

SPATIAL SOUND BASED SYSTEM FOR IMPROVING ORIENTATION AND MOBILITY SKILLS IN THE ABSENCE OF SIGHT

Oana BĂLAN¹, Alin MOLDOVEANU², Florica MOLDOVEANU³, Alexandru BUTEAN⁴

For the visually impaired people, navigating in unfamiliar environments represents a considerable challenge. In recent years, virtual reality based applications for the blind people have used the auditory channel as a way to convey perceptual and spatial information. This paper presents a sound localization experiment having the purpose of studying the orientation and mobility skills of 10 blindfolded subjects. Our research concluded that the use of the hearing sense as an alternate compensatory modality has a positive impact on the improvement of the localization skills and discrimination perceptual abilities in virtual reality based auditory environments.

Keywords: Virtual Reality; 3D sound; blind people; sound localization; orientation and mobility

1. Introduction

The use of audio interfaces for the development of rehabilitative tools designed to meet the needs of the visually impaired people proved to be highly efficient, as it promotes useful contextual information for navigation and independent mobility in real-world situations [1] [2]. Our work is concentrated towards evaluating the sound localization and orientation abilities of blind or blindfolded subjects in virtual auditory environments based on 3D sounds [3] [4]. In particular, the localization performance was evaluated, i.e. the capability of searching for sound sources at predetermined positions in the space. The experiment is concentrated on the localization performance in relation to both direction and distance perception. This study is part of a more complex project that comprises the design and assessment of a navigational assistive device for the

¹ PhD Student, Faculty of Automatic Control and Computer Science, University POLITEHNICA of Bucharest, Romania, e-mail: oanab_2005@yahoo.com

² Prof., Faculty of Automatic Control and Computer Science, University POLITEHNICA of Bucharest, Romania, e-mail: alin.moldoveanu@cs.pub.ro

³ Prof., Faculty of Automatic Control and Computer Science, University POLITEHNICA of Bucharest, Romania, e-mail: alin.moldoveanu@cs.pub.ro

⁴ PhD Student, Faculty of Automatic Control and Computer Science, University POLITEHNICA of Bucharest, Romania, e-mail: oanab_2005@yahoo.com

blind users. The experiment described here is a preliminary study that focuses on analyzing the basic performances that a sound localization application should have in order to provide accurate spatial information and to expand the navigational skills of the visually impaired. The experiment led to significant results concerning the perceptual behavior of 10 blindfolded subjects who were required to perform simple navigational tasks in a virtual reality based auditory environment.

2. Orientation and Mobility in Sound-based Virtual Environments

Orientation and mobility issues deal with the ability of blind users to determine their position in the environment, to recognize surrounding objects and to avoid obstacles or unsafe situations. The navigation process is divided into several phases: perception - the acquisition of sensory information, analysis - the management of the acquired data according to their relevance, selection - the selection of the most important information at a moment of time, planning - which involves decision-making tasks for performing a safe navigation and finally execution, which consists of carrying out the planned action [5].

2.1. Spatial sound

Spatial sound technology offers researchers the opportunity to deliver directional sounds at any position in space and to build synthesized auditory environments, without using any physical equipment [5]. The aim of this technique is to perfectly replicate the sound environment in the 3D virtual space. In order to develop a powerful audio virtual environment, we have to consider the sound localization performance and accuracy.

The use of virtual reality based environments helps blind people to develop strong spatial representation skills, as it enables them to gather, manipulate and transfer the acquired information from the virtual space to the real world. Furthermore, immersive simulators have proved to be useful for the enhancement of contextual learning and for the development of navigation and orientation abilities. The purpose of auditory virtual environments is to guide the visually impaired subjects towards the buildup of solid spatial cognitive maps that would help them to navigate safely in day-to-day situations.

The sound carries significant information concerning the constitutive elements of the environment and integrates strong clues regarding the location of nearby objects. 3D binaural sounds incorporate information about the position of sound sources in space - distance, trajectory and localization in both the horizontal and the vertical plane. Nonetheless, an important limitation lies in the lack of head-externalization and in the high frequency of “front-back confusions”, situation in which the listener perceives the sounds that are coming from the front as coming from the back and vice-versa [2] [3]. Spatial sound has been widely

used to study sound localization accuracy, as well as to explore more advanced cognitive behaviors, such as spatial map representation, mental imagery and distance estimation.

2.2. Spatial Cognitive Maps Development

Golledge [6] defined the concept of cognitive map as “the internal spatial representation of environmental information” [4]. A similar notion is “visual imagery”, which stands for “the representation of perceptual information in the absence of visual input” [5] [7]. For the sighted people, the visual sense is the primary modality to construct the spatial cognitive maps. In the case of the visually impaired individuals, there is a need for a compensatory strategy and for an alternative sensory channel that would counterbalance the lack of sight. As it has been already demonstrated, the hearing sense is very useful not only for spatially mapping the environment, but also for the transfer of localization skills from the virtual space to the real environment, with the aim of performing real-world navigation tasks, raising spatial awareness, promoting searching skills and enhancing orientation and mobility [8].

2.3. Crossmodal Auditive-Spatial Compensatory Strategies

When the brain lacks its natural sensory inputs, it undergoes a complex process of rewiring which consists in a wide range of neuroplastical changes. For instance, recent studies demonstrated that in the case of congenitally blind people, the occipital cortex (Fig. 1) is engaged in processing nonvisual sensory information. Moreover, the primary auditory cortex is strongly connected to the occipital cortex of the blind people, suggesting that this conjunction leads to the decoding of audio information in the visual cortex. Moreover, the spatial sound processing takes place in the regions of the occipital cortex that are specialized in performing visual spatial tasks in the case of sighted individuals [9].

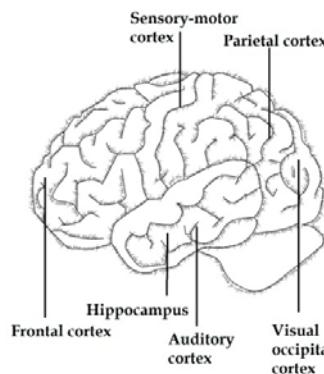


Fig. 1. The sensory regions of the brain [18]

In what concerns the perception of the distance to the sound source, the parietal cortex plays an important role in the localization of audio cues in both the

blind and the sighted subjects, while the inferior parietal lobe is designated to mediate cross-modal localization [10] [11]. Cross-modal plasticity and brain rewiring is emphasized by the complementary perceptual approach that is specific to blind and sighted individuals. For instance, in the case of sighted individuals, audio localization takes place in the frontal and temporal lobe, while blind people perform the same activities in the hippocampus (responsible for topographic orientation and memory) and in the visual occipital cortex (responsible for visual input processing) [12] [13]. Cross-modal learning and plasticity involve specific neural networks that adapt themselves in order to perform functions similar to the ones that are characteristic of the remaining senses. In this way, blind people develop an adaptable functional processing network that is unique and highly specialized [14].

3. The sound localization experiment

Our 3D audio software application, entitled Binaural Navigation, consists of two interface modules: the first module, called Binaural Navigation Test (Fig. 3), is a graphical interface that allows the user to move freely in order to locate the position of the incoming sound source; the second module, suggestively named Binaural Navigation Analyzer (Fig. 4), offers statistical tools to evaluate the results of the tests. The application was developed entirely in the C# programming language, using CSound as the main 3D sound processing tool.

In the actual experiment:

- We tested the sound localization ability in the horizontal plane (elevation 0 degrees in polar coordinates) of 10 blindfolded subjects, both normal sighted or nearsighted (aged 19-29) (Fig. 2);
- We included two sessions separated by a time interval of four days; the sessions were identical in regard to the performed tasks: the blindfolded users were required to locate in the horizontal plane (the plane of the user's ears) the source of the 3D sound that they heard over the headphones;
 - Each session consisted of twenty rounds; at each round, the sound source was positioned randomly at a fixed distance from the center of the screen (the starting point of navigation for the user), i.e. on a circle with a fixed radius (for this experiment, we set the radius at 150 pixels);
 - The audio cues used were continuous white Gaussian noise and an intermittent “ding” type signal; they were convolved in real time with the corresponding HRTFs (Head Related Transfer Function) from the CSound HRTF database in order to obtain a 3D binaural sound which corresponded to the direction of the sound source

relative to the position of the listener. CSound uses the MIT HRTF database which consists of KEMAR dummy-head impulse response measurements of 710 different positions that range in azimuth from 0 to 360 degrees and in elevation from -40 to +90 degrees. The white noise is a random signal that is considered to offer the best localization accuracy from all the randomly generated sounds. The spatial sounds were distributed like this: the rounds 1-5 and 11-15 used white noise, while rounds 6-10 and 16-20 used the “ding” sound;

- The pause between two consecutive rounds was one second for the sounds of the same type and three seconds when passing from white noise to “ding” or vice versa;
- The subject communicated with the interface by moving the mouse in the direction in which he considered that the sound source was located (conceptually, this meant he was moving his avatar in that direction);
- The localization approach was based on the employment of 3D binaural sounds and on the distance coding of the sound intensity: when the user got nearer to the source, the sound intensity increased, while, on the other hand, when he got farther, the intensity decreased until complete silence for a distance greater than 200 pixels. The formula for calculating the current intensity is:

$volume = d > dmax ? 0 : VOL_MIN + (VOL_MAX - VOL_MIN) * Math.Pow(1 - d / dmax, 2)$, where d is the current distance, dmax=200 pixels, VOL_MIN=0.05, VOL_MAX=1

- The range of the sound perception was of 200 pixels;
- For each round, we calculated and evaluated the following parameters:
 - o P1: the ratio of the distance travelled by the user to the minimum possible distance of 150 pixels (the radius of the circle that encompasses the sound source);
 - o P2: the percentage of good movements (“towards” the sound source);
 - o P3: the number of mouse movements;
 - o P4: the round completion time (in seconds).
- We also calculated the mean values of these parameters for the 10 rounds that used white noise and the 10 rounds that used the “ding” sound, as well as for all the 20 rounds.



Fig. 2. Blindfolded subjects during the sound localization experiment

The research stage consisted of a usability evaluation that followed the 7 phases settled by Shneiderman [15] in 1992: introduction to the virtual environment, software interaction, anecdotal record of the observations made by the end users, usability evaluation (qualitative evaluation based on a questionnaire applied at the end of the test session), session record (photographs taken during the interaction with the software), protocol reports of the session (comments, feedback) and suggestions for software redesign [16] [17].

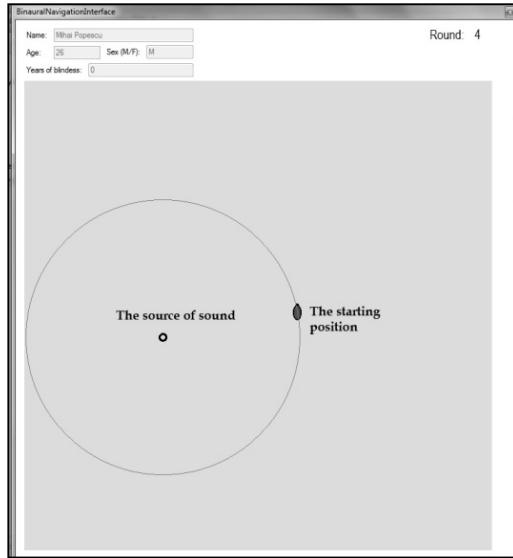


Fig. 3. The Binaural Navigation Test

The Binaural Analyzer (Fig. 4) module allows real-time visualization and audio playback of the performance of the users for each round. Segments considered as “good moves” were colored in green, while those which

corresponded to “wrong moves” were colored in red; an alternative procedure paints the road in progressive grayscale from source to destination. In addition, parameters P1-P4 can also be visualized and analyzed. Finally, the export button allows the data to be saved to an Excel file in order to be assessed for further analysis.

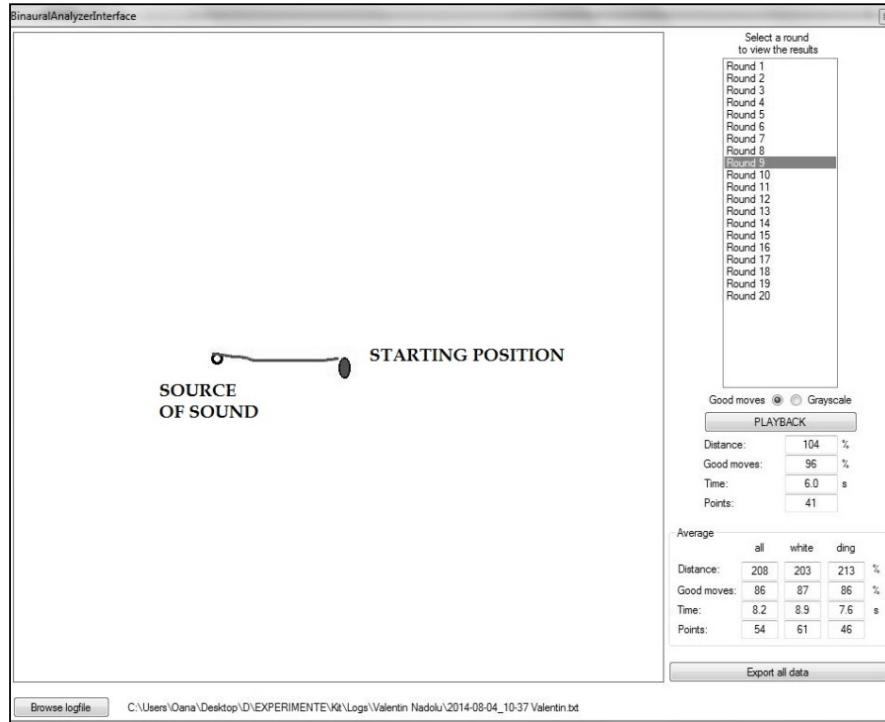


Fig. 4. The Binaural Navigation Analyzer. In this case, in the 9th round of the session (corresponding to the “ding” sound), the user succeeded to obtain a percentage a 96% of correct travel decisions (P2), a ratio of 1.04 of distance travelled to minimum possible distance (P1) and a round completion time of 6 seconds (P4)

4. Results and discussion

Table 1 lists summary descriptive statistics (mean, standard deviation – SD, minimum and maximum value) for each of the 4 studied parameters for both types of sound (white noise and “ding”) in the two sessions of the experiment.

First thing to note is that average results obtained during sound localization based on the “ding” signal are significantly weaker than those obtained using the white noise, for both sessions of the experiment (we used a Student t-test at a significance level $p \leq 0.1$) (Table 2). As a result, we conclude

that the “ding” sound is more difficult to be localized than the white noise, offering an uncertain acoustic directional representation.

Table 1

Descriptive statistics (mean, standard deviation - SD, minimum and maximum value) for each type of sound and experiment session

	WHITE NOISE								
	Session 1				Session 2				
	P1	P2 (%)	P3	P4 (s)	P1	P2 (%)	P3	P4 (s)	
Mean	3.62	77	87	12	3.57	79	89	12	
SD	1.51	6	24	4	3.02	8	35	5	
Min	1.78	66	62	8	1.63	66	30	4	
Max	6.09	87	129	23	10.87	89	140	21	
	“DING” SOUND								
	Session 1				Session 2				
	P1	P2 (%)	P3	P4 (s)	P1	P2 (%)	P3	P4 (s)	
	Mean	7.66	73	124	18	4.13	76	105	15
	SD	7.41	10	59	8	2.36	8	39	6
Min	2.05	59	47	7	1.68	61	33	5	
Max	23.26	86	217	31	8.95	90	166	31	

Table 2

Average results of the researched parameters

Type of sound	Session 1				Session 2			
	P1	P2 (%)	P3	P4 (sec)	P1 (%)	P2 (%)	P3	P4 (sec)
White noise	3.62	77	87	12	3.58	80	89	13
“Ding”	7.66	73	124	18	4.14	76	105	15
Ratio “ding”/white noise	2.1	0.9	1.4	1.4	1.2	0.9	1.2	1.2

The results of the second session show a significant increase of the values of P2 - percent of correct travel decisions ($t=1.706$ at a significance level $p\leq 0.1$). Thus, the subjects succeeded to improve their localization accuracy between the two sessions. Also, there have been recorded important enhancements for the “ding” sound localization in the second session – the results getting closer to the values for white noise. Thus, there are significant improvements for P1 ($t=-1657$, $p\leq 0.1$) and for P2 ($t=4.40$, $p\leq 0.1$). In average, the percent of correct travel decisions increased with 3% (within the experiment’s error range) and the ratio of distance travelled to the minimum possible distance reduced with 44% between

the two sessions of the experiment. We can conclude that although it offers a weaker initial spatial perception, the “ding” sound still presents a good potential of enhancement through training.

Nonetheless, most users did score improvements in the second session for both types of sound (Table 3, Fig. 4); overall, at least 60% of the subjects improved in all the 4 studied parameters. The percentage of subjects who scored better results in the second session range from 60% for P3 to 80% for P1 for both types of sound. In general, the percentage of correct travel decision, which we consider the most important parameter of our study, as it offers clear clues regarding the auditory spatial judgment of the listener, increased between the two sessions for 70% of the subjects, conducting to a better localization accuracy and directional spatial discrimination.

Table 3
Improvements in the sound localization ability, measured in the 2nd session, compared to 1st one

Type of sound	P1	P2	P3	P4
	Percentage of subjects with improvements in 2 nd session			
White noise (10 rounds)	80%	70%	60%	70%
“Ding” (10 rounds)	80%	70%	60%	70%
Both white noise and “ding”	80%	70%	60%	70%

Both the qualitative results and the applied questionnaire (Table 4) showed that most of the users (90%) believed that the white noise is more efficient for localization than the “ding” sound.

Table 4

Usability questionnaire applied at the end of the test session			
CLASSIFICATION	STATEMENT	TRUE	FALSE
MOTIVATION & INTERACTION	The sound localization application is interesting and motivating	100%	0%
	I consider that this type of application would be useful for the future development of an assistive device for the blind people	100%	0%
	I would like to use in the future this application, in order to train my sound localization abilities	70%	30%
USABILITY	The interface is accessible and easy to use.	100%	0%
	The software is intuitive	100%	0%
3D SOUND	3D sounds are efficient for spatial	100%	0%

PERCEPTION	localization.		
	I easily identified the sound source in space.	100%	0%
	It was easier to me to identify the source using white noise.	90%	10%
	It was easier to me to identify the source using the “ding” sound.	10%	90%
	It was difficult to me to identify the sound sources coming from the back and from the front.	70%	30%

Furthermore, in a short interview that took place after each test session, some subjects stated that even if they believe that the white noise is efficient for navigation, it can become quite annoying if used for long time. At the same time, most of subjects considered that the other sound is misleading because of the short breaks between the “ding” sounds, which causes them to lose the direction.

Furthermore, 70% of the subjects were confronted with front-back confusions, a typical issue for binaural sounds [18] [19]. Visual analysis with the dedicated tool (Fig. 3) indicated that this confusion influenced most of the wrong navigational decisions.

The subjects, who use headphones on a large scale to listen to music- 80% of the subjects use them frequently, unanimously appreciated that the 3D sound localization application is challenging and motivating and that it would be useful for the development of an assistive device for the blind people. Moreover, they believed that the 3D binaural sounds are efficient for navigation and orientation tasks and that there is no major difficulty in what concerns the localization of the sound source. Also, the questionnaire indicated that the subjects considered the interface to be intuitive and easy to be learnt and used. In addition, 70% of the subjects appreciated that they would like to use the software in the future to train their sound localization abilities (Table 4).

5. Conclusions

The purpose of this study was satisfactorily accomplished by evaluating the localization abilities of 10 blindfolded subjects who navigated in a 3D virtual auditory environment in order to find the location of a target sound source. The sonification technique relied on two types of sound (Gaussian white noise and a repetitive “ding” sequence) and on the inverse proportional sound intensity encoding of distance (the sound intensity decreases with distance). The results showed that learning and training can significantly improve the orientation and mobility skills and the general spatial representation of blindfolded sighted subjects. This improvement is independent of the type of sound employed for localization, although white noise was obviously preferred by our subjects.

Besides, white noise leads to a more accurate localization within a shorter time limit. Thus, the blindfolded subjects actually reduced the length of the path travelled from the starting point to the sound source and recorded an increase of the rate of correct directional decisions towards the actual position of the incoming sound. Moreover, they got accustomed with the perception of 3D sounds, became able to find optimal routes, enhanced their directional and decision-making abilities and developed a spatial auditory representation of the environment.

Acknowledgement

The work has been funded by the Sectoral Operational Programme Human Resources Development 2007-2013 of the Ministry of European Funds through the Financial Agreement POSDRU/159/1.5/S/132395 and POSDRU/159/1.5/S/134398. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 643636 "Sound of Vision".

R E F E R E N C E S

- [1]. *J. Sanchez*, "User-Centered Technologies for Blind Children", in Human Technology Journal, Vol. 4(2), 2008, pp. 96-122
- [2]. *A. Moldoveanu, O. Balan, F. Moldoveanu*, "Training System for Improving Spatial Sound Localization", Proceedings of the eLSE Conference, Bucharest, 2014
- [3]. *O. Balan, F. Moldoveanu, A. Morar, V. Asavei*, "Experiments on Training the Sound Localization Abilities: A Systematic Review", Proceedings of the eLSE Conference, Bucharest, 2014
- [4]. *J. Sanchez, M. Campos, M. Espinoza, L. Merabet*, "Audio Haptic Videogaming for Developing Wayfinding Skills in Learners Who are Blind", Proceedings of the 19th International Conference on Intelligent User Interfaces, Haifa, Israel, February 2014
- [5]. *B.F.G Katz, L. Picinali*, "Spatial Audio Applied to Research with the Blind", Advances in Sound Localization, INTECH, 2011, pp. 225-250
- [6]. *R.G. Golledge*, "Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes", JHU Press, 1999
- [7]. *D. Kasky*, "Revision: is visual perception a requisite for visual imagery?" in Perception, 2002, 31(6), pp. 717-31
- [8]. *E. Connors, E. Chrastil, J. Sanchez, L. Merabet*, "Virtual environments for the transfer of navigation skills in the blind: a comparison of directed instruction vs. video game based learning approaches", Frontiers in Human Neuroscience, 8:223, 2014
- [9]. *S. Maidenbaum, S. Abboud, A. Amedi*, "Sensory substitution: Closing the gap between basic research and widespread practical visual rehabilitation", Neuroscience and Biobehavioral Reviews, 2014
- [10]. *J. Sanchez, M. Saenz, A. Pascual-Leone, L. Merabet*, "Navigation for the blind through audio-based virtual environments", Proceedings of the CHI Conference, 2010, Atlanta, Georgia, USA

- [11]. *S. Levy-Tzedek, S. Hanassy, S. Abboud, S. Maidenbaum, A. Amedi*, “Fast, accurate reaching movements with a visual-to-auditory sensory substitution device”, Restorative Neurology and Neuroscience, 30 (2012), pp. 313-323
- [12]. *E. Striem-Amit, A. Amedi*, “Visual Cortex Extrastriate Body-Selective Area Activation in Congenitally Blind People “Seeing” by Using Sounds”, Current Biology, 24(6), 2014, pp. 687-92
- [13]. *O. Collignon, G. Vandewalle, P. Voss, G. Albouy, G. Charboneau, F. Lepore*, “Functional Specialization for Auditory–Spatial Processing in the Occipital Cortex of Congenitally Blind Humans”, PNAS Early Edition, 108 (11), 2011, pp. 4435-4440
- [14]. *C.C.H Chan, A.W.K Wong, K.H Ting, S. Whitfield-Gabrieli, J. He, T.M.C. Lee*, “Cross Auditory-Spatial Learning in Early-Blind Individuals”, Human Brain Mapping, 33(11), 2012, pp. 2714-2727
- [15]. *B. Schneiderman*, “Designing the user interface: strategies for effective human-computer interaction”, Addison-Wesley, 1992
- [16]. *J. Sanchez*, “User-Centered Technologies for Blind Children”, Human Technology 4(2), 2008, pp. 96-122.
- [17]. *G. Wersenyi*, “Auditory representations of a graphical user interface for a better human-computer interaction”, Springer Verlag, Berlin, 2010, pp. 80–102
- [18]. *O. Balan, A. Moldoveanu, F. Moldoveanu, M.I. Dascalescu*, “Navigational 3D audio-based game- training towards rich auditory spatial representation of the environment”, in The International Conference on System Theory, Control and Computing, Sinaia, Romania, 2014
- [19]. *G. Wersenyi*, “Virtual localization by blind persons”, in J. Audio Eng. Soc., vol. 60, 2012, pp. 568–579