VIRTUAL REALITY GAME FOR TRAINING THE VISUALLY IMPAIRED IN SENSORY SUBSTITUTION

Silviu IVAȘCU[1], Alin MOLDOVEANU[2], Florica MOLDOVEANU[3], Anca MORAR[4], Victor ASAVEI[5], Cristian LAMBRU[6], Ana-Maria ȚUGULEA[7]

This paper presents a virtual reality application intended as a training tool for the visually impaired in sensory substitution. Although the proposed solution is dedicated to visually impaired people (VI), it can be used by sighted people as well. The application can be classified as an orientation and mobility (OM) game, where the players perceive the virtual environment only through audio and haptic stimuli. Since the visual representation of the environment was purposely eluded, special attention was paid to the sonification strategies and to the haptic feedback. Furthermore, to increase immersion by simulating real phenomena, advanced sound design has been incorporated. The game logic guides the user to navigate a labyrinth and collect several items, by using only sound and vibration. We performed a multi-modal evaluation, to assess both the opinion of users regarding sound orientation capability quality and intuitiveness, and their performance in accomplishing the game tasks. The test findings show that the users adapted with relative ease to the sonification models and therefore were able to solve straightforward the game problematique.

Keywords: VIP, visually impaired person, virtual reality, virtual environment

1. Introduction

According to the World Health Organization (WHO) report on vision in 2019[1], there are 2.2 billion people who have vision impairment. Of these people, there are at least 1 billion that have a vision impairment which could have been prevented or that can be addressed using rehabilitation procedures. In the

[1] Founder of Augmented Arts, Romania, e-mail: silviu.ivascu@augmentedarts.ro
[2] Prof., Dept. of Computer Science, University POLITEHNICA of Bucharest, Romania, e-mail: alin.moldoveanu@cs.pub.ro
[3] Prof., Dept. of Computer Science, University POLITEHNICA of Bucharest, Romania, e-mail: florica.moldoveanu@cs.pub.ro
[4] Assoc. Prof., Dept. of Computer Science, University POLITEHNICA of Bucharest, Romania, e-mail: Anca.morar@cs.pub.ro
[5] Assist. Prof., Dept. of Computer Science, University POLITEHNICA of Bucharest, Romania, e-mail: victor.asavei@cs.pub.ro
[6] Teaching Assist., Dept. of Computer Science, University POLITEHNICA of Bucharest, Romania, e-mail: andrei.lambru@upb.ro
[7] Dept. of Computer Science, University POLITEHNICA of Bucharest, Romania, e-mail: ana_maria.tugulea@upb.ro

past 30 years, by virtue of concentrated action, progress has been made regarding data accuracy due to an increased number of population-based surveys.

It is reasonable to say that vision, being the most dominant of our senses, plays a salient role in our lives. From a social standpoint, interpersonal interactions are in the lack of non-verbal cues such as facial expressions and gestures in the case of visually impaired participants. If the visual problems are present early in life, for people in this conjecture, the negative impact on their formative process is greater, as the available diversity of educational materials and methods is narrowed, which leads to limited development of social skills, lower self-esteem, and perturbations in their well-being.

Vision also plays an important role in social activities and sports, essential to mental and physical health, socialization, and personal character definition.

As it is the case for many adults, having the sense of vision, furthers employment and contributes to the economy as most work-related available activities cannot be accomplished in lack of it. Occupation exercising facilitates personal independence and social contact, maintaining individual well-being while contributing to societal development.

As Dela Torre and Khaliq mention in their work, games are usually lacking accessibility for the VI community [2]. Some of the techniques presented in their paper (auditory navigation, binaural audio and sonification techniques) have been exploited in our work, to achieve accessibility for the VI.

2. Related work

Several early works on gamification for visually impaired, as audio and sometimes audio-haptic games and virtual environments, are described in [3-13].

Corresponding to our application, is the virtual environment which has been utilized for the training of VI users, during the alpha testing phase of the Sound of Vision (SOV) device [14] [15]. The Sound of Vision project [16-23] paid utmost attention to developing games and virtual environments as an efficient mean to train visually impaired users in how to use the device [20-36]. This device has the propriety of guiding the user to be able to freely navigate in an unknown environment, by means of translating wearable camera images into audio signals and vibrations.

As relevant example of a VI game that is currently available and which is based on similar principles of functionality as the ones we employ, is the Nintendo Switch exclusive, called 1-2 Switch [2]. This application features a series of 28 small games, which all use the same hardware specific game mechanics. In case of the VI users, the visual aspect of these games can be removed but then, all the levels can still be played because as main output, the feedback for winning or losing is communicated verbally and in the case of navigation, the accelerometer and buttons from each gamepad are necessary. An important feature is that the 1-2 Switch allow to be played by either one or two people, so they are competitive. Since the games do not feature real-world situations, their main purpose is entertainment.

Another virtual environment entertainment application, with a focus on sound design, is the work of Fizek et al., the Audio Game Hub [37]. This application features 8 games which are available on both mobile and desktop platforms.

Other related work to be mentioned for the purpose of the specific methodologies used during their design, are the applications studied by Mattheiss et al. [38] and Baker et al. [39].

The originality of the work presented in this paper comes from multiple factors. First, the sonification of the scene uses Google Resonance audio. This type of environment sonification is very close to reality as it simulates sound occlusion, reflection and reverberation, physical wave phenomena, detailed in the following chapters. The sonification modelling also considers the type of material which the obstacles are made of. As a practical aspect, the virtual vehicle controls and behavior have been modeled using real quadcopter device design, which the VIP can effectively train for using this application and pilot after.

With respect to the virtual training environment used by the Sound of Vision project, our current project presents two fundamental differences: the entertainment feature of our application and, the applied sonification methods compatible to our virtual vehicle (quadcopter). If the Sound of Vision software has been designed to serve only special purposely designed hardware, ours is modelled to train pilots with or without the sense of vision, taking into consideration an abstraction of a quadcopter and the generic controls specific to these real vehicles. Furthermore, by comparison to the Sound of Vision project, our application can be used for flight pilot training and also for entertainment purposes, hence the sonification methods are specifically adapted to the virtual vehicle quadcopter design and environment. If Sound of Vision is a hardware and software solution for VIP, designed as an aid for independent living, the aim of this project is to provide an interactive entertainment solution, for bettering the quality of social life (the VIPs can play a complex game with other -impaired or not- people).

3. Application design

The proposed application was designed for both desktop and mobile (android only for the time being). The technologies used during the implementation are Unity game engine, the Google Resonance Audio SDK for Unity, Ableton Live 11 - as our Digital Audio Workstation (DAW) - and XAMPP for our web server which stores the statistical data. The only difference between the two application versions - desktop and mobile - is that only the mobile version has support for the head tracking functionality. The core game loop of the application is centered around the user navigating a labyrinth and collecting several items. The user navigates with a virtual flying vehicle which is controlled with a gamepad to collect objects scattered in each scene.

The choice of app deployment on mobile platforms has considered the game scenario of a person sitting inside a flying vehicle, whilst the tracking of the rotation movements of his head constitutes relevant perception freedom parameters. In this case, the users cannot move, they can only rotate their head to hear the sounds from a different perspective.

An important design objective was to make use only of consumer electronics. Custom hardware will always be more efficient, as in the case of the Sound of Vision project [40], yet by using consumer electronics we hope for a fast and wide software distribution, hence the cost-effective solution. The required hardware components for the mobile version of the proposed application are:

- Android smartphone (Fig. 1 a)
- Stereo headphones (Fig. 1 b)
- Gamepad (preferably Xbox One) (Fig. 1 c)
- Mobile phone headset (Optional) (Fig. 1 d)

Fig. 1. Hardware consumer electronics

An over the ear type of headphones is recommended to achieve a higher degree of external noise dampening, although any type of headphones will work.

The mobile version features the capability of using the phone's in-built accelerometer for tracking the rotation of the user's head.

The desktop version of the proposed application requires:

- Computer
- Stereo headphones
- Gamepad (preferably Xbox One)

When using the desktop version, the hardware lacks head-tracking support, so we recommend that the users maintain a fixed position of their head, because only rotating with the virtual vehicle (using the gamepad) is possible.

4. User interaction

The initial tests included a virtual vehicle control scheme, designed to be conceptually close to an airplane type of control, as shown in Fig. 2.

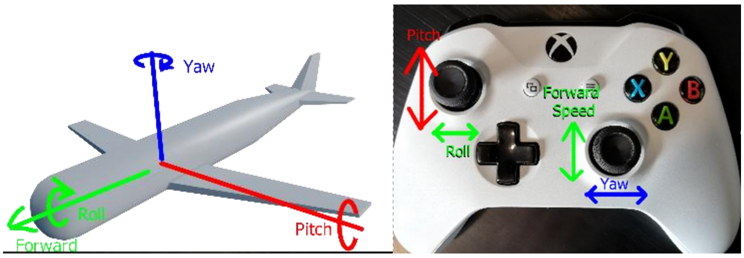

Fig. 2 Airplane control scheme

The virtual airplane model had, by default, a slow forward velocity. This velocity could be altered by tilting the right joystick on the Y axis. Tilting the right joystick up on the Y axis increased the forward velocity, tilting it down on the Y axis reduced the velocity to the minimum default. The main disadvantage observed in this case has been that the user could not maintain a fixed position necessary for orientation with audio feedback. Tilting the same right joystick on the X axis to the right, would change the Yaw of the plane. From the user's point of view, this caused the tip of the plane to rotate toward right.

Tilting the left joystick on the Y axis was used to alter the pitch and roll of the airplane like an airplane handle. Tilting the left joystick up on the Y axis made the tip of the airplane move down, lowering the altitude. Tilting the left joystick right on the X axis rotates the airplane right (rolls the airplane right). By tilting the left joystick right or left, the user could rotate to the point of flying upside down. The roll of the aircraft was automatically adjusted periodically to avoid the situation in which the user was flying upside down. After a few days of testing, we concluded that the average user could not adjust to this control scheme. The user’s feedback was constantly confusing as they felt like the aircraft movement was not correlating with their commands. Based on this preliminary feedback, we adjusted the controls to the second test version which is closer to the filming/photography drones’ control scheme.

Using the right joystick, the user can move the vehicle on the X and Y axis (red and green axis in Fig. 3). By tilting the right joystick on the vertical axis, the vehicle moves forward and backward, and by tilting the right joystick on the horizontal axis, the vehicle moves left and right, like a quadcopter.

The left joystick is used to rotate the vehicle along its Z axis. By tilting the left joystick on its horizontal axis, the user can modify the yaw of the vehicle. Tilting the joystick on the vertical axis, the altitude of the vehicle is modified (Fig. 3).
The main advantages of this control scheme are the following:

- The user could stay still and focus on the sounds
- Speed could be controlled more efficiently
- User's roll is constant, thus the problem of flying upside down is fixed

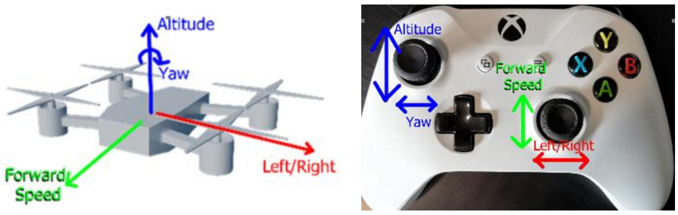

Fig. 3 Quadcopter/Drone control scheme

5. Level design

The virtual environment consists of simple scenes, which are paths containing only 90° turns. The difficulty of each scene is given by the number of targets needed to be collected, the length of the path, and the number of turns. To get a sense of orientation, the opening of the path is 20 meters wide by 20 meters high. The target to collect is a cube of $2m^3$ in volume and it is the same magnitude of the virtual vehicle controlled by the user. Targets are always placed away from the walls. Fig. 4 illustrates views of the environment, to offer a sense of the object magnitudes involved as they could be perceived with sight. The software does not render any scene geometry on the screen.

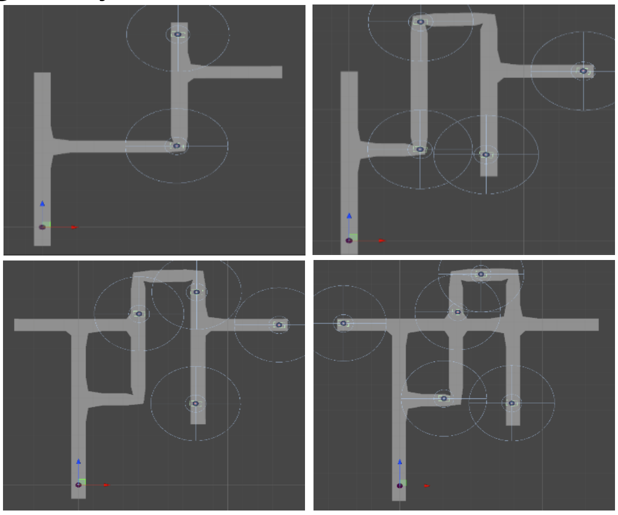

Fig. 4 Scenes seen from above in order of difficulty from easy (top left) to hard (bottom right)

6. Sonification and Sound design

Sonification

For sonification we have used the Google Resonance Audio SDK for Unity which supports HRTFs (Head Related Transfer Functions), with different types of sound occlusion and convolution reverb calculated based on the volume and the material of the walls. Further we will succinctly present the relevant details of our sonification design and functionalities.

The HRTFs are used as provided by Google, with no further formulae or code alterations. For simplifying of the VI user training process, we decided to apply a single material for all the walls of the scenes, so the occluded sounds could be rendered similarly in terms of the frequency spectrum. The Google Resonance Audio SDK helps the design to automatically add frequency filters to sounds which have an obstacle placed between the listener and the sound source inside the virtual environment. Predefined filters for each material assigned for the obstacles have been employed. The HRTFs simulates the shape of the human head, and the way sound reflects and passes through the human head before it is perceived. Convolution reverb is a special type of sound effect that is used to simulate the reverberation of sound when it is circulating through a room. It is based on a mathematical convolution operation, and it usually needs a pre-recorded sample of the space it will simulate. The pre-recorded sample is called an Impulse Response. Resonance Audio SDK provides a method of automatically calculating the convolution reverb settings without the need of an Impulse Response sample. Only the definition of the volume in which to calculate the convolution reverb and the materials that the virtual objects are made of like concrete, wood etc. is necessary. By using the obstacles with their predefined volumes and materials, an accurate simulation of the reverberation in that virtual room has been created.

In Addition to the aid Resonance Audio SDK provides, we have implemented a supplementary sonification model, with the purpose of guiding the aircraft through the scene. The sounds which are used, are as follows:

- **Collision** sounds have 4 levels of intensity designated based on the speed at which the user collided with an obstacle and are triggered at the moment of impact. A collision also triggers haptic feedback. The phone vibrates briefly when the user collides with an obstacle.
- **The Distance** between the user and the closest front obstacle is modeled as a beeping sound like car parking sensors. The amplitude of the signal is modulated based on a digital wave. The volume is either 0 or 1. The frequency of the modulator wave is increased based on the distance between the user and the obstacle.

- **Targets** constantly emit sound which is reflected by the surfaces of other objects in the scene, like walls or obstacles. Once the user is directly facing the target, an **additional sound wave** is played which varies in amplitude based on the angle between the forward direction of the user and the position of the target. The closer the target is to the center of the user view screen – by looking directly at it-, the higher the amplitude of the additional sound gets.

The sonification model described acts as a sensory substitution device [40] and thus extensive user training is needed to play the game at its full capacity. Our training sessions consisted of verbally explaining how the audio feedback should be used and short 20 min sessions of in-app specific exercises. During the usage of the application, statistical data has been collected and stored.

For storing the statistical data, we used a web server and a MySQL database. For each training session, the following data is collected:

- Anonymous user information: age, height, weight
- Session information: application start time (including time spent using the menu), end time, elapsed time (time spent only playing game levels)
- Scene information: name of the scene, time spent in the scene, average target collection time, total number of targets, collected targets, number of collisions, distance travelled.
- User's trajectory in each scene (each level) in case the user got stuck in a certain part of the scene due to poor level design or faulty collision.

Sound design

For the sound design process, we have used Ableton as our main DAW, as it offers the main tools to process the sound efficiently without 3rd party software. Ableton provides a digital frequency equalizer, compression, a synthesizer to generate the tones needed and the possibility to process pre-recorded sounds for simulating real processes like crashing the vehicle at different speeds. Techniques described in the works of Rovithis et al. [41], Zattra et al. [42] and Delle Monache et al. [43] have been utilized.

A Wavetable Synthesizer is one of the simplest ways to generate and process sounds. The reason we chose a wavetable synthesizer instead of one based on frequency modulation is due to the ease of use by comparison to the former. As in our case, the wavetable synthesizer uses premade waveforms thus having a quicker process of generating sounds. Frequencies for the audio stimuli suitable for our project are chosen based on the research of Nan. et al. [44], Spoendlin and Schrott [45], Riecke et al. [46] and Sininger et al. [47]. For the sonification model that we are using, the repetitive sounds harmonics are situated predominantly on the lower half of the hearing spectrum range (<1000 Hz) and present scarcely.

This translates into a “smoother” waveform Fig. 5 (left). For sounds that need the user’s immediate attention we use the 1000 – 2000 Hz range or higher with multiple harmonics, projecting a “harsher” sound and waveform Fig. 5 (right). The frequency intervals have been chosen based on the above-mentioned research, and on the tests with users.

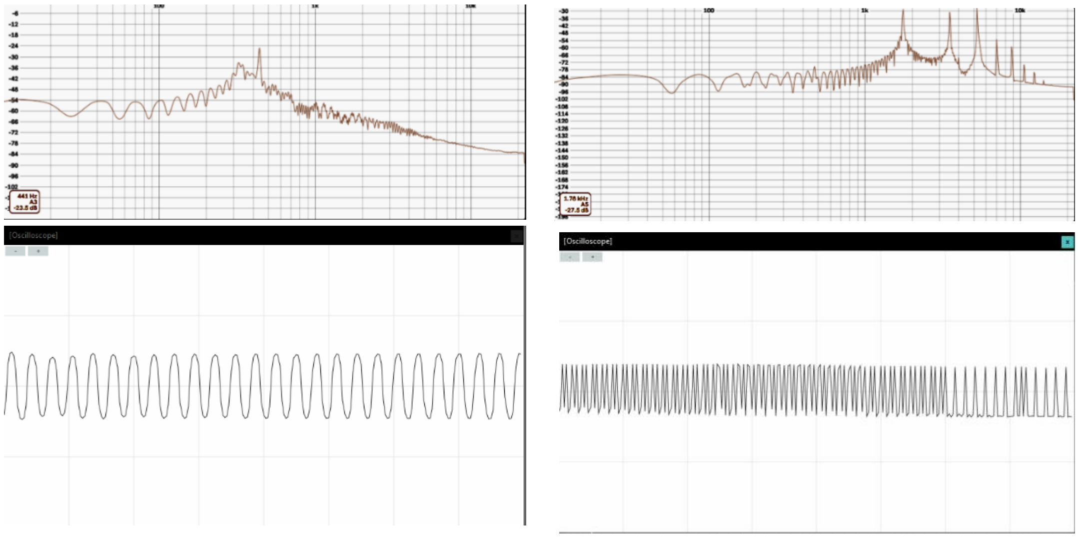

Fig. 5 (**Left**) Repetitive "smooth" sound. Top – audio frequency spectrum. Bottom – waveform. (**Right**) Rare "harsh" sound. Top – audio frequency spectrum. Bottom – waveform.

7. Testing and assessment

We performed several tests on three blindfolded users for one month and, based on their feedback, we adjusted the application accordingly. The users were asked for feedback regarding sound quality and level of satisfaction and their experience regarding video games ranged from none to very experienced. After adjusting the sound's frequency spectrums, the result was less ear fatigue resulting in longer training sessions. Tests were also made with different scene complexities leading to the final scenes which were like the ones presented in Fig. 6. Testing concluded that by increasing the complexity of the scenes more fatigue is perceived by the players. As users became fatigued, training sessions shortened due to user decision flaws and inconclusive test responses and results. The same feedback form has been used for evaluating the interaction methods and as explained by Fig. 2 and Fig. 3, the second design of the interaction method has proven to work better.

Our statistics system collects data regarding the movement of the users through the virtual environment, thus the 3D renders of the trajectories can be analyzed, to methodically interpret the collected data. In Fig. 6 we can see examples of movement trajectories, for different users, on the last two testing scenes.

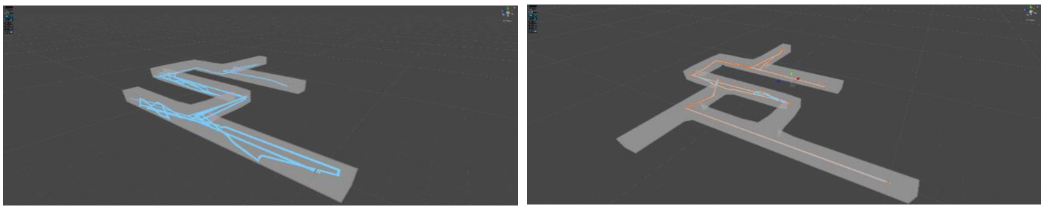

Fig. 6 Examples of Scene 1 and Scene 2 sections with trajectories

Another problem regarding the level design was found in the case of the 3rd testing scene, shown in Fig. 7. The users have a tendency of moving directly towards the target sound, thus having a problem of finding the correct path. For example, in Fig. 7 one user starts at the bottom left of the figure. The first two targets are easy to find. From this point, as seen from the blue trajectory, the user needs a lot of time to find the last two targets due to the lack of sonification of the path. Obstacles are not a problem for the sonification model in most cases.

The problem of finding the correct path could be addressed in two ways. The first method would be to fix the design of the scene and add another target pickup to guide the user. The second method would be to add a distinct sound playable when a correct path is chosen, towards the next target. This user orientation issue will be addressed in a future version of the application.

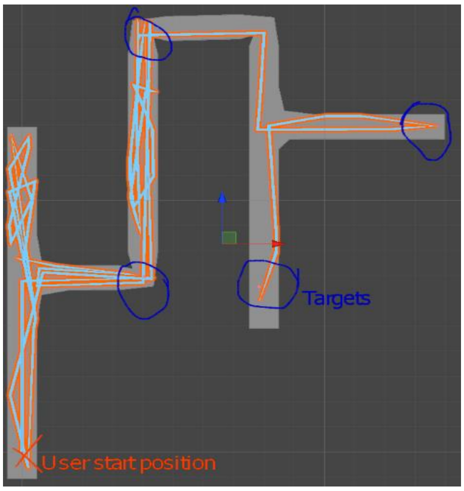

Fig. 7 Level design problem. Scene 3 seen from above

Regarding the results of our tests, our main performance indicator is the Average Time to Collect Target (ATCT). This parameter is calculated per scene when the user collects all the targets. It represents the total time it took to collect all the targets in the scene starting from the moment of the first collected target, divided by the number of targets in the scene. As illustrated in Table 1, the user with the ID 2 had an increase of ATCT for scene 2. The number of collisions for Scene 2 is decreasing as the user gets accustomed with the sonification methods. Scene 2 also has a total of 4 test sessions. For scene 3 ATCT is inconclusive, and the collisions are increasing. This is influenced by the prevalence of fatigue symptoms like wall collision at a higher rate, decision flaws like forgetting the mapping of the controls, etc.

Although the game itself may seem unnatural for the users which are not lacking the sense of sight, it is engaging from the point of view of attainable targets. The immersion within the virtual environment is rapid as the controls and the logic are quite simple but, to become proficient, a fair amount of play and practice are needed. Once the sonification significance becomes nervous reflex, the user can demonstrate skill and tactics to reach each target inside the labyrinth. Qualitatively, the game addresses well the problem of training users to use this sonification model. The targets, obstacles and paths are recognizable and, virtualization helps with offering the players the lack of fear when navigating without sight on an obstacle course.

8. Conclusions and future work

After the preliminary tests and adjustments, we concluded that our application and sonification model can be used to navigate a virtual environment with a quadcopter type vehicle using only audio and rudimentary haptic feedback. The time it took users to get accustomed to the system was relatively short. After 3 sessions, most users get accustomed to the sonification model. Considering the feedback received from our users, we believe that longer sessions of using the application can greatly improve performance. The game can be used for training the visually impaired to navigate in a real environment based on audio and haptic feedback.

For future work regarding this type of application, our objectives include adding collaborative logic to the game, like having multiple users coordinating each other towards a common goal and adding a diversity of complex levels.

Table 1

Statistics for User 2

Time Level (seconds)	ATCT (seconds)	Total Targets	Targets Collected	Collisions	Distance (meters)	Scene ID	Session ID	Elapsed Time (seconds)	User ID
117.009	58.5046	2	2	4	263.173	2	5	459.214	2
94.8204	47.4102	2	2	2	260.585	2	6	409.613	2
62.5521	31.2761	2	2	1	252.956	2	7	692.818	2
58.0107	29.0053	2	2	0	244.106	2	8	1380.52	2
115.367	38.4557	3	3	1	419.689	3	6	409.613	2

137.357	45.7856	3	3	6	419.997	3	7	692.818	2
110.719	36.9064	3	3	4	407.297	3	8	1380.52	2
296.416	74.104	4	4	8	794.485	4	8	1380.52	2
473.689	118.422	4	4	23	1117.52	5	8	1380.52	2
369.916	73.9833	5	5	18	696.821	6	8	1380.52	2

Time Level is the time in seconds spent in the scene, **Elapsed time** is the amount of time spent using the application.

Acknowledgments

Financial support from the Competitiveness Operational Program 2014-2020, Action 1.1.3: Creating synergies with RDI actions of the EU's HORIZON 2020 framework program and other international RDI programs, MySMIS Code 108792, Acronym project "UPB4H", financed by contract: 250/11.05.2020.

R E F E R E N C E S

[1] World Health Organization, "World report on vision," World health Organization, vol. 214, no. 14, pp. 1–160, 2019.

[2] I. Khaliq and I. dela Torre, "A study on accessibility in games for the visually impaired," in ACM International Conference Proceeding Series, 2019, pp. 142–148. doi: 10.1145/3342428.3342682.

[3] Bălan Oana, A. Moldoveanu, F. Moldoveanu, and O.-M. Ferche, "O scurta clasificare a jocurilor audio în contextul îmbunatatirii interactiunii si accesibilitatii pentru persoanele nevazatoare - ProQuest," Romanian Journal of Human - Computer Interaction, MatrixRom, vol. 7, no. 3, pp. 225–238, 2014.

[4] A. Sicaru, A. Morar, A. Moldoveanu, F. Moldoveanu, and M. Mocanu, "Computer gaming for the visually impaired - Status and perspectives," in Proceedings - 2018 IEEE 14th International Conference on Intelligent Computer Communication and Processing, ICCP 2018, Oct. 2018, pp. 145–151. doi: 10.1109/ICCP.2018.8516605.

[5] A. Moldoveanu et al., "Virtual Mini-Games - A Serious Learning Tool for Sensory Substitution Devices," Proceedings of the EduLearn Conference, Barcelona, Spain, pp. 3–5, Jul. 2017.

[6] O. Bălan, A. Moldoveanu, F. Moldoveanu, H. Nagy, G. Wersényi, and R. Unnþórsson, "Improving the audio game–playing performances of people with visual impairments through multimodal training," Journal of Visual Impairment and Blindness, vol. 111, no. 2, pp. 148–164, Mar. 2017, doi: 10.1177/0145482X1711100206.

[7] A. Csapo et al., "Usability and Effectiveness of Auditory Sensory Substitution Models for the Visually Impaired," Conference on Audio Engineering Society Convention 142, Audio Engineering Society, May 2017

[8] A. Butean, O. Balan, A. Moldoveanu, and F. Moldoveanu, "Developing a navigational 3D audio game with hierarchical levels of difficulty for the visually impaired players," in Romanian Conference on Human Computer Interaction - RoCHI 2015, Sep. 2015, pp. 49–54.

[9] S. Spagnol et al., “Model-based obstacle sonification for the navigation of visually impaired persons,” in Proc. 19th Int. Conf. Digital Audio Effects (DAFx-16), Sep. 2016, pp. 309–316.

[10] O. Bălan, A. Moldoveanu, F. Moldoveanu, and M.-I. Dascălu, “Navigational 3D audio-based game-training towards rich auditory spatial representation of the environment,” in 2014 18th International Conference on System Theory, Control and Computing (ICSTCC), 2014, pp. 682–687. doi: 10.1109/ICSTCC.2014.6982496.

[11] O. Bălan, A. Moldoveanu, F. Moldoveanu, and M.-I. Dascalu, “Audio Games - A Novel Approach Towards Effective Learning in the Case of Visually- Impaired People,” in In Proceedings of Seventh international conference of education, research and innovation, Seville, Nov. 2014, pp. 6542–6548.

[12] O. Bălan, A. Moldoveanu, and F. Moldoveanu, “Navigational audio games: An effective approach toward improving spatial contextual learning for blind people,” International Journal on Disability and Human Development, vol. 14, no. 2, pp. 109–118, 2015, doi: 10.1515/IJDHD-2014-0018/HTML.

[13] O. Ferche et al., "From neuromotor command to feedback: A survey of techniques for rehabilitation through altered perception" in 2015 E-Health and Bioengineering Conference (EHB), Nov. 2015, pp. 1-4, doi: 10.1109/EHB.2015.7391454.

[14] A. D. B. Moldoveanu et al., “Mastering an advanced sensory substitution device for visually impaired through innovative virtual training,” in 2017 IEEE 7th International Conference on Consumer Electronics - Berlin (ICCE-Berlin), Sep. 2017, pp. 120–125. doi: 10.1109/ICCE-Berlin.2017.8210608.

[15] O. Bălan, A. Moldoveanu, F. Moldoveanu, H. Nagy, G. Wersényi, and R. Unnþórsson, “Improving the audio game–playing performances of people with visual impairments through multimodal training,” Journal of Visual Impairment and Blindness, vol. 111, no. 2, pp. 148–164, 2017, doi: 10.1177/0145482x1711100206.

[16] A. Moldoveanu, F. Moldoveanu, and O. Bălan, “Binaural Sound Analysis and Spatial Localization for the Visually Impaired People,” in 9th International Conference on Interfaces and Human Computer Interaction, Jul. 2015, pp. 331–335.

[17] S. Caraiman et al., “Computer Vision for the Visually Impaired: The Sound of Vision System,” in 2017 IEEE International Conference on Computer Vision Workshops (ICCVW), Oct. 2017, pp. 1480-1489, doi: 10.1109/ICCVW.2017.175.

[18] M. Bujacz et al., “Sound of Vision - Spatial Audio Output and Sonification Approaches,” in Computers Helping People with Special Needs, 2016, pp. 202–209.

[19] A. Moldoveanu et al., “Sound Of Vision 3d Virtual Training Environments - A Gamification Approach For Visual To Audio-Haptic Sensory Substitution,” Rev. Roum. Sci. Techn.-Électrotechn. et Énerg, vol. 63, pp. 112–117, 2018.

[20] Ó. Jóhannesson, O. Balan, R. Unnthorsson, A. Moldoveanu, and Á. Kristjánsson, “The Sound of Vision Project: On the Feasibility of an Audio-Haptic Representation of the Environment, for the Visually Impaired,” Brain Sciences, vol. 6, no. 3, p. 20, Jun. 2016, doi: 10.3390/brainsci6030020.

[21] A. Butean, A. Moldoveanu, and O. Mitruț, “Introducing Basic Geometric Shapes to Visually Impaired People Using a Mobile App Sound of Vision View project 3D Smith Chart View project,” in Romanian Journal of Human - Computer Interaction, 2015, pp. 91–94.

[22] O. Bălan, A. Moldoveanu, F. Moldoveanu, and A. Butean, "Spatial Sound Based System for Improving Orientation and Mobility Skills in the Absence of Sight," UPB Scientific Bulletin, vol. 77, no. 3, Dec. 2015.

[23] O. Bălan, A. Moldoveanu, F. Moldoveanu, and A. Butean, "Auditory and Haptic Spatial Cognitive Representation in the Case of the Visually Impaired People," The 22nd International Congress on Sound and Vibration, Florence, Italy, pp. 12–16, Jul. 2015

[24] A. Butean, B. Troancă, O. Bălan, F. Moldoveanu, A. Moldoveanu, and D. Chirita, "Applications on Touchscreen Mobile Devices for Visually Impaired People," Romanian Journal of Human-Computer Interaction, Published by MATRIX ROM, vol. 8, no. 2, pp. 121–138, 2015.

[25] O. Mitruț, A. Moldoveanu, and F. Moldoveanu, "Spatial Auditory Representation in the Case of the Visually Impaired People," Journal of Information Systems & Operations Management, vol. 8, no. 1, pp. 1–11, May 2015

[26] O. Bălan, A. Moldoveanu, A. Butean, and F. Moldoveanu, "Comparative Research on Sound Localization Accuracy in the Free-field and Virtual Auditory Displays," The 11 th International Scientific Conference eLearning and software for Education Bucharest, " Carol I" National Defence University, vol. 1, pp. 540–548, Apr. 2015, doi: 10.12753/2066-026X-15-079.

[27] A. D. B. Moldoveanu et al., "Virtual environments for training visually impaired for a sensory substitution device," in 2017 Zooming Innovation in Consumer Electronics International Conference: Galvanize Your Creativity, ZINC 2017, Jul. 2017, pp. 26–29. doi: 10.1109/ZINC.2017.7968654.

[28] O. Bălan, A. Moldoveanu, F. Moldoveanu, A. Morar, and S. Ivaşcu, "Perceptual feedback training for improving spatial acuity and resolving front-back confusion errors in virtual auditory environments," in 2017 40th International Conference on Telecommunications and Signal Processing, TSP 2017, Oct. 2017, vol. 2017-January, pp. 334–337. doi: 10.1109/TSP.2017.8075999.

[29] O. Bălan, A. Moldoveanu, and F. Moldoveanu, "A Systematic Review of the Methods and Experiments Aimed to Reduce Front-Back Confusions in the Free-Field and Virtual Auditory Environments," in Romanian Journal of Human - Computer Interaction, 2018, pp. 24–29.

[30] M. Dascalu et al., "Brain activation during virtual and real-world mobility tasks with single and multimodal sensory input provided by an assistive navigational device for visually impaired people," in 2018 25th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS), May 2018, pp. 1–6. doi: 10.23919/ICINS.2018.8405870.

[31] D. Flamaropol, A. Moldoveanu, F. Moldoveanu, M. I. Dascalu, I. Stanica, and I. Negoi, "User Behavior Analytics in Virtual Training Environments for Sensory Substitution Devices," 2018 Zooming Innovation in Consumer Technologies Conference, ZINC 2018, pp. 22–26, Aug. 2018, doi: 10.1109/ZINC.2018.8448514.

[32] R.-G. Lupu, O. Mitruț, A. Stan, F. Ungureanu, K. Kalimeri, and A. Moldoveanu, "Cognitive and Affective Assessment of Navigation and Mobility Tasks for the Visually Impaired via Electroencephalography and Behavioral Signals," Sensors, vol. 20, no. 20, 2020, doi: 10.3390/s20205821.

[33] O. Bălan, A. Moldoveanu, and F. Moldoveanu, "Multimodal Perceptual Training for Improving Spatial Auditory Performance in Blind and Sighted Listeners," Archives of

Acoustics, vol. Vol. 40, No. 4, no. 4, pp. 491–502, Dec. 2015, doi: 10.1515/AOA-2015-0049.

[34] O. Bălan et al., "Haptic-auditory perceptual feedback based training for improving the spatial acoustic resolution of the visually impaired people," in International Conference on Auditory Display (ICAD), Jul. 2015, pp. 73–79.

[35] A. Moldoveanu, O. Balan and F. Moldoveanu, "Training system for improving spatial sound localization," in The International Scientific Conference eLearning and Software for Education, 2014, vol. 4, pp. 79-84.

[36] O. Bălan, A. Moldoveanu, F. Moldoveanu, and I. Negoi, "The Role of Perceptual Feedback Training on Sound Localization Accuracy in Audio Experiments," eLearning & Software for Education, " Carol I" National Defence University, no. 1, pp. 502–509, 2015.

[37] S. Fizek, J. Woletz, and J. Beksa, "Playing with sound and gesture in digital audio games," in Mensch und Computer 2015 - Workshop, De Gruyter, 2015, pp. 423–430. doi: 10.1515/9783110443905-061.

[38] E. Mattheiss, G. Regal, D. Sellitsch, and M. Tscheligi, "User-centred design with visually impaired pupils: A case study of a game editor for orientation and mobility training," International Journal of Child-Computer Interaction, vol. 11, pp. 12–18, 2017, doi: 10.1016/j.ijcci.2016.11.001.

[39] R. M. Baker, K. Ramos, and J. R. Turner, "Game design for visually-impaired individuals: Creativity and innovation theories and sensory substitution devices influence on virtual and physical navigation skills," Irish Journal of Technology Enhanced Learning, vol. 4, no. 1, p. 36, 2018, doi: 10.22554/ijtel.v4i1.51.

[40] A. D. B. Moldoveanu et al., "Mastering an advanced sensory substitution device for visually impaired through innovative virtual training," in IEEE International Conference on Consumer Electronics - Berlin, ICCE-Berlin, 2017, vol. 2017-Septe, pp. 120–125. doi: 10.1109/ICCE-Berlin.2017.8210608.

[41] E. Rovithis, N. Moustakas, A. Floros, and K. Vogklis, "Audio legends: Investigating sonic interaction in an augmented reality audio game," Multimodal Technologies and Interaction, vol. 3, no. 4, p. 73, 2019, doi: 10.3390/mti3040073.

[42] L. Zattra, N. Misdariis, F. Pecquet, N. Donin, and D. Fierro, "Analysis of Sound Design Practices [ASDP]. Research methodology," in Machine sounds, Sound machines - Proceedings of the XXII CIM Colloquium on Music Informatics, 2018, pp. 168–175.

[43] S. Delle Monache, D. Rocchesso, F. Bevilacqua, G. Lemaitre, S. Baldan, and A. Cera, "Embodied sound design," International Journal of Human Computer Studies, vol. 118, pp. 47–59, 2018, doi: 10.1016/j.ijhcs.2018.05.007.

[44] Y. Nan, E. Skoe, T. Nicol, and N. Kraus, "Auditory brainstem's sensitivity to human voices," International Journal of Psychophysiology, vol. 95, no. 3, pp. 333–337, 2015, doi: 10.1016/j.ijpsycho.2014.12.013.

[45] H. Spoendlin and A. Schrott, "Analysis of the human auditory nerve," Hearing Research, vol. 43, no. 1, pp. 25–38, 1989, doi: 10.1016/0378-5955(89)90056-7.

[46] L. Riecke et al., “Frequency-specific attentional modulation in human primary auditory cortex and midbrain,” NeuroImage, vol. 174, pp. 274–287, 2018, doi: 10.1016/j.neuroimage.2018.03.038.

[47] Y. S. Sininger, C. Abdala, and B. Cone-Wesson, “Auditory threshold sensitivity of the human neonate as measured by the auditory brainstem response,” Hearing Research, vol. 104, no. 1–2, pp. 27–38, 1997, doi: 10.1016/S0378-5955(96)00178-5.