Improving the Audio Game–Playing Performances of People with Visual Impairments Through Multimodal Training

Oana Bălan, Alin Moldoveanu, Florica Moldoveanu, Hunor Nagy, György Wersényi, and Rúnar Unnþórsson

Structured abstract: *Introduction:* As the number of people with visual impairments (that is, those who are blind or have low vision) is continuously increasing, rehabilitation and engineering researchers have identified the need to design sensory-substitution devices that would offer assistance and guidance to these people for performing navigational tasks. Auditory and haptic cues have been shown to be an effective approach towards creating a rich spatial representation of the environment, so they are considered for inclusion in the development of assistive tools that would enable people with visual impairments to acquire knowledge of the surrounding space in a way close to the visually based perception of sighted individuals. However, achieving efficiency through a sensory substitution device requires extensive training for visually impaired users to learn how to process the artificial auditory cues and convert them into spatial information. *Methods:* Considering all the potential advantages game-based learning can provide, we propose a new method for training sound localization and virtual navigational skills of visually impaired people in a 3D audio game with hierarchical levels of difficulty. The training procedure is focused on a multimodal (auditory and haptic) learning approach in which the subjects have been asked to listen to 3D sounds while simultaneously perceiving a series of vibrations on a haptic headband that corresponds to the direction of the sound source in space. *Results:* The results we obtained in a sound-localization experiment with 10 visually impaired people showed that the proposed training strategy resulted in significant improvements in auditory performance and navigation skills of the subjects, thus ensuring behavioral gains in the spatial perception of the environment.

More than 285 million people worldwide suffer from a certain degree of visual impairment (that is, blindness or low vision), of which 40 million are totally blind, according to statistics from the World Health Organization (2015). Some people with vision loss are affected by the lack of social involvement, a sedentary, home-confined lifestyle, as well as the inability to take care of themselves (Constâncio, 2010). Effective assistive solutions can improve mobility, cardinality, and social integration. A detailed representation of the environment can improve the navigational skills of visually impaired people (Afonso et al., 2010). In compensation for the lack of sight, visually impaired individuals must rely more on alternative sensory modalities such as hearing and touch. This need led to the idea of integrating auditory and haptic (tactile and kinesthetic) stimuli into the development of assistive devices in order to create and deliver a thorough representation of the surrounding space (Jaijongrak, Kumazawa, & Thiemjarus, 2011).

The purpose of this paper is to investigate the improvements that occur in the sound-localization performance and elementary virtual-environment navigational skills of people with visual impairments as a result of training. Through sound-localization training, visually impaired subjects develop improved spatial-auditory performances which are reflected by more accurate representations of environments and enhanced navigational skills. We propose a realistic, immersive, feedback-rich, user-centered, and motivational ludic-based approach to training through virtual games. This approach will enable people with visual impairments to develop sound-localization skills, acquire situational knowledge from audio cues, collect directional information, and develop auditory-based spatial understanding and virtual navigational skills in a dynamic and interactive way.

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

Training sound localization through audio games

Many experiments have demonstrated that audio game–based training can be a reliable strategy for improving spatial-auditory resolution in virtual environments (Bălan, Moldoveanu, & Moldoveanu, 2015; Sanchez, Saenz, Pascual-Leone, & Merabet, 2010; Zahorik, Bangayan, Sundareswaran, Wang, & Tam, 2006).

In the experiment presented by Blum, Katz, and Warusfel (2004), a proprioceptive feedback method was used to help the blindfolded subjects manipulate sound sources located at their hand position (a position that has the advantage of being egocentric for the user). In the training session (designed as a game-like scenario), the subjects were required to search for animal sounds hidden around them. Using a positional hand-tracked ball (the sound being spatialized at the center of the ball), they were required to search the perceived direction of the sound. When the subject identified the correct location, pink noise (a signal with power spectral density inversely proportional to the signal frequency) was used as primary sound stimulus was replaced by a sound that was similar to that produced by an animal, and the next direction was generated having as a reference point the previous sound source location. In this way, the subjects were provided with continuous auditory feedback and could explore their entire sphere of perception and action. The experimental results concluded that the angular error was reduced by 6.1 degrees.

In Honda et al. (2007), the transfer effects of playing an audio game were analyzed in a sound-localization experiment in which two groups of subjects participated: a group that used generic head-related transfer functions (HRTFs), and another group that used individualized HRTFs. The HRTFs are a measure of the sound transformation from the source to the listener's ears (Meshram et al., 2014; Spagnol, Geronazzo, Rocchesso, & Avanzini, 2013) that are highly dependent on the anatomical characteristics of the listener's body (size and shape of the pinna, head, and torso). Computer-generated 3D binaural sounds are made by convolving a monaural signal with the corresponding HRTFs of a particular direction in space. The training method consisted of an audio game entitled Bee Bee Beat, in which the sound of a honeybee appeared at an arbitrary direction in space and the listener was required to use a plastic hammer to hit the noise. After the listener identified the location of the target, immediate vibration feedback was received from the hammer. The trained group was asked to play the game for 30 minutes per day for a period of two weeks. Consequent to training, the accuracy of sound localization of the subjects increased by 20%.

The Audio Doom game (Lessard, Pare, Lepore, & Lassonde, 1998) aimed to enhance the cognitive abilities and virtual navigational skills of visually impaired children. The players needed to explore a virtual environment that consisted of a 3D labyrinth of walls and corridors, to avoid monsters, and to find the exit to the next level—all while listening to auditory cues such as the sound of footsteps or of a door opening. Seven visually impaired children who played the game succeeded in improving their spatial cognitive skills. The transfer of learning into real-world situations was demonstrated by their ability to recreate the route they traveled in the game using Lego blocks (Bălan, Moldoveanu, & Moldoveanu, 2015).

In Pyvox (Gaudy, Natkin, & Archambault, 2009), the virtual environment was represented by a tower with 70 floors, corresponding to the 70 levels of the game. The player was required to find the exit for each floor without hitting the walls. The sonification approach encoded the walls with the sound of an unpleasant noise, while to the exit of each level it assigned a stereo rendering that varied in pitch according to the distance to the current position of the player. An experiment performed with two groups of visually impaired people showed that they understood the principles of the game and succeeded in completing at least three levels of it (Bălan, Moldoveanu, Moldoveanu, & Negoi, 2015).

In BlindSide (Astolfi, n.d.; Parker, 2013; Reinhard, 2014), the virtual environment consisted of the settings of a building and a city, while the sonification technique included 3D binaural and prerecorded sounds, narration sequences, and many realistic effects. In an experiment that included both sighted people and persons with visual impairments, the results showed that the latter were able to play the game more accurately and to finish it earlier than their sighted counterparts

(Bălan, Moldoveanu, Moldoveanu, & Dascălu, 2014).

In Balan et al. (2015), the sound-localization performance of nine subjects with visual impairments was assessed before and after a series of haptic-auditory training procedures aimed at enhancing the perception of 3D sounds. The test results showed that the subjects improved their sound-localization performance and reduced the incidence of angular precision and reversal errors. Our experiment was composed of a pretest session; a training session (based on haptic-auditory perceptual feedback, aimed at helping the subjects get used to the perception of 3D sounds delivered via headphones); and a posttest session (identical to the pretest session), which had the purpose of assessing the degree of acoustic spatial resolution improvement achieved as a result of the perceptual training procedure.

Methods

Overview

The study presented here comprises the assessment of the sound-localization skills of 10 subjects with visual impairments who were required to play an audio game with hierarchical levels of difficulty. In the game, the players had to identify the location of several hidden auditory targets while trying to avoid obstacles. The experiment was composed of a pretest session (in which the subjects were asked to play the game twice, for two different sets of levels); a training session (aimed at helping individuals with visual impairments to adapt to the perception of 3D sounds through multimodal interaction, both auditory and haptic); and a posttest session (similar in structure and difficulty to the pretest), in which the degree of improvement in sound localization following training was evaluated.

Participants

Ten individuals with visual impairments (five women and five men, aged 27 to 63 (mean age, 43), with a percentage of residual vision ranging between 0% and 15%, participated in our experiment. Prior to the experiment, each subject declared his or her level of residual vision, determined by their ophthalmologists. The subjects' visual acuity ranged from 20/200 to 20/1000, with a visual field of 20 degrees or less. Two of the subjects were congenitally blind, one was congenitally visually impaired (with 10% vision), and the others were late-onset visually impaired. The experiments followed the tenets of the World Medical Association Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects. The Blind People Association from Györ, Hungary, approved the experiments and written informed consent was obtained from all subjects.

Sound stimuli

The auditory stimuli were continuous 3D binaural sounds synthesized with nonindividualized HRTFs from the Massachusetts Institute of Technology (MIT) dataset (Gardner & Martin, 2000), auditory icons (sounds that create an analogy with real-world events and situations), and "earcons" (abstract, symbolic sounds used to facilitate the players' navigation through the complex content of the game) (Csapó and Wersényi, 2013). The location of a hidden auditory target was sonified using a combination of white and

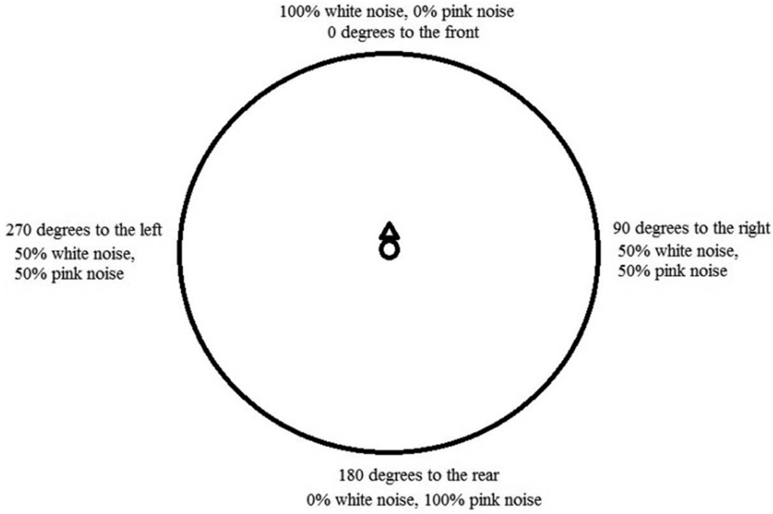

Figure 1. Diagram of the hemifield.

pink noise in varying proportions (both noises could be perceived simultaneously), so that at 0 degrees to the front the listener hears only white noise and at 180 degrees to the back the listener hears only pink noise. In the right hemifield, the proportion of white noise decreased and that of pink noise increased, reaching equal levels at 90 degrees (see Figure 1). On the other hand, in the left hemifield, the percentage of pink noise decreased and that of white noise increased, reaching equal levels at 270 degrees. The white noise presents a constant-power spectral density (equal power per Hertz) (Norton and Karczub, 2003), which means that it contains every frequency within the range of human hearing (generally from 20 hertz to 20 kHz) in equal amounts (Hansen, 2001). The pink noise is a random signal with a power spectral density that is inversely proportional to the frequency of the signal, so that each octave contains an equal amount of noise energy (Hansen, 2001). The formula for calculating the proportion of white and pink noise for a given direction in space in our experiment is (Bălan et al., 2015):

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

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

The purpose of this combination of white and pink noise was to reduce the incidence of front-back confusions by enabling the listener to differentiate the direction of source, based on the spectral profile of the sound.

The location of the obstacles was encoded in an alarm sound that was spatialized by using the nonindividualized HRTFs from the MIT HRTF database, so that the players could identify the direction of both targets and obstacles. The auditory icons used were the alarm sounds aimed at raising the player's awareness of approaching the obstacles and the sound of a crash (or accident) when running into an obstacle. Earcons were represented by the sound of a click, when the player succeeded in identifying the position of a target, and by a bell ringing that announced the end of the game. When the subjects approached targets or obstacles, the perceived sound intensity increased; when they receded from the objects, the intensity decreased.

Procedure

In the pretest session, the sound-localization performance of the subjects with visual impairments was evaluated in an audio game in which the subjects were required to identify the location of several hidden auditory targets while trying to avoid blocking obstacles. The game had 10 levels of difficulty designed in the 2D space, with varying numbers of targets and obstacles:

- Level 1: 1 target and 1 obstacle
- Level 2: 1 target and 2 obstacles
- Level 3: 2 targets and 2 obstacles
- Level 4: 2 targets and 3 obstacles
- Level 5: 3 targets and 3 obstacles
- Level 6: 3 targets and 4 obstacles
- Level 7: 4 targets and 4 obstacles
- Level 8: 4 targets and 5 obstacles
- Level 9: 5 targets and 5 obstacles
- Level 10: 5 targets and 5 obstacles

At each level, the target sound sources were positioned starting from the bottom border of the playing window to the top of it, so that once the player discovered one target, the player was required to look for the next ones above, thus reducing the searching area and further preventing the occurrence of front-back localization misjudgments. Only one sound target was active (that is, could be heard) at a moment of time. Once the position of the current target was identified, it consequently became inactive and the next target (in ascending order, from the bottom side of the playing window to the top of it) became audible. An obstacle became audible when the player was within a range of 150 pixels from the obstacle. When two obstacles were less than 150 pixels away from the player, two sound stimuli with different directional cues, corresponding to each obstacle, were presented.

Figure 2. A subject playing the game during the pretest session.

The players were required to navigate freely, using the mouse or touchpad movement as interaction modality. Prior to the experiment, the subjects were instructed to use the mouse or the touchpad, according to preference or prior experience. Because they were modifying their position in respect to the active target and the surrounding obstacles, they could perceive changes in both the spectral content of the sound, its intensity, and its localization simulated with HRTFs. During the experiment, the subjects with low vision were blindfolded.

During the pretest session, the subjects were required to play the game twice, with two different sets of levels (set 1 and set 2) (see Figure 2). Before the pretest,

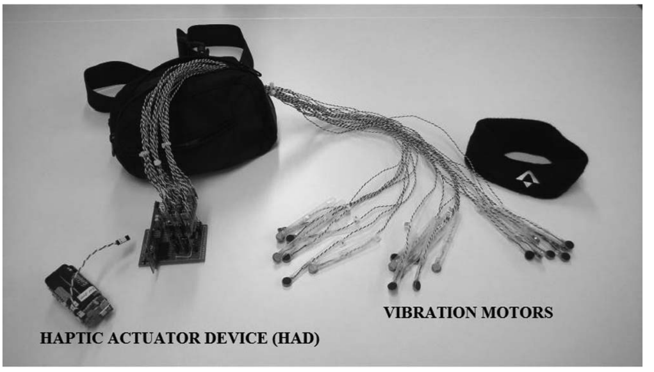

Figure 3. The haptic system.

the subjects were presented with the purpose of the game and the main auditory cues used in the sonification strategy. Moreover, in order to allow them to become familiar with the aim of the game and with the perception of 3D sounds, the subjects were allowed to practice playing the game as long as they considered it was necessary prior to the start of the tests. Usually, most of them played the game once (10 levels), for an average time of 10 minutes, until they were accustomed to the sonification approach and the aim of the game.

The studied parameters were:

- P1: The ratio of the distance traveled by the player (from the starting position until discovery of the location of the current target) to the minimum possible distance (the Euclidean distance between the starting point and the position of the current target). For the first target of any level, the starting position was the center of the bottom border of the playing window. For the other targets, the starting position was the location of the previously identified target.
- P2: The percentage of correct travel decisions, defined as movements effectuated towards the sound source (minimizing the distance between the user's virtual location and the position of the target).
- P3: The average level completion time (in seconds).
- P4: The average number of obstacle hits.

During the training session, the subjects were provided with haptic feedback, which conveyed the direction of the sound source. Thus, the subjects wore stereophonic headphones and a haptic headband that contained 24 vibration motors (evenly distributed around the head) that transmitted vibrations corresponding to the direction of the sound source in space.

The haptic system was composed of several parts (see Figure 3):

- The USB wireless gateway device (UWGD);
- A haptic actuator device (HAD) that effectively controlled the haptic actuators (eccentric rotating mass [ERM] motors); and
- Vibration motors, fixed along a stick in order to allow easy handling. The 24

motors were evenly distributed around the head of the subject. As head size and shape varied significantly among subjects, the experimenter personally checked and ensured accurate placement of the vibration motors at their corresponding directions on the haptic belt for each listener.

The training session took place on two consecutive days. On both, the subjects went through three training periods, each with a duration of three minutes, in which they were required to listen to a series of 24 sounds (emulated in clockwise order from 0 to 345 degrees) and then to randomly generated auditory stimuli for four seconds. The sounds used were a combination of white and pink noise in varying proportions—the same encoding used for the game described in the previous section. Each sound stimulus perceived through the headphones was accompanied by a train of four vibrations on the haptic headband (one vibration per second), corresponding to the direction of the sound in space. The purpose of the training session was to help the subjects get used to the perception of 3D sounds and to create an effective crossmodal association (haptic and auditory) that would help them easily identify the direction of the 3D sounds. The high resolution of the haptic headband (the vibration motors were evenly distributed around the subject's head, resulting in a fifteen-degree angle between the motors) allowed for accurate training, a strategy aiming to offer a more accurate spatial perception of the environment (see Figure 4).

The posttest session, which took place one day after training, was carried out in exactly the same conditions as the pretest

Figure 4. A subject playing a game while wearing the haptic headband.

session, using the audio-based game. The goal of the posttest session was to assess the level of sound-localization and virtual-navigational skills improvement achieved after training.

Software tools

The game presented a functionality that allowed saving the current gameplay records (the records of the player's movements on the computer screen) into a log file that could be uploaded later on for analyzing and visualizing the results.

For designing each level (number and location of targets and obstacles), a tool called Game Editor was developed. It allowed the experimenter to set the difficulty of the levels by manually placing targets and obstacles inside the playing window.

The Game Analyzer tool allowed visualization, synchronized playback of the subjects' game performance, and basic statistical analysis (mean values for all four of the studied parameters) for each level and for each target. The segments that were colored in green (black in the colored version of the paper) represented good movements (getting closer to the

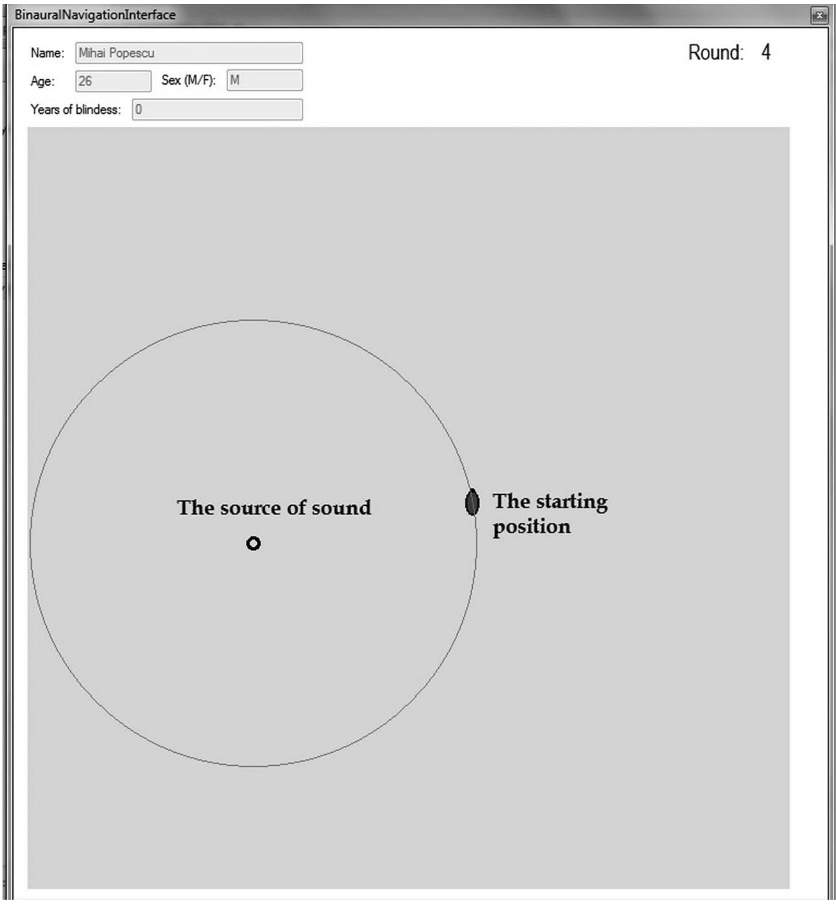

Figure 5. Image of the binaural navigation interface showing the starting position and the source of the sound.

target), and those painted in red (black in the colored version of the paper) represented incorrect travel decisions (getting farther away from the target). The "Export all data" button allowed the results to be exported to an Excel file for further analysis. Overall, the Game Analyzer tool proved to be an excellent analysis solution, which, besides allowing direct observation of the subjects during the experiment, facilitated better understanding of the in-game behavior and the issues that the subjects were confronted with.

Seven of our subjects participated in a sound-localization experiment that took place two months before the current one (Bălan et al., 2015). In the previous experiment, the sound-localization accuracy of nine people with visual impairments was tested before and after a training session based on auditory and haptic feedback. In both the pretest and the posttest sessions, the subjects were required to identify the position of a sound source by freely navigating from the starting position (the center of a circle of 150 pixels in radius) to the actual location of the target (which was randomly generated on the margin of the circle) (see Figure 5). The sound stimuli were a combination of

Table 1
Brief statistics of the results.

Variable	Pretest session P1 (distance ratio)	Pretest session P2 (% correct moves)	Pretest session P3 (seconds)	Pretest session P4 (obstacle hits)	Posttest session P1 (distance ratio)	Posttest session P2 (% correct moves)	Posttest session P3 (seconds)	Posttest session P4 (obstacle hits)
Mean	4.2	70.6	41.6	0.4	3.0	77.4	27.7	0.2
SD	1.8	6.6	10.4	0.3	1.5	8.1	17.5	0.2
Min	1.8	58.5	23.5	0.1	1.5	66.2	10.9	0.0
Max	8.6	83.1	61.9	1.3	6.4	90.0	60.7	0.5

white and pink noise in varying proportions, according to the direction of the sound source in space (previously described in this paper) and a "ding" type signal with a narrower spectral profile. In the training session, the subjects received auditory and haptic feedback about the direction of the sound source in space through the haptic belt they were required to wear on their heads. In that case, the vibration motors were placed at a 30-degree difference around the head, offering a lower spatial resolution than in the current experiment. The results obtained demonstrated a more accurate sound-localization performance and enhanced localization skills in the posttest session of the experiment, notably for the trials where the white-pink noise combination was employed as the primary auditory cue.

Results

Table 1 briefly presents a statistical overview of the results obtained by the subjects (mean, *SD* [standard deviation], minimum and maximum value) in both the pretest and posttest sessions of the experiment for parameters P1 through P4 (see Figures 6, 7, 8, and 9). Table 2 presents the raw results of the experiment.

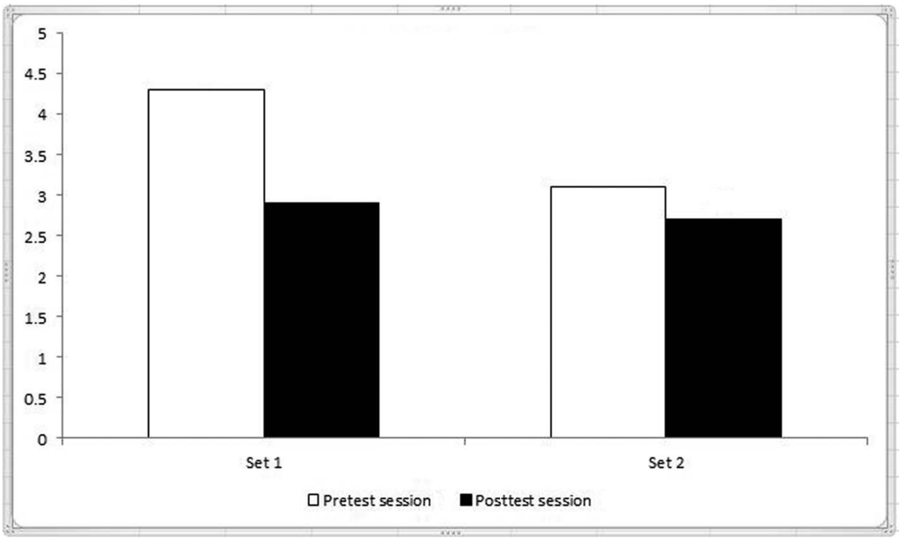

Figure 6. Evolution of P1.

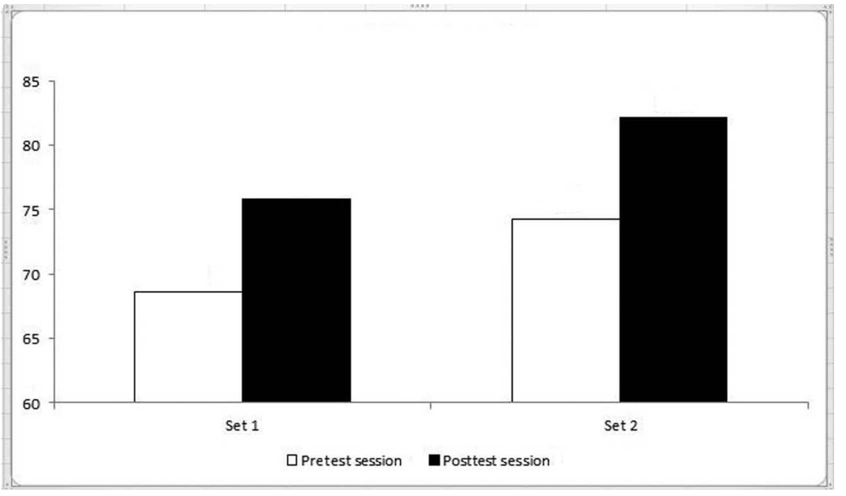

Figure 7. Evolution of P2 (%).

More than 80% of the subjects succeeded in enhancing their sound-localization and navigational performance in the posttest session of the experiment (see Table 3). All the participants recorded a higher rate of correct travel decisions toward the target sound source.

Table 4 presents the mean results for all four parameters, for both sets of levels, in the pretest and posttest sessions of the experiment.

Regarding the results obtained for the second set of levels, the results obtained in both the pretest and posttest sessions are higher than those recorded for the first set. Thus, the mean value of parameter P1 decreased by 14% (from 3.1 to 2.7, although the results are not statistically significant). The mean rate of correct travel decisions increased by 7.7% (from 74.3% to 82.1%); the differences between the performance in the pretest and post-

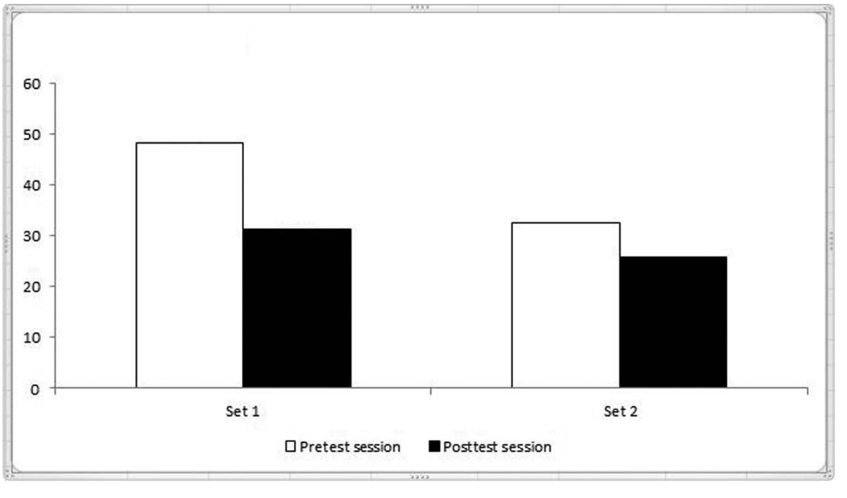

Figure 8. Evolution of P3 (seconds).

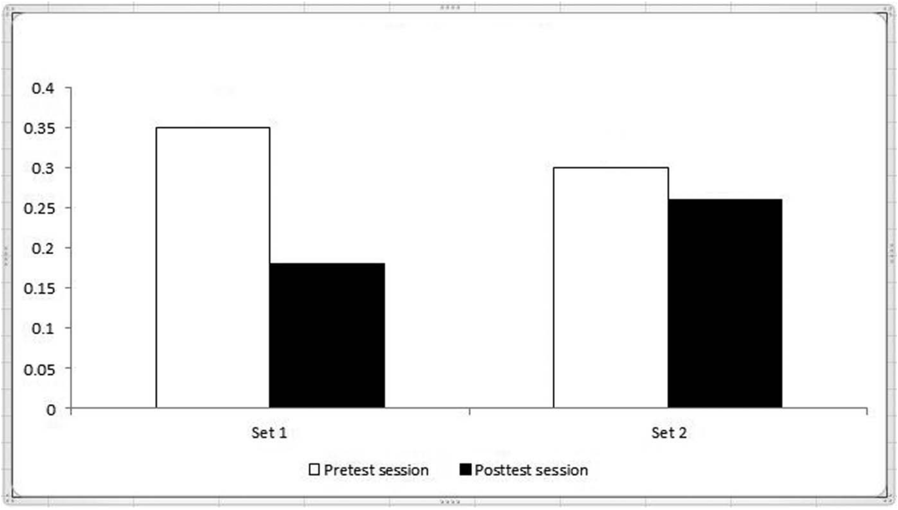

Figure 9. Evolution of P4.

test sessions are statistically significant in an ANOVA test at $p \leq 0.05$ and in a student *t*-test where $t = 3.4$ at $p \leq 0.05$). The level completion time was reduced by 20.4% (from 32.5 seconds to 25.9 seconds in the posttest session; the differences are not statistically significant). And the mean rate of obstacle hits per level decreased by 14.2% (from 0.30 to 0.26; the results are not statistically significant) (see Table 5).

The results presented in Table 6 show that the mean rate of parameters P1 and P2 is higher in the current experiment than in the one performed two months prior (Bălan et al., 2015), demonstrating the persistence of the spatial-auditory adaptation of the human hearing system and its continuous improvement in time.

The mean rate of parameter P1 improved by 43.1% ($t = 1.68$, $p \leq 0.1$)

Table 2
The raw results of the experiment.

	Pretest session				Posttest session			
Subject	P1 (distance ratio)	P2 (% correct moves)	P3 (seconds)	P4 (obstacle hits)	P1 (distance ratio)	P2 (% correct moves)	P3 (seconds)	P4 (obstacle hits)
Subject 1	1.8	83	23.5	0.1	1.7	84.4	20.8	0.1
Subject 2	3.7	70.1	41.2	0.4	2.5	78.3	19.6	0.2
Subject 3	4.0	69.6	39.2	0.4	1.9	83.6	12.6	0.08
Subject 4	5.5	66.3	46.7	0.1	6.4	66.8	57.6	0.4
Subject 5	3.9	70.3	35.9	0.3	1.8	84.4	17.4	0.02
Subject 6	4.4	66.8	61.9	0.4	3.8	72.3	60.7	0.1
Subject 7	2.9	76.1	36.6	0.1	1.5	89.9	10.9	0
Subject 8	3.5	69.4	37.5	0.4	3	71.9	28.9	0.4
Subject 9	8.6	58.5	53.1	1.3	4.8	66.2	27.3	0.1
Subject 10	3.4	75.5	39.6	0.1	3	75.9	20.8	0.3

Table 3
Improvements recorded in the posttest session.

Variable	P1 (% distance ratio)	P2 (% correct moves)	P3 (% seconds)	P4 (% obstacle hits)
Average improvement in the posttest session	27	6.81	33.3	50.3
Percentage of subjects with improvements in the posttest session	90	100	90	80

and that of parameter P2 by 5.64% (the results are statistically in an ANOVA test at $p \leq 0.1$ and in a student t-test where $t = 2.08$ at $p \leq 0.1$) between the posttest session of the previous experiment and the pretest session of the current one. Also, 85% (six out of seven of the subjects) recorded more accurate sound-localization abilities in the pretest session of the current experiment than in the posttest phase of the previous one, a fact that can be explained by their familiarity with the tasks and with the sound stimuli—the perception of directionality conveyed by the 3D sounds, the spectral characteristics of the white and pink noise, and the continuous change in sound intensity that is inversely proportional to the distance between the listener and the sound source.

The final results demonstrate significant improvements in the posttest session of the present experiment, surpassing those obtained in the previous experiment by 57.4% (the results are statistically significant in an ANOVA test at $p \leq 0.1$) for P1 (from 6.68 to 2.84) and with 13.8% for P2 (the differences are statistically significant in an ANOVA test at $p \leq 0.05$ and in a student t-test for dependent means in which $t = 3.5$ at $p \leq 0.5$) (from 66.15% to 80%).

Moreover, the players who participated in the previous experiment were more proficient in accomplishing the virtual navigational and sound-localization tasks required in the current game than were their inexperienced counterparts. At the same time, the experienced users recorded a higher improvement in the posttest session of the current experiment (compared to the pretest session), increasing their percentage of correct travel decisions by 8.3% (from 71.7% to 80%; the results are statistically significant in an ANOVA test at $p \leq 0.05$ and in a student t-test where $t = 3.66$ at $p \leq 0.05$).

Discussion

The results show that by using the proposed multimodal (auditory and haptic) training strategy, people with visual im-

Table 4
Results for the first and the second set of levels.

Set	Pretest session				Posttest session			
	P1	P2 (%)	P3 (seconds)	P4	P1	P2 (%)	P3 (seconds)	P4
Set 1	4.3	68.6	48.1	0.35	2.9	75.8	31.3	0.18
Set 2	3.1	74.3	32.5	0.30	2.7	82.1	25.9	0.26

Table 5
Improvements in the posttest session of the experiment for both sets of levels.

Improvement in the posttest session	P1 (%)	P2 (%)	P3 (%)	P4 (%)
Set 1	30.9	7.2	34.9	49.2
Set 2	14	7.7	20.4	14.2

pairments succeeded in achieving a rapid improvement of their sound-localization abilities and auditory-based virtual navigational skills. Extensive gameplay in the pretest and posttest sessions of the experiment provided an important contribution to the improvement of the accuracy of sound localization and virtual navigational skills, as it enabled the subjects to gain experience with the game.

The training procedure helped the subjects adapt to stimulus conditions that presented a mismatch between the spectral cues and the sound direction, such as the use of 3D binaural sounds synthesized with nonindividualized HRTFs in virtual auditory environments. Moreover, the perceptual training enabled the subjects to map the virtual settings and to perform simple navigational tasks (target localization, obstacle avoidance, and effective gameplay).

The subjects based their game-playing strategy on the perception of both the directional binaural sounds which gave clear clues about the location of both targets and obstacles in space, and on the perception of continuous changes in sound intensity. The improvement in sound localization accuracy is explained by the perceptual training method that used broadband noises containing more spectral cues for the learning and retrieving process. Moreover, even if these stimuli are not natural, they are effective for training because of their enhanced externalization features (Mendonça, Campos, Dias, & Santos, 2013). Another argument that supports the efficiency of the training session is that the subjects were only trained using 24 virtual sound source positions, whereas error reductions and improved spatial perception have been recorded for many other stimuli directions, including the untrained positions. One of the most remarkable results reported in this study is the long-lasting effects of the training sessions performed in the previous experiment. However, the results of the pretest session of the current experiment are more accurate than those recorded in the posttest phase of the previous one, demonstrating that the spatial-auditory remapping is a continuous process and that a new, solid and persistent head model was developed for localizing altered sound cues, in accordance with Hofman's theory (Hofman, van Riswick, & van Opstal, 1998; Mendonça et al., 2013).

In summary, the ratio of the total distance traveled by the listeners to the minimum possible distance decreased by 27%, the rate of correct travel decisions towards identifying the location of the target sound sources improved by approximately 7%, the mean level com-

Table 6
Comparative analysis between the results of the posttest session of the previous experiment and the results of the current one.

Session	P1	P2 (%)
Posttest session of the previous experiment	6.6	66.1
Pretest session of the current experiment	3.8	71.7
Posttest session of the current experiment	2.84	80

pletion time was reduced by 33%, and the number of obstacle hits per level decreased by half in the posttest session of the experiment.

Since the results obtained in both the pretest and posttest sessions of the current experiment are higher than those achieved in the previous one, we shed some light on the fact that the human auditory system is able to continuously improve its sound-localization abilities and that experience-driven learning plays a fundamental role in enhancing the navigational and spatial cognitive skills of individuals with visual impairments.

Conclusions

The experimental results of this research demonstrated that visually impaired people are able to perform route-navigational tasks (such as searching for auditory targets or avoiding obstacles) in virtual reality environments using 3D binaural sounds as the only means for navigation. Moreover, the brief multimodal (auditory and haptic) training session helped the subjects adapt to altered hearing conditions (such as the use of 3D sounds filtered with nonindividualized HRTFs) and contributed to creating an association between the auditory stimuli perceived in the headphones and the vibrations corresponding to the direction of the sound source on the haptic belt.

One possible limitation to the study presented here is the prior experience of the seven subjects who participated in a similar sound-localization experiment that took place two months prior to the one depicted in this article. Thus, they were familiar with the procedure, the haptic device, and the perception of 3D binaural sounds. Another limitation was the small sample size of subjects that led to statistically insignificant results for some parameters of the study.

This study has applicability for the Sound of Vision project (Sound of Vision, n.d.), which is a European research project that intends to develop an assistive device specifically for visually impaired people. The system will be designed to encode the environmental information into auditory and haptic stimuli. It will not use traditional supra-aural headphones, but custom-designed bone conduction or multispeakers, in order to prevent the user from losing the ability to hear environmental information. The results of the current study will be used for the development of a training strategy based on multimodal perceptual feedback and virtual reality that will enable the users of the Sound of Vision device to become familiar with the system before using it in the real-world environment.

Furthermore, as audio games are not restricted to only individuals in the visually impaired community, this game can be played by sighted people who want to try an alternative to traditional video games for entertainment purposes and to train their sound-localization skills at the same time.

References

Afonso, A., Blum, A., Katz, B. F. G., Tarroux, P., Borst, G., & Denis, M. (2010). Structural properties of spatial representations in blind people: Scanning images constructed from haptic exploration or from locomotion in a 3-D audio virtual environment. *Memory & Cognition*, *38*, 591–604.

Astolfi, M. T. (n.d.) *BlindSide*. Retrieved from http://michaeltastolfi.com/#/blindside

Bălan, O., Moldoveanu, A., & Moldoveanu,

F. (2015). Navigational audio games: An effective approach towards improving spatial contextual learning for blind people. *International Journal on Disability and Human Development*, *14*(2), 109–118.

Bălan, O., Moldoveanu, A., Moldoveanu, F., & Dascălu, M. I. (2014). Navigational 3D audio-based game-training towards rich auditory spatial representation of the environment. *Proceedings of the 18th International Conference on System Theory, Control and Computing*, *Sinaia, Romania,* 688–693.

Bălan, O., Moldoveanu, A., Moldoveanu, F., & Negoi, I. (2015). The role of perceptual feedback training on sound localization accuracy in audio experiments. *Proceedings of the 11th International Scientific Conference: eLearning and Software for Education*, *Bucharest, Romania,* 502–509.

Bălan, O., Moldoveanu, A., Nagy, H., Wersényi, Gy., Botezatu, N., Stan, A., & Lupu, R. G. (2015, July). *Haptic-auditory perceptual feedback-based training for improving the spatial acoustic resolution of the people with visual impairment.* Paper presented at the 21st International Conference on Auditory Display (ICAD-2015), Graz, Austria.

Blum, A., Katz, B. F. G., & Warusfel, O. (2004). Eliciting adaptation to non-individual HRTF spectral cues with multi-modal training. *Proceedings of Joint Meeting of the German and the French Acoustical Societies (CFA/DAGA '04), Strasboug, France,* 1225–1226.

Constâncio, J. L .L. (2010). The visually impaired: involvement in physical activity and sport. *Revista Digital–Buenos Aires*, *146*. Retrieved from http://www.efdeportes.com/efd146/the-visually-impaired-in-sport.htm

Csapó, Á., & Wersényi, Gy. (2013). Overview of auditory representations in human-machine interfaces, *Journal ACM Computing Surveys (CSUR)*, *46*(2).

Gardner, B., & Martin, K. (2000). *HRTF Measurements of a KEMAR Dummy-Head Microphone.* Retrieved from http://sound.media.mit.edu/resources/KEMAR.html

Gaudy, T., Natkin, S., & Archambault, D. (2009). Pyvox 2: An audio game accessible to visually impaired people playable without visual or verbal instructions. *Transaction on Edutainment II*, *LNCS 5660*, 176–186.

Hansen, C. H. (2001). *Fundamentals of acoustics* (Unpublished lecture notes). Department of Mechanical Engineering, University of Adelaide, Australia.

Hofman, P. M., van Riswick, J. G. A., & van Opstal, A. J. (1998). Relearning sound localization with new ears. *Nature Neuroscience*, *1*, 417–421.

Honda, A., Shibata, H., Gyoba, J., Saitou, K., Iwaya, Y., & Suzuki, Y. (2007). Transfer effects on sound localization performances from playing a virtual three-dimensional auditory game. *Applied Acoustics*, *68(8)*, 885–896.

Jaijongrak, V., Kumazawa, I., & Thiemjarus, S. (2011). A haptic and auditory assistive user interface: Helping the blind on their computer operations. *The 12th IEEE International Conference on Rehabilitation Robotics*, *Zurich, Switzerland,* 154–159,

Lessard, N., Pare, M., Lepore, F., & Lassonde, M. (1998). Early-blind human subjects localize sound sources better than sighted subjects. *Nature*, *395*, 278–280.

Mendonça, C., Campos, G., Dias, P., & Santos, J. A. (2013). Learning auditory space: Generalization and long-term effects. *PLOS ONE, 8*(10). 8:e77900. doi: 10.1371/journal.pone.0077900

Meshram, A., Mehra, R., Yang, H., Dunn, E., Frahm, J.M., & Manocha, D. (2014). P-HRTF: Efficient personalized HRTF computation for high-fidelity spatial sound. *Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR), Munich, Germany.*

Norton, M. P., & Karczub, D. G. (2003). *Fundamentals of noise and vibration analysis for engineers*. Cambridge, UK: Cambridge University Press.

Parker, L. (2013). *A video game that you can't even see*. Retrieved from http://www.newyorker.com/tech/elements/a-video-game-that-you-cant-even-see

Reinhard, B. (2014). *A sound solution: History of audio games for the visually impaired*. Retrieved from http://artistryingames.com/sound-solution-history-audio-games-visually-impaired

Sanchez, J., Saenz, M., Pascual-Leone, A., & Merabet, L. (2010). Navigation for the blind through audio-based virtual environments. *Proceedings of the CHI Conference*, *Atlanta, Georgia,* 3409–3414.

Sound of Vision. (n.d.) *Home*. Retrieved from http://www.soundofvision.net

Spagnol, S., Geronazzo, M., Rocchesso, D., & Avanzini, F. (2013). Extraction of pinna features for customized binaural audio delivery on mobile devices. *Proceedings of International Conference on Advances in Mobile Computing & Multimedia*, *Vienna, Austria,* 514.

World Health Organization. (2015). *Statistics*. Retrieved from http://www.who.int/mediacentre/factsheets/fs282/en

Zahorik, P., Bangayan, P., Sundareswaran, V., Wang, K., & Tam, C. (2006). Perceptual recalibration in human sound localization: Learning to remediate front-back reversals. *Journal of the Acoustical Society of America, 120*, 343–359. doi: 10.1121/1.2208429

Oana Bălan, Ph.D., *assistant teacher, University Politehnica of Bucharest, Splaiul Independentei, 313, Bucharest, Romania; e-mail: oana.balan@cs.pub.ro.* ***Alin Moldoveanu, Ph.D.,*** *professor, University Politehnica of Bucharest, Bucharest, Romania; e-mail: alin.moldoveanu@cs.pub.ro.* ***Florica Moldoveanu, Ph.D.,*** *professor, University Politehnica of Bucharest, Bucharest, Romania; e-mail: florica.moldoveanu@cs.pub.ro.* ***György Wersényi, Ph.D.,*** *associate professor, Széchenyi István University, Győr, Egyetem tér 1., 9026 Hungary; e-mail: wersenyi@sze.hu.* ***Hunor Nagy, M.Sc.,*** *Ph.D. student, Széchenyi István University, Egyetem tér 1., Hungary; e-mail: nhunorz@gmail.com.* ***Rúnar Unnþórsson, Ph.D.,*** *associate professor, University of Iceland, School of Engineering and Natural Sciences—Faculty of Industrial Engineering, Mechanical Engineering and Computer Science, VR-2/V02-237, Reykjavik, Iceland; e-mail: runson@hi.is. Address all correspondence to Oana Balan.*

