Towards Understanding the Capability of Spatial Audio Feedback in Virtual Environments for People with Visual Impairments

Miao Dong and Rongkai Guo
Kennesaw State University

ABSTRACT

This research analyzes if and how the Head Related Transfer Function (HRTF) can be used to support effective Human-Computer Interaction when people in a Virtual Environment (VE) without visual feedback. If sounds can be located in a VE by using HRTF only, designing and developing considerably safer but diversified training environments might greatly benefit individuals with visual impairments. To investigate this, we ran 2 usability studies: 1) to ascertain whether the HRTF could provide sufficient position information in VEs; 2) to learn whether the HRTF could provide sufficient distance and direction information in VEs. The results showed that a continuous audio feedback could help navigate in a VE without vision feedback.

Keywords: Assistive technology; HRTF; 3D Audio; user study

Index Terms: Human-center computing -> Interaction devices; Accessibility; Accessibility technologies

1 INTRODUCTION

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]. The 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].

Virtual Reality (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, 6]. One of VR technology's most important contributions is its use to ensure safety [6], such as the pilot training simulator 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 study of VR for people with visual impairments [7] warrants exploration. We believe that although people with visual impairment cannot enjoy the advantage of VR in visual aspects, but they might be able utilize auditory functions of VR.

This paper presents research that investigates *if* and *how* audio feedback can be used to support effective Human-Computer Interaction for people with visual impairments by means of virtual environments. The use of VR technology could greatly benefit individuals with visual impairments by designing and developing considerably safer but diversified training environments.

Specifically, in this paper we present results of two studies. 1) Investigating if/how the participants could navigate themselves with different types of 3D audio feedback in the same VE. 2) Investigating if the participants could recognize the distance and direction of a virtual sound source in the virtual environment (VE) effectively. The results of this project will help us understand more if/how we could build a 3D audio based VE for visually impaired users. In the meantime, the results could also suggest the design of VEs for general purposes that might increase the presence of VEs with 3D audio feedback.

2 BACKGROUND AND RELATED WORK

Most VR research work and applications use visual feedback to present virtual environments to users. For example, research work investigated how avatars could affect behaviours [8]. 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 a more immersive display are a few of the new techniques. While these are cutting edge innovations, most of these areas are focused on visual feedback and cannot be used for people with visual impairments.

2.1 Audio Feedback in VR

Audio feedback is the second most popular response in virtual environments. However, it is always in addition to visual feedback [9]. Head Related Transfer Function (HRTF) allows the developer to create a low-cost 3D sound system. HRTF is a function that describes how a sound source in a specific location will arrive to user's ear. 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 showed that users were able to identify the location of sounds simulated by using HRTF, with the exception that the top position was not well recognized [10]. In 2010, Haraszy, et al., presented an Acoustic Virtual Reality (AVR) implemented by an improved HRTF [11]. Their work demonstrated that with the aid of the artificial neural network, the generation of the HRTFs for people with visual impairments is possible. Nevertheless, there have been no usability studies investigating user experience of people with visual impairments.

2.2 VR for People with Visual Impairments

Unlike sighted persons, people with visual impairments use only their sense of hearing, touch and smell. Gareth et al. 2008, interviewed eight visually impaired expert users concerning an increase in the accessibility of 3D virtual environments for the blind and visually impaired [12].The suggested approaches for 3D navigation used audio and haptics. David et al., used haptic and 3D audio to develop BlindAid, which enabled blind people to learn about new environments on their own [13]. Orly et al., added spatial audio information to BlindAid for blind people navigating in virtual environments [14] to improve user experiences. Iglesias et al., developed the GRAB system to investigate the interaction process of visually impaired persons with haptic environments [15]. 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 [16]. However, 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 [17], but it was not for interaction. Oana et al., developed a navigational 3D audio-based game [18]. 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.

2.3 A Gap in Knowledge

One of the reasons visual feedback is necessary is to provide *continuous* feedback to users, which has remained unresolved for the visually impaired. To our knowledge, whether or how to use spatial audio techniques to develop continuous feedback in virtual environments for people with visual impairments 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 issues for training purposes. Based on the previous research, it is likely that spatial audio feedback might help people with visual impairments but the ability of a stand-alone spatial audio feedback virtual environment is unknown.

2.4 Our Contributions

Ludwig Wallmeier and Lutz Wiegrebe have since demonstrated that simulating echolocation in virtual environments did not bring the same perception as in the real world [19]. How, then, could we build a virtual world that people with visually impairments can use? We ran a depth and location recognition study. Additionally, we used object-generated sounds to provide information in VE as the navigation interfaces.

3 SYSTEM DESCRIPTION

We designed two studies. The task of the first study used different audio feedback to help participants walk through a virtual gallery, and the second study provided 3D audio at different positions and directions before allowing participants to determine location. Thus, we used UNITY to build our VE.

Audio: Participants wore an on-ear headphone during the study. The headphone blocked out all of the ambient sound from the real world to keep the participant focused on the VE. All of the sounds in the VE were generated by the Head Related Transfer Function (HRTF).

Xbox 360 Controller: Participants used an Xbox 360 controller to move in the virtual environment. We 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.

Headphone: Participants wore the on-ear headphone during the experiment.

Eye Cover: During the experiment, participants wore the eye cover to simulate the visual impairments.

4 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 the Figure 1 Left. There were three sessions provided randomly and each session had different audio reminders or feedback for participants. We 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 directions they were facing.

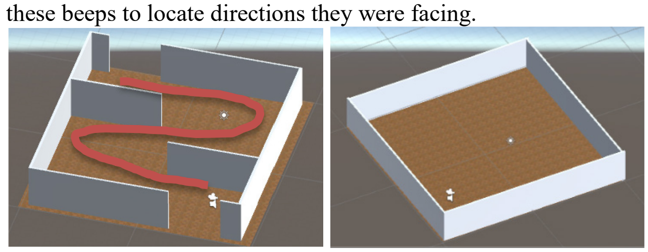

Figure 1: Left: The virtual gallery; Right: The training scene

4.1 The virtual gallery Environment

The study was conducted in a quiet, air-conditioned laboratory environment. Only the participant and experimenter were present.

4.2 Participant

We recruited 28 college students from our CSE department for participation in this study. We discarded two participants' data, because one of them had a hearing problem; the other was disrupted by a fire alarm test while participating in the study. Starting from the training session, participants` eyes were covered during the entire study. Each participant would earn 1 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, we 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): We 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 a 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): We generated one bell sound at each turning point. These sound sources were presented one by one, 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 last sound source.

Condition 3 – Continuous Wall Alert & Path Sound Audio Feedback (CWAPSAF): In this condition, we ALWAYS provided 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 where the path sound source was, 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 path to navigate.

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. We also briefly introduced the study to them.

II *Headphone, Eye Cover and Initialization* – The participants wore the headphones and the eyeshades before starting the study. Then they were requested to hold the x-box controller and try to use it. Furthermore, we checked if they could hear the sound from headphone 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. We used the same VE (Figure 1 Left) for participants in this training to let the participants get familiar with the path firstly.

IV *Training for Orientating in the VE* –We designed a user interface to assist the participants to identify the orientations in the VE. This training helped the participants understand this user interface. We used a virtual room which just had walls (Figure 1 Right) as the training scene for participants.

1) Session1

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. We used the same training scene (Figure 1 Right) 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.

2) Session 2

VII *Training of Session 2* – We 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 signalled finishing the game. They had a 3-5 minute break before the next session.

3) Session 3

IX *Training of Session 3* – The 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

4.5 Metrics

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

Questionnaire: We 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) differences in experiences between the different interfaces, 3) confidence of using different interfaces, and 4) comments about the virtual environment will be asked as well.

4.6 Hypotheses

We expected participants to walk through this virtual gallery successfully by using different 3D audios.

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

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

ARCAF-H3: Participants will prefer the continuous audio feedback (Condition 2 and 3) without any continuous audio feedback (Condition1).

4.7 Result and Discussion

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

4.7.1 Total Time

We 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, there was weak evidence 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, we asked the participant 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. We assume the most two plausible reasons may be: 1) unfamiliar/complicated interaction methods, 2) too much information to be processed at the same time. We 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

We 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

We 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 we 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
1st Session - 2nd Session	0.992	0.010
1st Session - 3rd Session	0.444	0.777
2nd Session - 3rd Session	0.371	-0.911

4.7.2 Questionnaire

We could not find any significant differences from the questionnaire. *So 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, we assume the most plausible reasons may be that someone has a keener sense of direction so that they could master the task in a short time.

5 STUDY 2: DISTANCE AND DIRECTION RECOGNITION (DDR)

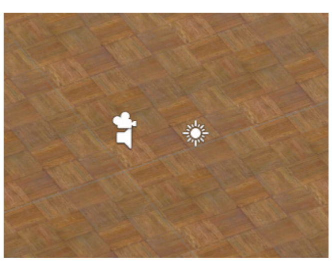

Figure 2: The virtual space.

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 direction by using an Xbox 360 controller. We 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, we 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.

5.1 Environment

Study 2 was conducted in the same laboratory as Study 1.

5.2 Participants

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

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. We 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.

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

II *Headphone, Eye Cover and Initialization* – The participants wore the headphone 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 headphone and adjusted a suitable volume for them.

1) Session 1

III Training of Session 1 – Subjects were presented with 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. We used the same virtual space (Figure 2) as a training scene for participants in this training.

IV Session 1 – We randomly presented 5 different sound sources at 5 different distances in front of participants one by one, which were 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). We 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 movement 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.

2) Session 2

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, we presented 4 sound sources clockwise at four positions which surrounded the participant a circle with a 10 feet (3.048 meters) radius. These four sound sources were at 4 different degrees: 0, 90 180, 270 which existed 10 for seconds. Within these 10 seconds, the participant could make a rotation to face to the sound by using the Xbox 360 controller. We used the same virtual space (Figure 2) as Session 1 for participants in this training.

VI Session 2 – We randomly generated ten sound sources one by one from 16 positions, which were around the participant on a circle with 10 feet (3.048 meters) radius. These sound sources were at 10 different directions. 5 were well distributed in the 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 will not have significant differences.

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

5.6 Result and Discussion

We recorded the errors between the participants' estimations and exact locations and directions of each sound source. We 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 finished the entire study, we 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 arrive at the location they estimated location. Therefore, they thought if they could hear reminders at fixed distances in the estimated distance session, they would do better.

6 GENERAL DISCUSSION

After participants completed the experiment, we asked them "Do you think the audio feedback is helpful?" Most thought the self-generated audio feedback was more helpful than the automatically generated one, because they felt that it was challenging to find the correct direction when they found the automatically generated one, while it was easier for them to find the one that generated themselves because it was in front of them. However, from the results, we found the participants finished the task in significantly shorter time with the automatically generated audio sources. It seems that the participants felt efficiently 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 might focus more on the single audio feedback in the VE. 2) They might 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 headphone was not sensitive enough, which leads to 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 are similar. To simplify the figure, we picked two participants' walking trajectories. For example, Figure 3 shows two participants' walking trajectories under Condition CPSAF.

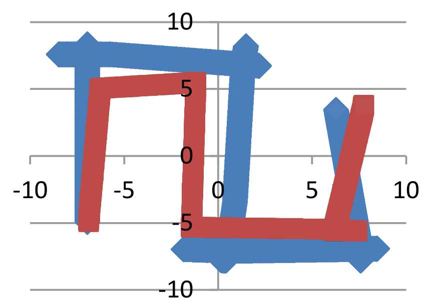

Figure 3: The walking trajectories of two participants (blue and red) under Condition CPSAF (meter)

The walking trajectories of Condition WAAR and CWAPSAF were varied. In Figure 4 and Figure 5, there are two walking trajectories from the same two participants as showed in Figure 4. Obviously, the two trajectories in both Figure 4 and Figure 5 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 similar understanding of using CPSAF. We should consider CPSAF when design interaction methods for non-vision VEs.

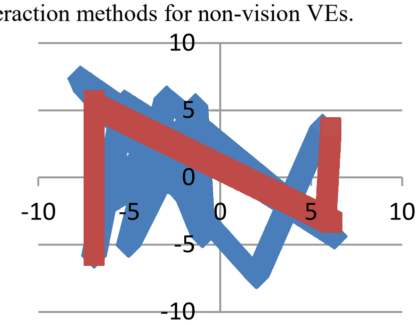

Figure 4: The walking trajectories of the same two participants (blue and red) as showed in Figure 3 under Condition WAAP (meter)

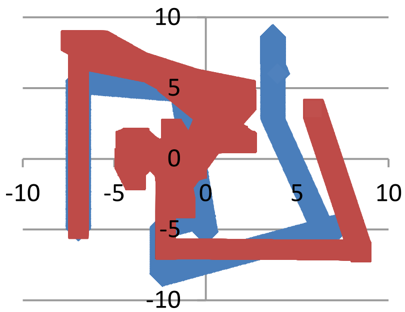

Figure 5: The walking trajectories of the same two participants (blue and red) as showed in Figure 3 under Condition CWAPSAF (meter)

7 CONCLUSION

From the Study 1 Audio Reminder vs Continuous Audio Feedback: we 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: we noticed that although the 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, we 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.

8 FURTHER WORK

Since available visually impaired participants are very limited, we did not recruit such participants for the purpose of this study. After this research, our next step is to recruit visually impaired participants in a future study to investigate the differences between the visually impaired participants and the non-visually impaired participants. We will continue investigating how to provide continuous audio feedback in VEs for visually impaired users, such as differences between using multiple of audio feedback vs single audio feedback for different tasks in VEs, using extra audio cues to provide walking pace information, etc. Furthermore, we intend to add comparisons with positive feedback from the people with visual impairments. Finally, we plan to learn how to build VEs specifically for the visually impaired, such as virtual training environments.

REFERENCES

[1] *National Center for Health Statistics*. Available: http://www.cdc.gov/nchs/, 2012. *Last accessed on March the 2nd, 2016.*

[2] *IAPB*. Available: http://www.iapb.org/vision-2020/global-facts, 2010. *Last accessed on March the 2nd, 2016.*

[3] *CDC*. Available: www.cdc.gov/visionhealth/basic_information/vision_loss_burden.htm, 2001. *Last accessed on March the 2nd, 2016.*

[4] *National Federation of the Blind*. Available: https://nfb.org/blindness-statistics, 2014. *Last accessed on March the 2nd, 2016.*

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

[6] K. Brütsch, T. Schuler, A. Koenig, L. Zimmerli, S. Merillat, L. Lünenburger, R. Riener, L. Jäncke, and A. Meyer-Heim, "Influence of virtual reality soccer game on walking performance in robotic assisted gait training for children," *Journal of neuroengineering and rehabilitation,* vol. 7, p. 1, *2010*.

[7] B. F. Katz, F. Dramas, G. Parseihian, O. Gutierrez, S. Kammoun, A. Brilhault, L. Brunet, M. Gallay, B. Oriola, and M. Auvray, "NAVIG: guidance system for the visually impaired using virtual augmented reality," *Technology and Disability,* vol. 24, p. 163, 2012.

[8] R. Guo and J. Quarles, "Differences in presence between healthy users and users with multiple sclerosis," in *Proceedings of the IEEE VR Workshop on Perceptual Illusions in Virtual Environments (PIVE)*, 2012, pp. 1-6.

[9] F. Ruotolo, L. Maffei, M. Di Gabriele, T. Iachini, M. Masullo, G. Ruggiero, and V. P. Senese, "Immersive virtual reality and environmental noise assessment: An innovative audio–visual approach," *Environmental Impact Assessment Review,* vol. 41, pp. 10-20, 2013.

[10] K.-U. Doerr, H. Rademacher, S. Huesgen, and W. Kubbat, "Evaluation of a low-cost 3D sound system for immersive virtual reality training systems," *IEEE Transactions on Visualization and Computer Graphics,* vol. 13, pp. 204-212, 2007.

[11] 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.

[12] 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.

[13] 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 *Proceedings of the IEEE Haptics Symposium*, 2010, pp. 363-370.

[14] 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.

[15] R. Iglesias, S. Casado, T. Gutierrez, J. Barbero, C. Avizzano, S. Marcheschi, and M. Bergamasco, "Computer graphics access for blind people through a haptic and audio virtual environment," in *Proceedings of the 3rd IEEE International Workshop on Haptic, Audio and Visual Environments and Their Applications (HAVE).* 2004, pp. 13-18.

[16] 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 *Proceedings of the Computer Science and Electronic Engineering Conference (CEEC), 2011 3rd*, 2011, pp. 29-34.

[17] 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.

[18] 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 *Proceedings of the 18th International Conference: Control and Computing (ICSTCC) System Theory*, 2014, pp. 682-687.

[19] L. Wallmeier and L. Wiegrebe, "Self-motion facilitates echo-acoustic orientation in humans," *Royal Society Open Science,* vol. 1, p. 140185, 2014.