HAPTIC-AUDITORY PERCEPTUAL FEEDBACK BASED TRAINING FOR IMPROVING THE SPATIAL ACOUSTIC RESOLUTION OF THE VISUALLY IMPAIRED PEOPLE

Oana Bălan, Alin Moldoveanu
University POLITEHNICA of Bucharest,
Faculty of Automatic Control and Computers,
060042, Bucharest, Romania
oanab_2005@yahoo.com

Hunor Nagy, György Wersényi
Széchenyi István University,
Department of Telecommunications,
H-9026, Győr, Hungary
wersenyi@sze.hu

Nicolae Botezatu, Andrei Stan, Robert-Gabriel Lupu
"Gheorghe Asachi" Technical University of Iaşi,
70050, Iaşi, Romania
robert@tuiasi.ro

ABSTRACT

3D binaural sounds play an important role in the development of navigational systems for the blind people. The use of generic HRTFs in virtual auditory displays significantly affects the acoustic spatial resolution and the listener's ability to make localization judgments regarding the sound sources situated inside the cone of confusion. The aim of this paper is to investigate whether haptic-auditory feedback based training can enhance sound localization performance, front-back discrimination and the navigational skills of the visually impaired people. In our experiments, we assessed the sound localization performance of nine visually impaired subjects before and after a series of haptic-auditory training procedures aimed to enhance the perception of 3D sounds. The results of our tests demonstrate that our subjects succeeded to improve their sound localization performance, reduced the incidence of angular precision and reversal errors and became able to build an effective spatial representation map of the acoustic environment.

1. INTRODUCTION

Just as vision, hearing is able to provide a wide range of stimuli and perceptual information concerning the objects and events that build up our surrounding space. Sound localization, the process of identifying the direction of a sound source in space, is influenced by the existence of binaural and monaural cues that give the listener significant information concerning the position of the acoustic objects [1]. The binaural cues are responsible for decoding the location of a sound source in the horizontal plane, while the monaural cues play an essential role in vertical discrimination and front-back disambiguation. The Head Related Transfer Function (HRTF), the response that describes how the ear perceives a sound coming from a particular direction in space, is highly influenced by the anatomical characteristics of the human body (such as the size and shape of the pinna, torso and shoulders), the head asymmetry and the ears placement [2]. The specific differences in the anatomy of the ears, head and body do not allow for the use of the same HRTFs for all the listeners.

As a result, the characteristics of the spectrum shape of the sound that points to a certain source location vary among subjects, leading to ambiguous interpretations of the same HRTFs [3]. In virtual auditory displays, the use of 3D sounds generated from non-individualized HRTFs conducts to a lower level of acoustic localization accuracy and to an increase in the incidence reversal errors – ambiguous localization judgments indicating to the opposite direction than the actual position of the sound source, such as front-back and back-front confusions [4], [5], [6]. Over the years, various experiments have proven that the brain possesses a high level of adaptation and training-driven plasticity in response to auditory and haptic cues [7]. Sensory substitution is a crossmodal process, as the information received either as auditory or haptic cues is communicated to the visual cortex to be built into a spatial cognitive image of the environment [8]. The user of a virtual auditory environment can adapt to the perception of 3D sounds synthesized from non-individualized HRTFs by learning to identify the spectral features of the sound. This can be achieved through training and by providing auditory and visual (in the case of the sighted subjects) or haptic perceptual training (in the case of the visually impaired people) [9]. In this way, the users can build a solid sensorial and cognitive association map between the location of the sound source in space and the 3D acoustic stimulus that is characteristic to that particular direction. This paper aims to present a new approach towards improving the spatial auditory resolution (sound localization accuracy and front-back disambiguation) and for enhancing the ability to navigate and orient in virtual auditory environments (based on 3D sounds synthesized from generic HRTFs) of the blind and visually impaired people, through haptic and auditory perceptual feedback based training. The results of our experiments concluded that the proposed training strategy improved the listeners' sound localization and navigational performance and facilitated the development of a solid spatial representation of the virtual acoustic environment.

2. METHOD

2.1. Participants

In our experiment, we trained and tested the sound localization accuracy of 9 visually impaired subjects (6 women and 3 men), aged 27-52 (mean age=42 years, SD=8.64). Their percent of residual vision was between 0% and 20%. All the participants were naïve to the purpose of the experiment. They reported normal hearing and gave their full consent before the start of the tests.

2.2. Sound stimuli

The sound stimuli consisted of a train of combined continuous Gaussian white and pink noises [10], [11], [12] that were perceived simultaneously, but at varying acoustic levels, according to the direction of the sound source in space [13]. Our sonification approach is aimed at reducing the incidence of reversal errors (front-back and back-front confusions) in the virtual auditory environment and uses a method based on the spectral coloration of the sound which consists of listening simultaneously to two types of noises with different spectral characteristics (white and pink noise). We used a combination of white and pink noise in varying proportions, so that at 0 degrees to the front the listener heard 100% white noise and 0% pink noise, consequently in the first quadrant the level of pink noise increased and the level of white noise decreased, reaching 50% white noise and 50% pink noise at 90 degrees (to the right side) and 0% white noise and 100% pink noise at 180 degrees (in the back). In the third and in the fourth quadrants, the rate of white noise increased and the percent of pink noise decreased, reaching 50% white noise and 50% pink noise at 270 degrees (to the left side). The formula for calculating the level of perceptible white and pink noise for a given angle from the polar coordinate system is:

- $pink = angle / 180;\ white = 1 - pink;\ (0 \leq angle \leq 180)$
- $white = (angle-180) / 180;\ pink = 1 - white;\ (180 < angle < 360)$

(1)

In addition to this, we used a continuous, repetitive (with very short breaks of 250 ms between two consecutive bursts) "ding" type signal (Figure 1).

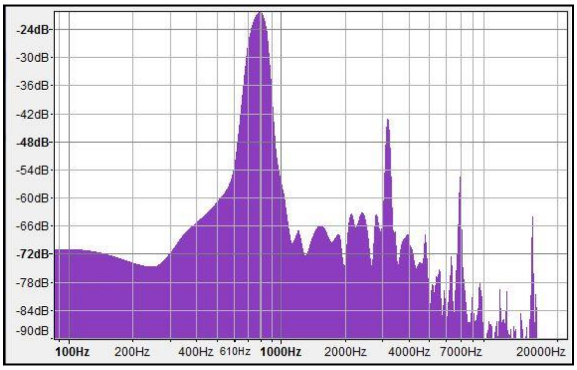


Figure 1: The spectral profile of the "ding" sound

Both the white-pink noise combination and the "ding" sounds have been convoluted in real time with the non-individualized HRTFs from the MIT database [14], using the CSound programming language for sound processing in order to obtain the 3D sound that corresponds to the desired angular direction in the horizontal plane [15]. It should be emphasized that all the 3D sounds are synthesized at 0 degrees elevation in the median plane, at the level of the listener's ears.

2.3. The haptic belt

The haptic device (Figure 2, Figure 3) that we used in our experiment is composed of the following parts:

- A USB-Wireless Gateway Device (UWGD), that allows the system to be controlled by the PC (it connects to the PC via USB).
- The Haptic Actuator Device (HAD), which controls the haptic actuators (Eccentric Rotating Mass – ERM motors) by receiving commands from the UWGD and executing them.
- 12 vibration motors – each vibration motor is fixed along a stick in order to provide easy handling. The vibration motors have been placed at 30 degrees distance on the haptic belt, all around the listener's head.

Our experiment was composed of a pre-test session (in which we evaluated the visually impaired subjects' sound localization accuracy and ability to navigate in a virtual auditory environment by listening to 3D binaural sounds that have been synthesized from non-individualized HRTFs), a training session (based on haptic-auditory perceptual feedback, aimed to help the visually impaired subjects to get used to the perception of 3D sounds delivered through headphones) and a post-test session (identical with the pre-test session), which had the purpose of assessing the degree of acoustic spatial resolution improvement achieved as a result of the perceptual training procedure. As the head does not have a perfect spherical shape and as the head sizes and shapes varied among the subjects, the experimenter arranged the position of the actuators on the head of the participants at each round, making sure they match their corresponding angular location.

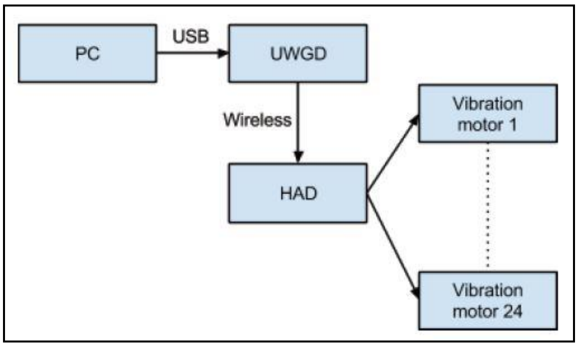


Figure 2: The structure of the haptic device

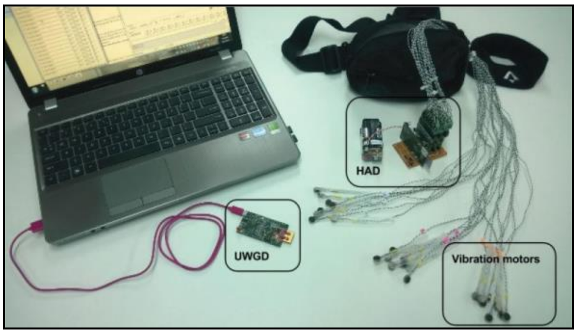


Figure 3: The physical components of the haptic device

2.4. The pre-test session

In the pre-test session, we tested the sound localization performance and navigational skills of our 9 visually impaired subjects in a virtual auditory application called Binaural Navigation Test (Figure 4). In the Binaural Navigation Test, the subjects were required to identify the position of the 3D sound source by freely navigating from the start position (which was randomly generated on the margins of a circle of 150 pixels radius) to the target sound source (which was situated exactly in the center of the circle). The main interaction modality was the mouse (for 6 of the subjects), while the other 3 used the touchpad. The audible area was a circle with 200 pixels radius around the position of the listener. The 3D sounds have been delivered to the listener through a pair of stereophonic headphones. The sonification technique used the inverse proportional encoding of distance, so that the intensity of sound increased as the user got nearer to the source and decreased as he got farther from it (reaching 0 outside the auditory area of 200 pixels). The formula for calculating the perceived sound intensity for the current distance d between the position of the listener to the sound source is the following:

$$SI = \begin{cases} 0, \mathrm{d} > dmax \\ SI_{MIN} + (SI_{MAX} - SI_{MIN}) * (1 - \frac{d}{dmax})^2, d \leq dmax \end{cases} \tag{2}$$

Where SI is the current perceived sound intensity, d is the current distance, dmax=150 pixels, SIMIN=0.05 (the minimum perceivable sound intensity), SIMAX=1 (the maximum sound intensity). The pre-test session consisted of two blocks of trials. Each block of trials had 20 rounds (the rounds numbered from 1 to 5 and from 11 to 15 used the white-pink noise 3D sound combination, while the rounds numbered from 6 to 10 and from 16 to 20 employed the “ding” signal as the main auditory stimulus).

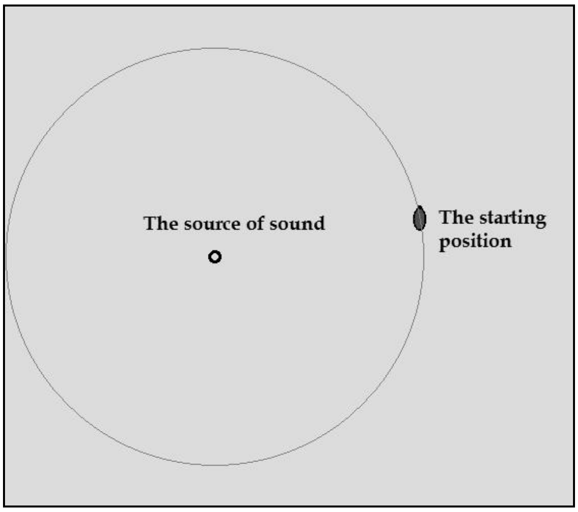


Figure 4: The Binaural Navigation Test

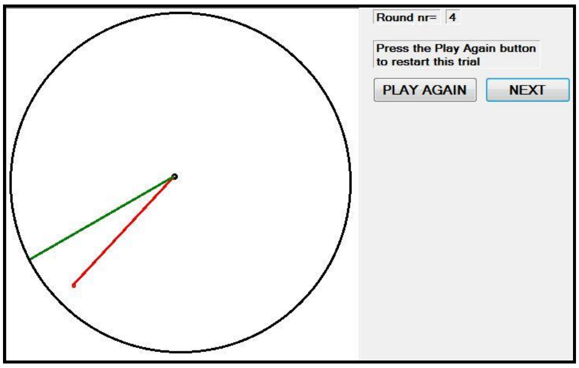
2.5. The training session

The training session offers haptic and auditory feedback based training regarding the correct direction of the sound source in space. The purpose of the training session is to help the subjects to get used to the perception of 3D sounds and to learn how to associate the sound with the corresponding direction from where it is originating. In order to reduce the errors caused by body or head motions, both during the tests and the training session, the subjects were asked to keep their posture in a fixed position. The training session is composed of the following two modules:

a). A free listening module, where the experimenter moved the mouse cursor inside a circle and the listener was able to hear the sound corresponding to the direction between the center of the circle (the virtual location of the listener) and the variable position of the mouse cursor, using as auditory stimuli both the white-pink noise combination and the “ding” sound;

b). A sound localization procedure (Figure 5), where the subjects were required to listen to different sound stimuli (corresponding to the directions 0, 30, 60, 90, 120, 150, 180, 210, 240, 270, 300 and 330 degrees) and to indicate the perceived sound direction using the conventional hour hand of the clock (for example, 12 o’clock for 0 degrees to the front, 3 o’clock for 90 degrees to the right or 6 o’clock for 180 degrees to the back). At each trial, the 3D sounds have been presented continuously, in order to offer the listener a complete perception of the acoustic space. Consequently to the response, the subjects received perceptual feedback about the correct direction of the sound through a series of vibrations produced by the haptic belt that they wore on the head. The subjects who had some degree of residual vision also received visual feedback (Figure 6) - the correct direction of the sound source was presented graphically on the screen (colored in green), together with the listener’s choice (colored in red) and acoustic feedback, in order to make an effective association between the 3D sound and the angular direction it is originating from. Each block of trials consisted of 12 rounds in which the listener was required to listen to the white-pink noise combination and to indicate the perceived direction of the sound, followed by another 12 rounds that employed the “ding” sound as main test stimulus.

Figure 5: Visually impaired subject during the training session



Round nr= 4

Press the Play Again button to restart this trial

Figure 6: Visual feedback about the correct response

2.6. The experimental procedure

In the pre-test and in the post-post sessions, each of the 9 visually impaired subjects ran two blocks of trials (each block having 20 rounds). The training session consisted of 4 blocks of trials (each block having 24 rounds), that were separated into 2 days (2 blocks in the first day and another 2 blocks in the second day). In total, each subject performed 80 sound localization trials during the pre-test and the post-test session and 96 auditory discrimination tasks (accompanied by haptic perceptual feedback) in the training session (Figure 7).

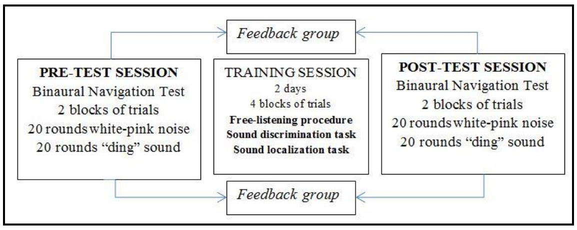


Figure 7: The experimental procedure

3. RESULTS

3.1. Results of the pre-test and of the post-test sessions

During the pre-test and the post-test session, we studied and evaluated the evolution of the following parameters:

P1: the ratio of the distance travelled by the listener (from the starting position that is randomly generated on the margin of the circle to the position of the sound source) to the minimum possible distance of 150 pixels (the radius of the circle).

P2: the percentage of good movements towards the sound source (movements that minimize the distance between the position of the user and the location of the sound source).

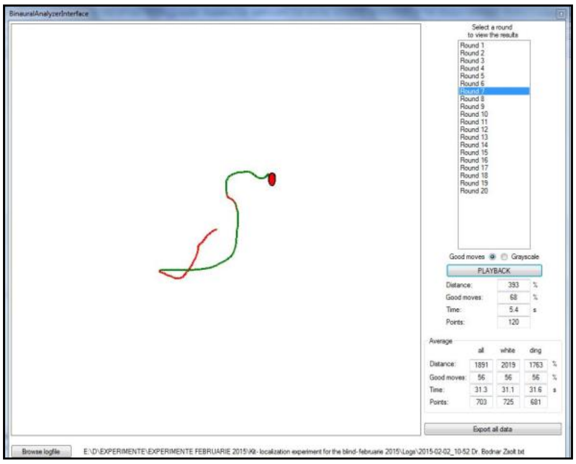
$$P2=\frac{Number\ of\ movements\ towards\ the\ sound\ source}{Total\ number\ of\ movements\ in\ a\ round}*100 \quad (3)$$

P1 and P2 give an insight into the navigational skills of the subjects, enabling us to evaluate their spatial auditory perception. These two parameters offer valuable information concerning the sound localization accuracy and decision-making abilities of the listeners.

P3: the number of mouse movements effectuated by the user.

P4: the round completion time (in seconds).

In order to assess the final results, we designed an application that provides statistical, visualization and audio playback functionality in what concerns the performance of the subjects. Thus, the mean results for all the four studied parameters are displayed, both for the 10 rounds that used white/pink noise, respectively the "ding" sound, as well as for the total number of 20 trials. The users' virtual displacements that are considered as good movements (minimizing the distance between the current position of the listener and the target sound source) are displayed as green-colored segments, while the incorrect movements (effectuated in the opposite direction, maximizing the distance to the target) are colored in red. At the same time, the experimenter can visualize a real-time playback of the subjects' performance for any of the 20 rounds of the test.



Select a round to view the results: Round 7

Good moves ● ○ Grayscale

Distance: 393 %
Good moves: 68 %
Time: 5.4 s
Points: 120

Average

	all	white	ding	
Distance:	1891	2019	1763	%
Good moves:	56	56	56	%
Time:	31.3	31.1	31.6	s
Points:	703	725	681	

E:\D\EXPERIMENTE\EXPERIMENTE FEBRUARIE 2015\Kit - localization experiment for the blind-februarie 2015\Logs\2015-02-02_10-52 Dr. Bodnar Zsolt.txt

Figure 8: Performance of a subject in the 7th round of the pre-test session

Table 1 presents some statistical analysis (mean, standard deviation – SD, minimum and maximum value), for all the researched parameters, for both types of sound stimuli, in the pre-test and post-test sessions of the experiment.

	WHITE-PINK NOISE							
	Pre-test session				**Post-test session**			
	P1	P2 (%)	P3	P4 (s)	P1	P2 (%)	P3	P4 (s)
Mean	10	64	413	26	6.8	66	339	22
SD	5.5	3.6	147	15	4	4	100	11
Min	4.4	57	229	11	2.9	56	175	9
Max	21	70	707	53	16	72	484	40
	"DING" SOUND							
	Pre-test session				**Post-test session**			
	P1	P2 (%)	P3	P4 (s)	P1	P2 (%)	P3	P4 (s)
Mean	12	61	525	32	7.5	65	379	23
SD	6.4	4	219	19	3.8	6	125	11
Min	4	4	219	12	2.6	6	125	9
Max	26	69	982	66	16	77	522	46

Table 1: Descriptive statistics of the researched parameters in the pre-test and in the post-test sessions of the experiment

The results show that the best sound localization performance has been obtained when the combination of white and pink noise (in varying proportions, according to the direction of the sound source in space) has been employed, for both the pre-test and the post-test sessions of the experiment. For the rounds that used the white-pink noise combination as the primary acoustic stimulus, there is a significant improvement in the results obtained in the post-test session for parameters P1 (in a Student t-test, $t=-3.64$, at $p\leq 0.05$) and P2 ($t=2.46$, at $p\leq 0.05$) (Table 2). The mean rate of parameter P1 increased with 32%, while the percent of correct travel decisions (P2) improved between the pre-test and the post-test sessions with 3.2%. We also recorded an improvement for parameters P3 and P4 of 17%, that is statistically significant at $p\leq 0.1$ (for P3, $t=-1.74$ and for P4, $t=-1.02$). For the rounds using the "ding" sound as the primary auditory cue, we obtained significant improvement for all the four studied parameters. Thus, the mean rate of parameter P1 improved with 41% (in an ANOVA test, we obtained $p=0.04$, which is lower than the critical p value of 0.05. In a consequent Student t-test in which we applied the Bonferroni correction in order to reduce the type I errors, we obtained $t=-4.44$, $p=0.002$, $p\leq 0.025$), for P2 we recorded an increase of 6.2% in the percent of correct travel decisions ($t=3.89$, $p\leq 0.05$), the number of mouse movements (P3) reduced with 27% ($t=-2.63$, $p\leq 0.05$), while the round completion time (P4) significantly decreased with 28% ($t=-2.1$, $p\leq 0.1$) between the pre-test and the post-test sessions of the experiment. Moreover, by using the Pearson correlation coefficient, we observed a strong negative correlation between the values of P1 and P2 ($R=-0.83$), demonstrating that a higher distance travelled by the listener from the randomly generated starting position to the location of the sound source will lead to a lower rate of correct travel decisions and to a less accurate perception of the acoustic space. The correlation between P2 and P3, respectively P4 is weak, showing that the round completion time or the number of mouse movements do not influence the sound localization accuracy or the navigational skills of the visually impaired users under altered hearing conditions (such as the use of 3D sounds synthesized from non-individualized HRTFs).

Furthermore, at least 78% of the subjects succeeded to improve their sound localization accuracy as a result of the haptic-auditory perceptual feedback based training (Table 3). The number of participants who obtained higher results in the post-test session is larger for the rounds that used the "ding" sound as the main auditory cue. Thus, all the participants recorded significant improvements for parameters P1 and P2 and 89% (8 of the 9 subjects) achieved better results for parameters P3 and P4. Parameter P2 (the percent of correct travel decisions towards the sound source) is the most significant sound localization criterion, as it offers a clear clue regarding the spatial acoustic resolution of our subjects. 7 of 9 subjects recorded an improvement in the white-pink noise combination rounds, while all the 9 participants succeeded to enhance their sound localization skills when the "ding" sound stimulus was used as the main auditory cue.

Type of sound	**P1**	**P2**	**P3**	**P4**
	Percentage of improvement after the second session			
White-pink noise	32%	3.2%	17.8%	17.1%
"Ding" sound	42%	6.3%	27.7%	28.2%

Table 2: Percentage of improvement after the second session

Type of sound	**P1**	**P2**	**P3**	**P4**
	Percentage of subjects with improvements in the second session			
White-pink noise	100%	78%	78%	89%
"Ding" sound	100%	100%	89%	89%

Table 3: Percentage of subjects with improvements in the second session

The results show clear improvements in the sound localization performance and navigational skills of our subjects between the pre-test and the post-test sessions of the experiment. These improvements are the result of the haptic-auditory perceptual feedback based training, demonstrating that the learning and adaptation play an important role in enhancing the spatial auditory resolution of the visually impaired people. In addition to this, the training procedure was highly efficient in helping the subjects to navigate in the virtual auditory environment under altered hearing conditions and to search for the target sound source while listening to 3D sounds synthesized with non-individualized HRTFs. Even if the "ding" sound offers initial poor localization accuracy (due to its narrowband spectral range and the disturbing short breaks between two consecutive stimuli), it presents a high potential for spatial localization improvement that can be achieved through perceptual learning.

3.2. Results of the training session

The training session took place in 2 different days (separated by a time interval of 4 days). Table 4 presents a statistical descriptive analysis (mean, SD, minimum and maximum value) of the most relevant parameters of our study – the percent of front-back confusions and the angular localization error (the unsigned difference between the direction indicated by the listener and the correct direction of the sound source).

	WHITE/PINK NOISE			
	Day 1		**Day 2**	
	FBC (%)	Err (degrees)	FBC (%)	Err (degrees)
Mean	12	37.3	6	27.7
SD	5.2	12.7	5.1	14.3
Min	4.1	23.5	0	11
Max	16.6	57	12.5	53.5
	"DING" SOUND			
	Day 1		**Day 2**	
	FBC (%)	Err (degrees)	FBC (%)	Err (degrees)
Mean	14.3	44.8	12.5	42.4
SD	7.8	9.9	10.2	15.4
Min	4.1	32	0	23
Max	25	61	29.1	66

Table 4: Descriptive statistics of the researched parameters in the training session

The results show that the best directional sound discrimination performance has been achieved when listening to the white-pink noise combination, in both days of the training session. On the other hand, the "ding" sound offered a less accurate spatial perception that is reflected in the high incidence of front-back confusions and precision localization errors. There has been recorded a significant improvement for the two studied parameters between the two days of the training session, for both types of sounds. Thus, in the case of the white-pink noise, the percent of front-back confusions reduced with 50% (from 12% in the first day of training, to 6% in the second day of training. In an ANOVA test, we obtained p=0.02, which is lower than the critical p value of 0.05. In a consequent Student t-test in which we applied the Bonferroni correction in order to reduce the type I errors, we obtained t=-3.83, p=0.005, p≤0.025), while the angular localization error decreased from 37 degrees to 27 degrees (t=-2.57, p≤0.05). In the case of the "ding" sound, the improvements are not statistically significant (the front-back confusion rate reduced just with 2%, from 14% to 12%, while the sound localization error improved with 2 degrees, from 44 to 42 degrees). In addition to this, the percentage of subjects who recorded improvements in the second day of training was higher under the white-pink noise listening conditions (78% of the total number of subjects succeeded to obtain improvements for both parameters when listening to the combination of white and pink noise in varying proportions, while only 33% of the participants were able to enhance their sound localization accuracy and front-back disambiguation skills when listening to the "ding" sound) (Table 5).

Type of sound	**Front-back confusions**	**Localization error**	**Both**
	Percentage of subjects with improvements in the second day of training		
White-pink noise	100%	78%	78%
"Ding"	55%	55%	33%

Table 5: Percentage of subjects with improvements in the second day of training

The lowest front-back confusion rate has been achieved for the sound targets situated in the second and in the fourth quadrants (Table 6, Figure 9), while the highest percent of front-back confusion errors has been recorded for the sound sources located in the first quadrant (when listening to the white-pink noise sound stimuli) and in the third quadrant (when using the "ding" sound as the main auditory cue). In what concerns the precision error, the highest errors have been recorded in the first quadrant (for the directions 0, 30 and 60 degrees, for both types of stimuli), while the best localization performance has been obtained in the third quadrant (for the white-pink noise stimuli) and in the second quadrant (for the "ding" sound).

For the white/pink noise combination, the largest rate of front-back confusions has been obtained at 0 degrees to the front (15%), while the lowest percent of reversal errors (0%) occurred at 150 degrees (in the second quadrant) and at 210 degrees (in the third quadrant). The lowest angular localization accuracy has been obtained at 0 degrees to the front, while the best performance was recorded at 180 degrees in the back (15 degrees).

For the "ding" sound, the largest rate of front-back confusions has been recorded also at 0 degrees to the front and at 180 degrees in the rear (16%), while the lowest percent of reversal errors (2%) occurred at 60 degrees (in the first quadrant). The lowest angular localization accuracy has been obtained at 30 degrees to the front, while the best performance (angular error of 30 degrees) was recorded at 60 degrees (in the first quadrant) and at 300 degrees (in the fourth quadrant).

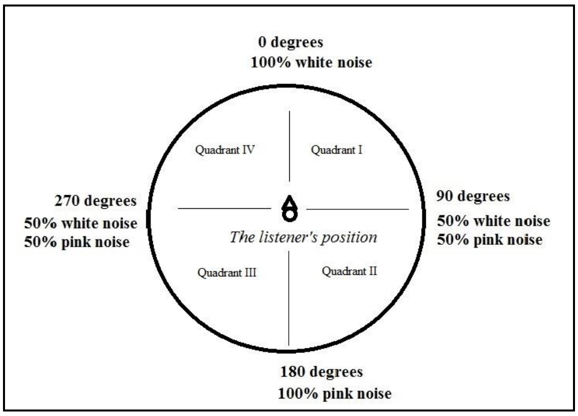


Figure 9: The azimuthal range for each quadrant of the auditory space

ANGLE	FBC W/P (%)	ERR W/P (degrees)	FBC DING (%)	ERR DING (degrees)
Quadrant I (from 0° in the front to 89° to the right)	18.7	55	10	45.6
Quadrant II (from 90° to the right to 179° in the back)	2	28.6	6.9	37.3
Quadrant III (from 180° in the back to 269° to the left)	9.3	25.3	13.1	41.3
Quadrant IV (from 270° to the left to 359° in the front)	3.1	32	3.1	40.6

Table 6: Distribution of front-back confusion errors and angular localization errors in the four quadrants of the 3D auditory space

3.3. Usability study

After the post-test session, the subjects were required to complete a usability questionnaire in which they expressed their opinion concerning the sound localization experiment. 78% of the subjects considered that the sonification technique based on the combination of white and pink noise helped them to avoid front-back confusions and to identify the sound sources originating from the frontal and from the rear plane. Also, 78% of the visually impaired participants expressed their satisfaction concerning the efficiency of the training session. Thus, they considered that the haptic-auditory perceptual feedback based training helped them to improve their sound localization performance and front-back disambiguation abilities. The subjects unanimously appreciated that the sound localization application is motivating and challenging, that it is easy to be learnt and used and that it can be the support for the future development of an assistive device for the visually impaired people. 89% of the subjects considered that the 3D sounds are efficient for localizing the sound source in space and that it was easy to identify the sound source by using the white-pink noise combination. In addition to this, 67% of our users appreciated that it is easy to identify the sound sources using the “ding” sound and that the combination of white and pink noise in varying proportions enabled them to correctly identify the sound sources situated inside the cone of confusion. Nonetheless, only 44% of the subjects considered that the “ding” sound is efficient for avoiding the front-back confusion errors.

4. DISCUSSION

The results of our experiment demonstrate a rapid adjustment of the auditory system of the visually impaired subjects to 3D sounds synthesized from non-individualized HRTFs, due to the crossmodal association between the haptic and the auditory senses. We reported significant improvements in the sound localization accuracy and general spatial acoustic resolution of our subjects, which is reflected in a lower rate of front-back confusions and in a decrease in the angular localization error. The perceptual learning process that took place during the training session enabled the subjects to focus on the spectral and temporal characteristics of the sound and to associate the perceived acoustic stimulus with the corresponding direction in space that was provided through haptic feedback. The adaptation and recalibration process was rapid, demonstrating that also other cognitive mechanism, such as attention and focus have been used during the perceptual training procedure. The improvements are more significant for the white-pink noise combination, due to their wider spectral profile. Nonetheless, we have demonstrated that also other types of sounds (with a narrower spectral profile, such as the “ding” sound) can be localized effectively and present a significant potential for sound localization enhancement as a result of perceptual feedback based training.

In a short interview that took place after the experiment, our subjects expressed their appreciation concerning the proposed application. They also argued that this can be the baseline for developing navigational audio games that would improve their sound localization abilities and that additional training would be highly useful for supporting enhanced orientation and mobility skills. Furthermore, one subject considered that the haptic belt can be used as a standalone device for improving navigation outdoors, by helping the blind people to keep a straight line when walking or by providing directional information.

The object recognition experiment presented in [16], [17] investigates the contribution of dynamic auditory cues (anchor sounds delivered through headphones, using individualized HRTFs) integrated with haptic feedback. The subjects were required to recognize simple virtual objects under two conditions: unimodal, where they were asked to use only a tactile mouse and multimodal, when they were provided both haptic and auditory feedback (continuous bursts of Gaussian white noise that were spatialized in the center of the virtual environment, giving global orientation clues). Just as in our experiment, the distance to the target was rendered through an inverse square law on sound attenuation level. The results of the experiment demonstrate that the bimodal condition led to a more effective exploration strategy. Thus, as in our tests, the use of 3D sound facilitated navigation and object recognition, conveying an experimental testbed for the use of spatialized sounds in virtual environments.

5. CONCLUSIONS

Our study demonstrated that the visually impaired people are able to adapt to altered hearing conditions, such as the use of 3D sounds generated from non-individualized HRTFs in virtual auditory environments. We found evidence that the adaptation process is the result of the haptic-auditory feedback based training session which helped the subjects to associate the auditory cues with the direction of the vibration delivered on the haptic belt placed on the head. Moreover, the visually impaired participants succeeded to improve their sound localization accuracy by reducing the rate of reversal and precision errors for both types of sound (with significant better results for the rounds that employed the synthesis of white and pink noise in varying proportions, according to the direction of the sound source in space). Furthermore, the results of the post-test session prove that our subjects have been able to use binaural 3D sounds as the only means for navigating in a virtual auditory environment. In addition to this, we have proved that they improved their spatial acoustic resolution, sound localization accuracy, orientation and mobility skills and directional decision-making abilities. In conclusion, the proposed approach can be considered a useful training and rehabilitation tool for the future development of audio-only games or for the design of an assistive device for the blind people. Our experiments will continue by refining this one and by providing a more extensive training session, in order to study the highest degree of sound localization accuracy that can be achieved as a result of haptic-auditory feedback based training. Moreover, the results that will be obtained will be integrated into a test procedure for the use of an auditory Sensory Substitution Device which intends to provide a rich representation of the environment.

6. ACKNOWLEDGMENT

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.This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 643636 "Sound of Vision".

7. REFERENCES

[1] T. Letowski, S. Letowski, Localization Error: Accuracy and Precision of Auditory Localization (Chapter 4) in Advances in Sound Localization, book edited by Pawel Strumillo, ISBN 978-953-307-224-1, published: April 11, 2011

[2] A. Meshram, R. Mehra, H. Yang, E. Dunn, J.M. Frahm, D. Manocha, "P-HRTF: Efficient personalized HRTF computation for high-fidelity spatial sound", *International Symposium on Mixed and Augmented Reality,* 2014

[3] M. Dellepiane, N. Pietroni, N. Tsingos, M. Asselot, R. Scopigno, "Reconstructing head models from photographs for individualized 3D-audio processing", *Comput. Graph. Forum*, 27(7), pp. 1719-1727, 2008

[4] Gy. Wersenyi, "Virtual localization by blind persons", *Journal of the Audio Engineering Society (AES)*, vol.60., no.7/8, pp. 579-586, July-August 2012

[5] S. Spagnol, M. Geronazzo, D. Rocchesso, F. Avanzini, "Extraction of Pinna Features for Customized Binaural Audio Delivery on Mobile Devices", *Proceedings of International Conference on Advances in Mobile Computing*, 2013.

[6] D.N. Zotkin, J. Hwang, R. Duraiswami, L.S. Davis, "HRTF Personalization Using Anthropometric Measurements", *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA'03)*, pp. 157-160, 2003.

[7] C. Mendonça, "A review on auditory space adaptations to altered head-related cues", *Front Neurosci.*, vol. 219, no. 8, 2014

[8] N. R. B. Stiles, S.Shimojo, "Sensory Substitution: A New Perceptual Experience", Chapter 43 in The Oxford Handbook of Perceptual Organization, Johan Wagemans, Ed., Oxford University Press, New York, (In Press).

[9] M. Ahissar, "Perceptual training: A tool for both modifying the brain and exploring it", *PNAS*, vol. 98, no.21, 2001

[10] B. Carter, Op Amps for Everyone, Texas Instruments. pp. 10–11. ISBN 0080949487, 2009.

[11] F.X. Diebold, Elements of Forecasting (4th ed.), Thomson, South-Western: Ohio, US, ISBN 978-0-324-35904-6, 2007

[12] A. Downey, Think Complexity, O'Reilly Media. p. 79. ISBN 978-1-4493-1463-7. "Visible light with this power spectrum looks pink, hence the name.", 2012

[13] Á. Csapó, Gy. Wersényi, "Overview of auditory representations in human-machine interfaces", *Journal ACM Computing Surveys (CSUR)*, Vol.46, No.2., Nov. 2013

[14] http://sound.media.mit.edu/resources/KEMAR.html

[15] http://www.csounds.com/downloads/

[16] M. Geronazzo, A. Bedin, L. Brayda, and F. Avanzini, "Multimodal Exploration of Virtual Objects with a Spatialized Anchor Sound," in *Proc. 55th Int. Conf. Audio Eng. Society, Spatial Audio*, Helsinki, Finland, 2014, pp. 1–8.

[17] M. Geronazzo, "Mixed structural models for 3D audio in virtual environments," Ph.D. Thesis, University of Padova, Padova, Italy, 2014.