Virtual reality application for blind people in unknown interior spaces

Nancy E. Guerrón[1,2] , Antonio Cobo[1,3], José J. Serrano Olmedo[1,3]

[1]Centre for Biomedical Technology. Universidad Politécnica de Madrid, 28223, Madrid-España
[2]Universidad de las Fuerzas Armadas ESPE. 170501, Sangolquí-Ecuador
[3]Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, 28029, Madrid- España

Abstract

Virtual reality applications (VRA) with 3D sounds on smartphones were used by six blind adults for indoor navigation. Each application provided users with three audio outputs with linear, logarithmic and customized response curves, that guided them during their tour within a real scenario that was previously virtualized using Unity 3D. The audio was transmitted, via low-power Bluetooth, from the mobile to the bone conduction headphones that were used to guide the user during his tour through an unknown space. The user's avatar moved through the virtualized environment simultaneously with the participant's movements in the real scenario. The movements of the user and the duration of the tour to go from a starting point to a target point were recorded each tour, in order to compare the efficiency and precision achieved by each sound during navigation. This research established that the VRA with a female voice and a customized response, improved the efficiency by 59% in locating target points within unfamiliar environments (p-value= 0.03614) compared to musical sound having a logarithmic response. Furthermore, it was found that the female voice and the 440Hz beeps increase the navigation precision compared to musical sound. 3D sound for indoor navigation could replace the sounds inside the place, with a small variation in efficiency and precision.

CCS Concepts

• ***Computing methodologies*** → Navigation, VRA • ***Hardware*** → Laser Sensor, remote control for video game

---

1. Introduction

When a blind person walks through a new place, he must overcome architectural barriers using a white cane or with the help of a guide dog. As possible solutions to these problems, social and adaptation technologies have emerged that compensate for functional limitations by using devices that provide information, which adapts to the user's capabilities, so they are able to access most locations and use available products and services [Soci00]. However, the non-availability or limitation of non-visual information in most public places, continues to cause insecurity, dependence and frustration for the visually impaired, who desires access to all available sites and services, without discrimination or barriers such as the sighted person [MFAP14]. Many assistive technology tools have been using ultrasonic devices or infrared light transceivers to gauge the distance to the obstacles. However, these devices exhibit several problems concerning the fact of alerting the user of possible obstacles during the navigation [Cald09].

To establish the position of an objective within an environment through haptic maps and audio-descriptions or verbal indications is usually a cognitively demanding task to blind people, since it depends on the user's self orientation ability in space[SH05].

Haptic-acoustic interactions ([PRBM07], [FRAD11], [CGMS17]) in VRAs are often used to train and educate blind people, which helps in creating a cognitive map of the site and locating points of interest within the map [PADK14], however the costs of haptic devices are gener-

ally high depending on the technology and services they provide [HWB06][PKAM16].

The use of 3D sonification on pre-elaborated relief maps allows the early exploration of exterior spaces [8] having the disadvantage of creating these supports manually in many cases [MGBP17].

Audio-touch interface [GoBP12a] uses sounds from musical instruments to encode the information concerning the obstacles and the borders of an environment, allowing users with visual disabilities to develop customized navigation strategies.

The strategy of using 3D audio as an output interface in VR applications allowed the identification of the object's position and structures of the environment [ŠRBV13] with little training. 3D sound considers the direction and distance between the user and the object to provide simple directional indications that defines the entire perceived auditory environment. The sound is received by the ear and then processed by the brain in the purpose of identifying and localizing the source of the incoming signals [BMMD14]. In our project, different 3D sounds were used on a smartphone to guide users to target points within a real space. A remote control of a video game was also used to change the type of audio response during the tour.

In this study, VRA with 3D sounds were used, as assistants in navigating blind people through unfamiliar spaces. The sound was transmitted via Bluetooth from a smartphone to the bone conduction headphones used by the participants. A structure sensor consisting of a laser and an

infrared emitter was connected to the mobile in order to track the user's movement in the test scenario and to update the position of the avatar in the application. We assume that VRAs with spatial sounds provide enough information for indoor navigation of blind people.

2. Material and methods

2.1 Participants

Six blind participants (5 men and 1 woman, average age 48.83, SD = 14.95) who had normal hearing conditions tested three VRAs (See Table 1). All the participants had become blind more than seven years ago. The participants had previous experience in the use of augmented reality applications, because they had previously participated in evaluating and testing of other applications.

2.2 Sound configuration

2.2.1 Sound 3D

3D Sound takes into consideration the direction and distance between the user and the object to provide simple directional indications for the location of objects within the place, eliminating the problem of generating a finite number of instructions during indoor navigation. Unity rendered the 3D sound using Audio Spatializer SDK with a generic HRTF (Head-related transfer function).
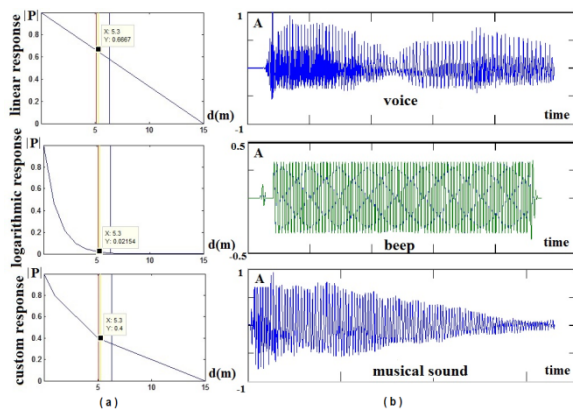

Fig. 1 *3D sounds a) auditory response b) audio signals*

The output audio responds to three power functions (linear, logarithmic and customized) (See Fig. 1a). The linear and the logarithmic response are software functions in Unity 3D, and the customized response was implemented with three linear sections to obtain an intermediate output power value between the linear and logarithmic response. The logarithmic response to 5.3 meters reaches a 4% of total power, while the linear response reaches 75% and the customized response reaches 40%. In addition, three elementary audio signals were defined: the female voice, the beep of 440Hz and the sound of a musical instrument (See Fig. 1.b), which resulted in nine combinations. The user received audio from one of these nine combinations.

The first experiment was performed using the female voice. During the first round, he/she received 3D audio which has a linear auditory power variation with the dis-

tance. During the second round, the user received 3D audio with a logarithmic response whereas during the third round he/she received audio with a customized response. Similarly, the second experiment was performed with a 440 Hz beep and the last experiment was made using a musical instrument's sound (See Table 1).

Table 1 *3D sound configuration in each experiment*

Experiment	Round 1	Round 2	Round 3
1	Voice-Lineal	Voice-Logarit.	Voice-Custom
2	Beep-Logarit.	Beep-Custom	Beep-Lineal
3	Cello-Custom	Cello-Lineal	Cello-Logarit.

2.2.2 Sound in the place

The female voice, the beep and the musical sound were recorded on an audio device and were used as independent sound sources. The sound sources were placed, one by one, in each of the target points established in the investigation.

The fourth experiment was performed using the female voice, throughout the entire tour. The fifth experiment used the 440Hz beep and the last experiment used the sound of the musical instrument.

2.3 Variables

Variables in Table 2 were assessed for three different sound types: female voice, beep of 440Hz and musical sound.

Table 2 *Variables assessed in the study*

Activity	Variables		
Navigation	efficiency	effectiveness	usefulness
Questionnaire	usability		

2.3.1 Navigation Process

Efficiency of the navigation process was determined by measuring the elapsed time to go from the starting point to the target point during a specific navigation route. The highest efficiency is obtained with shorter times of navigation.

Effectiveness was defined by the precision to reach a target point during a navigation route. The greater the number of target points reached using direct trajectories, the greater the effectiveness in the navigation process.

The usefulness was measured by navigation success within the scenario.

2.3.2 Questionnaire post-test

The usability of the navigation tool was determined as the average between the values: ease of use, user satisfaction, utility of the tool and user comfort. These values were obtained from a questionnaire designed to obtain information after each navigation experiment. The questionnaire was scored by users using the Likert scale where 5 is strongly agree, 4 like, 3 neutral, 2 dislike, 1 strongly disagree.

2.4 Implementation

The virtual reality application was implemented in Unity 3D as a video game for OS. The application shows the virtual representation of a smart house ($10 \times 7.5 \times 3m^3$) located within the Polytechnic University of Madrid (UPM). Despite the fact that visual information is irrelevant to blind people, it is useful for the process of development and for debugging. This living lab had four different spaces: a kitchen, a living room, a bedroom and a bathroom (See Fig. 2). In each space, there are some objects in different positions. These applications were installed on an iPhone 6S having an operating system of 10.2.1.
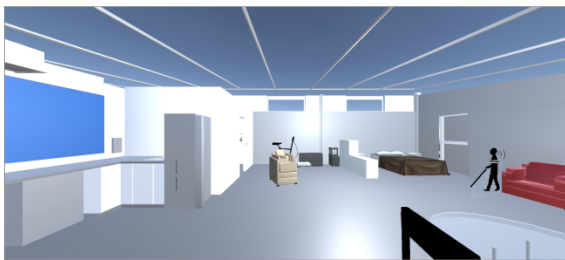

Fig. 2 *The virtual Smart House*

Both the Occipital structure sensor and the mobile phone tracked the movement of the participant in the smart house; this information was updated dynamically in the virtual world. Both devices were placed in a fanny pack at the height of the user's chest, so that he can walk freely around the house. A Bluetooth video game control was used to change the audio output response of the mobile. Each time the participant reached an objective point, the output audio was switched so that the user could walk to the next point.

2.5 Procedure

All participants signed an informed consent before beginning the study. The activities were recorded with two cameras, where two collaborators verified the sequence of the process and the registration of the information.

At the beginning of the session, each participant received a warm welcome and a member of the research team summarized to them the objectives of the study. Another member of the team explained how to use the application and placed the mobile phone with the sensor on the participant's chest.

Each participant was involved in six experiments; in three of them the user used an application with 3D sounds to go from one starting point to three target points, one by one; the distance established between two points was around 5.3m. In the next three experiments, the mobile application was not used; instead, an MP3 sound was activated at each target point. The experiments are described below:

First experiment: the female voice was reproduced in the bone conduction headphones. During the first round, the researcher selected the linear response with a remote control; when the participant reached the first objective, the researcher switched the response to logarithmic type, this response was maintained until the user reached the second objective; finally in the last round, the customized response was selected. When the user reached the last objective, there was a break of a few minutes, during which a member

of the team switched the application and verified the b capacity of the mobile. Next, the user was asked to answer a questionnaire related to previous experiment. Once finished, the user went on to the next experiment.

Second experiment: the 440Hz beep was reproduced, where similarly to the first experiment, the researcher switched the type of auditory response in each round, as shown in Table 1. Then, the user answered a questionnaire and when finished the user went on to the next experiment.

Third experiment: the application reproduced the 3D sound of a musical instrument. The researcher changed the type of response according to Table 1. Devices were removed from the participant's chest. Then, the user answered the questionnaire and continued with the next experiment.

Fourth experiment: here the female voice was used, but the mobile application was not used, instead, there were sound sources activated in each target point. The user walked from a starting point to three target points, one by one, where a member of the team activated one sound source at a time. When the participants reached the last target, they performed a questionnaire. Once completed, the user went on to the next experience.

Fifth experiment: the process of this experiment was similar to the previous one, except for the sound source where in the current experiment the 440Hz beep was used.

Sixth experiment: also, similar to the two previous experiences, except that here musical instrument was used as sound source.

During the six experiments, the researcher recorded the following parameters: duration of the tour, trajectory, proximity to the target point and success in the navigation process.

3. Results

3.1 Efficiency

The duration in seconds recorded by the user to reach each objective point is shown in Fig. 3. It could be seen that the shortest duration to reach the target points occurred with the customized audio responses using the female voice as 3D sound. The longest duration occurred with the logarithmic audio response.
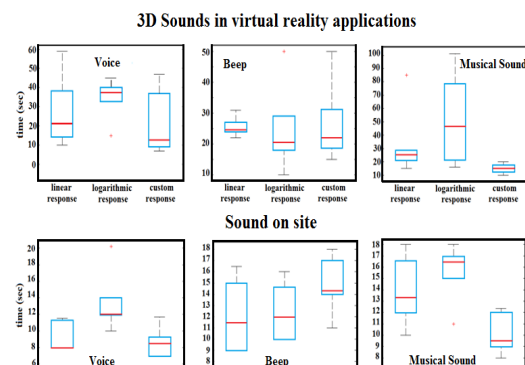

Fig. 3 *Duration of the navigation*

When the sound was reproduced on the target point, the blind participants advanced more securely towards the sound source. They overcame the obstacles with their white cane as they do in their daily lives. The lowest average time (10.4s, SD = 3s) used to reach a target point was obtained

using the female voice. The lowest standard deviation of 2.8 seconds was achieved when using a 440Hz tone as the sound source.

When 3D sound was reproduced in the bone conduction headphones, the lowest time was obtained with the customized response. The female voice having a customized response improved efficiency by 59% compared to the musical sound with logarithmic response.

The dependence of efficiency with the voice interface was assessed by a Friedman test that resulted statistically significant (Friedman chi-squared=4.27, df=3, p-value= 0.0022), where on the other hand, pairwise comparison was conducted by applying a Tukey Test. The voices in the place and in the bone conduction (with logarithmic response) showed a statistically significant difference (p-value= 0.0016), whereas no statistically significant differences were found among the other responses.

The dependence of efficiency with the beep interface was assessed with a Friedman test that resulted statistically significant (chi-squared=3.86, df=3, p-value= 0.0031), whereas pairwise comparison was conducted by applying a Tukey Test. The beeps in the place and in the bone conduction (with custom response) showed a statistically significant difference (p-value= 0.0045).

The dependence of efficiency with the sound musical interface was assessed with a Friedman test that resulted statistically significant (chi-squared=4.48, df=3, p-value= 0.0024), whereas pairwise comparison was conducted by applying a Tukey Test. The sounds in the place and the in the bone conduction (with logarithmic response) showed a statistically significant difference (p-value= 0.0034). The sounds with logarithmic and custom responses showed a statistically significant difference (p-value= 0.0038).

3.2   Effectiveness

When the sound was reproduced on the site, all participants reached the target point using direct trajectories. The following figure, Fig. 4, we can see the results of the type of trajectories and precision achieved during navigation using sounds 3D.
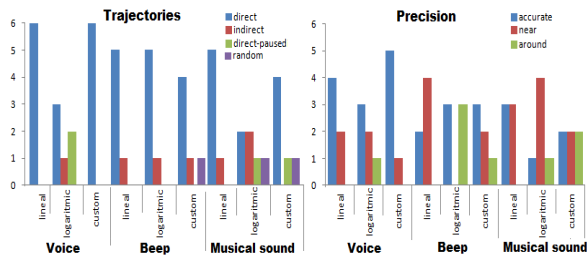

Fig. 4 *Trajectories and precision in navigation*

The female voice with customized response entailed a 66.66% improvement in precision compared to musical sounds with logarithmic response; it also entailed a 50% improved compared to beeps with lineal response. The voice with custom response improved a 50% in the number of direct trajectories, compared to musical sounds with

logarithmic response and supposed a 33.33% improvement compared to beeps with customized response.

The dependence of efficiency with the interfaces was assessed with a Friedman test, and the results are shown in the Table 3.

Table 3 *Efficiency results*

Voice	Beep	Musical Sound
F=4.16,df=2,p=0.048	F=4.31,df=2,p=0.044	F=2.54,df=2,p=0.127
Pair comparison	Pair comparison	was not found to be statistically significant
custom-logarithmic (p-value=0.048)	custom-logarithmic (p-value=0.037)	

3.3   Usefulness

Usefulness was modelled as a dichotomous variable for each target point. The Table 4 shows the results. (**1** represents the linear response; **2** the logarithmic response and **3** the custom response).

Table 4 *Usability test results*

Voice			Beep			Musical Sound		
1	2	3	1	2	3	1	2	3
F	F	F	F	F	F	F	F	F
F	F	F	F	F	F	F	F	F
F	F	F	F	U	F	F	F	F
F	F	F	F	F	U	F	U	U
F	F	F	F	F	F	F	U	F
F	F	F	F	F	F	F	U	U

f=finished task; u=unfinished task

All tasks were completed using the voice interface where there were no statistically significant results among the voice output responses. The voice interface implied a 16.66% improvement in utility, compared to the beep with logarithmic and personalized response; it also implies a 50% improvement in utility, compared to the musical sound having a logarithmic response and 16.66% compared to the musical sound with personalized response.

3.4   Usability

The results of the questionnaire are shown in the **¡Error! La autoreferencia al marcador no es válida.**.

Table 5. *Result tests usability*

Voice	a	b	c	d
Linear	3,67	3,67	4,08	3,67
Logarithmic	3,67	3,67	4,08	3,67
Custom	4,50	3,67	4,08	3,67
Beep				
Linear	3,58	4,33	4,33	4,33
Logarithmic	3,00	4,33	4,33	4,33
Custom	3,33	4,33	4,33	4,33

Musical Sound				
Linear	3,67	4,00	3,83	4,00
Logarithmic	3,25	4,00	3,83	4,00
Custom	3,00	4,00	3,83	4,00

a= Easy to use; b= Satisfaction; c = Utility; d= Comfort.

Sound of the female voice: The parameters that better scored with this interface are ease of use [GBP12b]. This could be due to the fact, that the female voice gave the name of target point. On the other hand, customized response reached the highest value.

Beep of 440Hz: Satisfaction, utility and comfort reached the best score, where some users mentioned that the beep remains clear despite the distance.

Musical Sound: Two of the participants mentioned that at the start of the application the sound was very low and they had to move in different directions until the sound gone clearer. According to some participants, the sound of a musical instrument is very pleasing to the ear.

4. Discussion and conclusions

The sound reproduced on the site gave all blind participants more confidence to navigate within the smart house. However, there are two important points to consider: if several sounds are reproduced inside a place, where other people coexist, these can be annoying because these invade their personal space and interrupt their concentration, which would be inappropriate and impractical. In addition, most people do not understand the limitations and needs of blind people [RDCG13]. Secondly, when several sounds are played at the same time, they can cause confusion and inadequate sound filtering, which leads to navigation errors [PADK14].

3D sounds in augmented reality applications allowed blind users to navigate through unknown interior spaces. The female voice and the beep reached from the bone conduction possess a level of precision and efficiency that is not significantly different from the sounds heard on the site. Greater efficiency and effectiveness was obtained through the use of customized response.

Users, who used virtual reality applications with 3D sounds to navigate within an unknown environment, reached the objective point 81.48% of times. The voice through bone conduction improved: accuracy, number of direct trajectories, the ease of use and greater percentage of success in completing the tour. This could be due to the fact that the user knows the name of the target point.

The 440Hz beeps were better heard by users. This interface obtained the highest score in the questionnaires subsequent to the usability test, except for the ease of use. As some users mentioned, the beep remained clear despite the distance.

The musical sound seemed to please the users as an output interface, but it did not achieve good results in the parameters of efficiency and effectiveness.

According to the participants, the use of 3D sounds in navigation through unknown interior spaces, allowed the user to be guided precisely to the objective points without requiring too much information.

References

[BMMD14] BALAN, OANA ; MOLDOVEANU, ALIN ; MOLDOVEANU, FLORICA ; DASCALU, MARIA IULIANA: Navigational 3D audio-based game-training towards rich auditory spatial representation of the environment. In: *2014 18th International Conference on System Theory, Control and Computing, ICSTCC 2014* (2014), S. 682–687 — ISBN 9781479946013

[Cald09] CALDER, DAVID J.: Assistive technology interfaces for the blind. In: *2009 3rd IEEE International Conference on Digital Ecosystems and Technologies, DEST '09* (2009), S. 318–323 — ISBN 9781424423460

[CGMS17] COBO, ANTONIO ; GUERRON, NANCY E ; MARTÍN, CARLOS ; SERRANO, JOSÉ JAVIER ; DEL POZO, FRANCISCO: Computers in Human Behavior Differences between blind people's cognitive maps after proximity and distant exploration of virtual environments Bd. 77 (2017), S. 294–308

[FRAD11] DE FELICE, FABIO ; RENNA, FLORIANA ; ATTOLICO, GIOVANNI ; DISTANTE, ARCANGELO: Hapto-Acoustic Interaction Metaphors in 3D Virtual Environments for Non-Visual Settings. In: *Virtual Reality* : InTech, 2011 — ISBN 978-953-307-518-1

[GoBP12a] GOMEZ, JUAN ; BOLOGNA, GUIDO ; PUN, THIERRY: Spatial awareness and intelligibility for the blind. In: *Proceedings of the 2012 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts - CHI EA '12*. New York, New York, USA : ACM Press, 2012 — ISBN 9781450310161, S. 1529

[GoBP12b] GOMEZ, JUAN ; DIEGO; ; BOLOGNA, GUIDO; ; PUN, THIERRY: Spatial awareness and intelligibility for the blind: audio-touch interfaces. In: *CHI '12 Extended Abstracts on Human Factors in Computing Systems*, 2012

[HeWB06] HEUTEN, WILKO ; WICHMANN, DANIEL ; BOLL, SUSANNE: Interactive 3D sonification for the exploration of city maps. In: *Proceedings of the 4th Nordic conference on Human-computer interaction changing roles - NordiCHI '06*. New York, New York, USA : ACM Press, 2006 — ISBN 1595933255, S. 155–164

[Kitc01] KITCHIN, R.: Cognitive Maps. In: *International Encyclopedia of the Social & Behavioral Sciences*, Pergamon (2001), S. 2120–2124 — ISBN 9780080430768

[MFAP14] MIESENBERGER, KLAUS ; FELS, DEBORAH ; ARCHAMBAULT, DOMINIQUE ; PEÑAZ, PETR ; ZAGLER, WOLFGAMG: Computers Helping People with Special Needs. In: SPRINGER (Hrsg.): *14th International Conference, ICCHP 2014*. Paris, 2014, S. 640

[MGBP17]   MASCETTI, SERGIO; ; GERINO, ANDREA; ; BERNAREGGI, CRISTIAN; ; PICINALI, LORENZO: On the Evaluation of Novel Sonification Techniques for Non-Visual Shape Exploration. In: *ACM Transactions on Accessible Computing* Bd. 9 (2017), Nr. 4

[PADK14]   PICINALI, LORENZO; ; AFONSO, AMANDINE; ; DENIS, MICHEL ; KATZ, BRIAN F G ; DENIS;, MICHEL ; KATZ, BRIAN F G: Exploration of architectural spaces by blind people using auditory virtual reality for the construction of spatial knowledge. In: *Elsiever: International Journal of Human-Computer Studies* Bd. 72, Elsevier (2014), Nr. 4, S. 393–407 — ISBN 10715819

[PKAM16]   POLZER, STEPHAN; ; KUNZ, ANDREAS; ; ALAVI, ALI; ; MIESENBERGER, KLAUS: Computers Helping People with Special Needs Bd. 9759 (2016), S. 587–593 — ISBN 978-3-319-41266-5

[PRBM07]   PUN, THIERRY ; ROTH, PATRICK ; BOLOGNA, GUIDO ; MOUSTAKAS, KONSTANTINOS ; TZOVARAS, DIMITRIOS: Image and video processing for visually handicapped people. In: *Eurasip Journal on Image and Video Processing* Bd. 2007 (2007)

[RDCG13]   RAFAEL, IVO ; DUARTE, LUÍS ; CARRIÇO, LUÍS ; GUERREIRO, TIAGO: Towards ubiquitous awareness tools for blind people. In: *Proceedings of the 27th International BCS Human Computer Interaction Conference* (2013), S. 38:1-38:5 — ISBN 9780000000002

[ScHe05]   SCOOTER, WILLIS ; HELAL, SUMI: RFID information grid for blind navigation and wayfinding. In: NINTH IEEE INTERNATIONAL SYMPOSIUM ON (Hrsg.): *Conference: Wearable Computers*, 2005

[Soci00]   SOCIALES, MINISTERIO DE TRABAJO Y ASUNTOS: *ERGONOMÍA Y DISCAPACIDAD*

[ŠRBV13]   ŠTOREK, DOMINIK; ; RUND, FRANTIŠEK ; ; BARÁTH, TOMÁŠ ; ; VÍTEK, STANISLAV: ICAD 2013: VIRTUAL AUDITORY SPACE FOR VISUALLY IMPAIRED. In: *METHODS FOR TESTING VIRTUAL SOUND SOURCE LOCALIZATION*, 2013, S. 33–36