CHI, 2010, pp. 2113-2122.

[7] J. Burke, et al., "Serious Games for Upper Limb Rehabilitation Following Stroke," 2009, pp. 103-110.

[8] J. Deutsch, M. Borbely, J. Filler, K. Huhn, and P. Guarrera-Bowlby, "Use of a low-cost, commercially available gaming console (Wii) for rehabilitation of an adolescent with cerebral palsy," Physical Therapy, vol. 88, p. 1196, 2008.

[9] D. Goude, S. Björk, and M. Rydmark, "Game design in virtual reality systems for stroke rehabilitation," Studies in health technology and informatics, vol. 125, p. 146, 2007.

[10] A. Whitehead, H. Johnston, N. Nixon, and J. Welch, "Exergame Effectiveness: What the Numbers Can Tell Us," in the 5th ACM SIGGRAPH Symposium on Video Games Los Angeles, CA, USA, 2010.

[11] R. Guo, G. Samaraweera, and J. Quarles, "The Effects of VEs on Mobility Impaired Users: Presence, Gait, and Physiological Response " in The 19th ACM Symposium on Virtual Reality Software and Technology Singapore, 2013.

[12] K. Brütsch, et al., "Influence of virtual reality soccer game on walking performance in robotic assisted gait training for children," 2010.

[13] M. Jannink, et al., "A low-cost video game applied for training of upper extremity function in children with cerebral palsy: a pilot study," Cyberpsychology & behavior, vol. 11, pp. 27-32, 2008.

[14] E. Richard, V. Billaudeau, P. Richard, and G. Gaudin, "Augmented Reality for Rehabilitation of Cognitive Disabled Children: A Preliminary Study," Virtual Rehabilitation, 2007 pp. 102 - 108 22 October 2007 2007

[15] M. Sandlund, S. McDonough, and C. Häger-Ross, "Interactive computer play in rehabilitation of children with sensorimotor disorders: a systematic review," Developmental Medicine & Child Neurology, vol. 51, pp. 173-179, 2009.

[16] P. N. Wilson, N. Foreman, and D. Stanton, "Improving spatial awareness in physically disabled children using virtual environments," Simulation in Medicine (Ref. No. 1998/256), IEE Colloquium on p. 6/1, 2 Dec 1998 1998

[17] G. Burdea, "Virtual rehabilitation-benefits and challenges," Methods of information in medicine, vol. 42, pp. 519-523, 2003.

[18] J. Burke, et al., "Optimising engagement for stroke rehabilitation using serious games," The Visual Computer, vol. 25, pp. 1085-1099, 2009.

[19] A. E. Chisholm, S. D. Perry, and W. E. McIlroy, "Inter-limb centre of pressure symmetry during gait among stroke survivors," Gait & Posture, vol. 33, pp. 238-243, 2011.

[20] B. F. Katz, et al., "NAVIG: guidance system for the visually impaired using virtual augmented reality," Technology and Disability, vol. 24, p. 163, 2012.

[21] Y. Seki and T. Sato, "A training system of orientation and mobility for blind people using acoustic virtual reality," Neural Systems and Rehabilitation Engineering, IEEE Transactions on, vol. 19, pp. 95-104, 2011.

[22] M. Torres-Gil, O. Casanova-Gonzalez, and J. Gonzalez-Mora, "Applications of virtual reality for visually impaired people," WSEAS Transactions on Computers, vol. 9, pp. 184-193, 2010.

[23] Anna Malwina Kamelska, Krzyszt of Mazurek, "The Assessment of the Quality of Lifein Visually Impaired People withDifferent Level of Physical Activity," in Physical Culture and Sport Studies and Research, DOI: 10.1515/pcssr-2015-0001.
[24] Available: http://blog.hookeaudio.com/hooke-blog1/2015/3/5/the-laymans-guide-to-3d-audio
[25] Martin Rothbucher, Matthias Kaufmann, Tim Habigt, Johannes Feldmaier, and Klaus Diepold, " Backwards compatible 3D audio conference server using HRTF synthesis and SIP," in 2011 Seventh International Conference on Signal Image Technology & Internet-Based Systems.
[26] German Ramos and Maximo Cobos, " Parametric head-related transfer function modeling and interpolation for cost-efficient binaural sound applications (L)," in 2013 Acoustical Society of America, Pages: 1735–1738.
[27] Impulsonic for Phonon 3D. Available: https://www.impulsonic.com/products/phonon3d/
[28] S. Lee, G. S. Sukhatme, G. J. Kim, and C.-M. Park, "Effects of haptic feedback on telepresence and navigational performance," in Proceedings of the 14th International Conference on Artificial Telexistence, Seoul, Korea, November - December, 2004.
[29] M. K. A. Usoh, et al., "Walking > Walking-in-Place > Flying in Virtual Environments," in Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques, Los Angeles, USA, August, 1999, pp. 359 - 364.
[30] S. Razzaque, D. Swapp, M. Slater, M. C. Whitton, and A. Steed, "Redirected walking in place," in Proceedings of the Workshop on Virtual Environments 2002, Barcelona, Spain, May, 2002, pp. 123 - 130.
[31] J. M. Riley and D. B. Kaber, "The effects of visual display type and navigational aid on performance, presence, and workload in virtual reality training of telerover navigation," in Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting, Houston, USA, September - October, 1999, pp. 1251 - 1255.
[32] K. Kilteni, I. Bergstrom, and M. Slater, "Drumming in immersive virtual reality: the body shapes the way we play," Visualization and Computer Graphics, IEEE Transactions on, vol. 19, pp. 597-605, 2013.
[33] D. Perez-Marcos, M. V. Sanchez-Vives, and M. Slater, "Is my hand connected to my body? The impact of body continuity and arm alignment on the virtual hand illusion," Cognitive Neurodynamics, vol. 6, pp. 295-305, 2012.
[34] K. Kilteni, J.-M. Normand, M. V. Sanchez-Vives, and M. Slater, "Extending body space in immersive virtual reality: a very long arm illusion," PloS one, vol. 7, p. e40867, 2012.
[35] R. Guo and J. Quarles, "Differences in presence between healthy users and users with multiple sclerosis," in Perceptual Illusions in Virtual Environments (PIVE), 2012 IEEE VR Workshop on, 2012, pp. 1-6.
[36] H. Farmer, A. Tajadura-Jiménez, and M. Tsakiris, "Beyond the colour of my skin: How skin colour affects the sense of body-ownership," Consciousness and cognition, vol. 21, pp. 1242-1256, 2012.
[37] J.-M. Normand, E. Giannopoulos, B. Spanlang, and M. Slater, "Multisensory stimulation can induce an illusion of larger belly size in immersive virtual reality," PloS one, vol. 6, p. e16128, 2011.
[38] Y. Yuan and A. Steed, "Is the rubber hand illusion induced by immersive virtual reality?," in Virtual Reality Conference (VR), 2010 IEEE, 2010, pp. 95-102.
[39] P. Moreno-Ger, D. Burgos, I. Martínez-Ortiz, J. L. Sierra, and B. Fernández-Manjón, "Educational game design for online education," Computers in Human Behavior, vol. 24, pp. 2530-2540, 2008.
[40] G. M. Lemole Jr, P. P. Banerjee, C. Luciano, S. Neckrysh, and F. T. Charbel, "Virtual Reality in Neurosurgical Education: Part-Task Ventriculostomy Simulation With Dynamic Visual and Haptic Feedback," Neurosurgery, vol. 61, pp. 142-149, 2007.

[41] G. Denis and P. Jouvelot, "Motivation-driven educational game design: applying best practices to music education," in Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology, 2005, pp. 462-465.
[42] R. Mintz, S. Litvak, and Y. Yair, "3D-virtual reality in science education: An implication for astronomy teaching," Journal of Computers in Mathematics and Science Teaching, vol. 20, pp. 293-305, 2001.
[43] K. Johnsen, A. Raij, A. Stevens, D. S. Lind, and B. Lok, "The validity of a virtual human experience for interpersonal skills education," in Proceedings of the SIGCHI conference on Human factors in computing systems, 2007, pp. 1049-1058.
[44] A. Stevens, et al., "The use of virtual patients to teach medical students history taking and communication skills," The American Journal of Surgery, vol. 191, pp. 806-811, 2006.
[45] F. Ruotolo, et al., "Immersive virtual reality and environmental noise assessment: An innovative audio–visual approach," Environmental Impact Assessment Review, vol. 41, pp. 10-20, 2013.
[46] R. Nordahl, S. Serafin, and L. Turchet, "Sound synthesis and evaluation of interactive footsteps for virtual reality applications," in Virtual Reality Conference (VR), 2010 IEEE, 2010, pp. 147-153.
[47] R. Nordahl, et al., "Preliminary experiment combining virtual reality haptic shoes and audio synthesis," in Haptics: Generating and Perceiving Tangible Sensations, ed: Springer, 2010, pp. 123-129.
[48] S. Zhao, et al., "A real-time 3D sound localization system with miniature microphone array for virtual reality," in Industrial Electronics and Applications (ICIEA), 2012 7th IEEE Conference on, 2012, pp. 1853-1857.
[49] A. Londero, et al., "Auditory and visual 3D virtual reality therapy for chronic subjective tinnitus: theoretical framework," Virtual Reality, vol. 14, pp. 143-151, 2010.
[50] B. Dalgarno and M. J. Lee, "What are the learning affordances of 3-D virtual environments?," British Journal of Educational Technology, vol. 41, pp. 10-32, 2010.
[51] Pinaki Shankar Chanda, Sungjin Park," A Binaural Synthesis with Multiple Sound Sources Based on Spatial Features of Head-related Transfer Functions," in 2006 International Joint Conference on Neural Networks Sheraton Vancouver Wall Centre Hotel, Vancouver, BC, Canada July 16-21, 2006.
[52] K.-U. Doerr, H. Rademacher, S. Huesgen, and W. Kubbat, "Evaluation of a low-cost 3D sound system for immersive virtual reality training systems," Visualization and Computer Graphics, IEEE Transactions on, vol. 13, pp. 204-212, 2007.
[53] Z. Haraszy, D.-G. Cristea, V. Tiponut, and T. Slavici, "Improved head related transfer function generation and testing for acoustic virtual reality development," in Proceedings of the 14th WSEAS international conference on Systems: part of the 14th WSEAS CSCC multiconference, 2010, pp. 411-416.
[54] T. Lokki and M. Gr ohn, “Navigation with auditory cues in a virtual environment,” IEEE Multimedia.,12(2):80{86, 2005.
[55] C.-H. Ho, C. Basdogan, and M. Srinivasan, "Efficient point-based rendering techniques for haptic display of virtual objects," Presence, vol. 8, pp. 477-491, 1999.
[56] D. C. Ruspini, K. Kolarov, and O. Khatib, "The haptic display of complex graphical environments," in Proceedings of the 24th annual conference on Computer graphics and interactive techniques, 1997, pp. 345-352.
[57] G. C. Burdea, G. C. Burdea, and C. Burdea, Force and touch feedback for virtual reality: Wiley New York, 1996.
[58] G. Jansson, et al., "Haptic virtual environments for blind people: exploratory experiments with two devices," International journal of virtual reality, vol. 4, pp. 10-20, 1999.
[59] J. M. Romano and K. J. Kuchenbecker, "Creating realistic virtual textures from contact acceleration data," Haptics, IEEE Transactions on, vol. 5, pp. 109-119, 2012.

[60] Y. Yanagida, S. Kawato, H. Noma, A. Tomono, and N. Tesutani, "Projection based olfactory display with nose tracking," in Virtual Reality, 2004. Proceedings. IEEE, 2004, pp. 43-50.

[61] H. Matsukura, T. Yoneda, and H. Ishida, "Smelling screen: development and evaluation of an olfactory display system for presenting a virtual odor source," Visualization and Computer Graphics, IEEE Transactions on, vol. 19, pp. 606-615, 2013.

[62] G. R. White, G. Fitzpatrick, and G. McAllister, "Toward accessible 3D virtual environments for the blind and visually impaired," in Proceedings of the 3rd international conference on Digital Interactive Media in Entertainment and Arts, 2008, pp. 134-141.

[63] D. W. Schloerb, O. Lahav, J. G. Desloge, and M. Srinivasan, "BlindAid: Virtual environment system for self-reliant trip planning and orientation and mobility training," in Haptics Symposium, 2010 IEEE, 2010, pp. 363-370.

[64] O. Lahav, D. W. Schloerb, and M. A. Srinivasan, "Rehabilitation program integrating virtual environment to improve orientation and mobility skills for people who are blind," Computers & education, vol. 80, pp. 1-14, 2015.

[65] R. Iglesias, et al., "Computer graphics access for blind people through a haptic and audio virtual environment," in Haptic, Audio and Visual Environments and Their Applications, 2004. HAVE 2004. Proceedings. The 3rd IEEE International Workshop on, 2004, pp. 13-18.

[66] M. Petridou, P. Blanchfield, and T. Brailsford, "Involving the user with low or no vision in the design of an audio-haptic learning environment for learning about 3D shapes: The first approach," in Computer Science and Electronic Engineering Conference (CEEC), 2011 3rd, 2011, pp. 29-34.

[67] L. Picinali, A. Afonso, M. Denis, and B. F. Katz, "Exploration of architectural spaces by blind people using auditory virtual reality for the construction of spatial knowledge," International Journal of Human-Computer Studies, vol. 72, pp. 393-407, 2014.

[68] O. Balan, A. Moldoveanu, F. Moldoveanu, and M.-I. Dascalu, "Navigational 3D audio-based game-training towards rich auditory spatial representation of the environment," in System Theory, Control and Computing (ICSTCC), 2014 18th International Conference, 2014, pp. 682-687.