Developing a Navigational 3D Audio Game with Hierarchical Levels of Difficulty for the Visually Impaired Players

Oana Balan
University Politehnica of Bucharest, Faculty of Automatic Control and Computer Science
Splaiul Independentei 313, Bucharest 060042, Romania
oanab_2005@yahoo.com

Alin Moldoveanu
University Politehnica of Bucharest, Faculty of Automatic Control and Computer Science
Splaiul Independentei 313, Bucharest 060042, Romania
alin.moldoveanu@cs.pub.ro

Florica Moldoveanu
University Politehnica of Bucharest, Faculty of Automatic Control and Computer Science
Splaiul Independentei 313, Bucharest 060042, Romania
florica.moldoveanu@cs.pub.ro

Alexandru Butean
University Politehnica of Bucharest, Faculty of Automatic Control and Computer Science
Splaiul Independentei 313, Bucharest 060042, Romania
alexandru@butean.com

ABSTRACT
The Binaural Navigation Game is a 3D audio game with hierarchical levels of difficulty that can be used by both normal sighted or visually impaired players. The purpose of the game is to test and train the sound localization skills, to entertain and to provide an alternative to the common video games available on the market nowadays. The sonification technique is based on the perception of 3D binaural sounds synthesized with non-individualized Head Related Transfer Functions (HRTFs) and on the inversely proportional sound intensity encoding of distance. Furthermore, we use an original method based on the simultaneous perception of two types of noise that aims to reduce the incidence of front-back confusions. This game has been tested in an experimental procedure (comprised of a pre-test, a haptic-auditory feedback based training and a post-test session) in which 10 visually impaired subjects (with a percent of residual vision ranging from 0% to 15%) have participated. This paper aims to present the design and development of the Binaural Navigation Game, the Game Editor (an application that allows the experimenter to set the layout of new sets of levels) and the Binaural Game Analyzer, a tool used to visualize and evaluate the players' performances..

Author Keywords
3D sound; HRTF; audio game; navigation skills; orientation and mobility skills.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms
Human Factors; Design; Measurement.

INTRODUCTION
Sound is an immersive and expressive medium that can convey a wide range of information on the meta-level of perception. In video or audio games, it provides powerful and reliable cues (concerning events and situations) that enhance the level of interactivity, engagement and the overall user game playing performance. Although in recent years many researchers and game developers have shifted their attention on the audio component, it is highly reduced in comparison to the visual content in most of the games available for PC or for mobile platforms [10]. Moreover, the audio-only games draw less interest than the video ones, being regularly developed by small groups of programmers or researchers and usually for experimental purposes [1].

The audio-only games provide accessibility features for the users who are partially or completely deprived of the visual sense. Furthermore, they should be designed with particular focus on the visually impaired players' particularities, skills and needs [4]. The audio games can be employed for educational purposes [2] or for training various abilities, for enhancing sound localization, improving the transfer of skills into real-world situations or for entertainment only [6, 14, 18, 21].

This paper presents the development of a navigational 3D audio game in which the entire navigation is based on the perception of 3D binaural cues that provide the listeners a complete spatial perception of the auditory setting. Although the sonification design is focused on the use of

3D binaural sounds as the only means of navigation, the acoustic environment is complemented by auditory icons and earcons [7] that enhance user interaction and provide a rapid response to the players' input. The purpose of the game is to find the location of several hidden targets while avoiding blocking obstacles. The number of targets and obstacles varies among levels, according to the difficulty.

The Binaural Navigation Game has been employed in an experimental procedure in which we tested the sound localization skills of 10 visually impaired people (5 women and 5 men, aged 27-63, with a percent of residual vision ranging from 0% to 15%) before and after a haptic-auditory feedback based training session. However, the focus of this paper is not on describing the sound localization experiment or its results (they will be discussed in another paper) but on introducing the developmental approach of the software tools used in our tests. The game level layout has been designed using the Game Editor software. Moreover, in order to assess the users' game playing performance, we have designed an application entitled Binaural Game Analyzer that provides visualization, real-time playback and statistical functionalities through an accessible and interactive interface.

THE SONIFICATION APPROACH

For encoding the direction of the target sound sources, we used a sonification technique based on the continuous and simultaneous perception of two types of noises (white and pink noise), in varying proportions, according to the direction of the sound source in space. To the front, the listener hears only white noise, while to the back he perceives solely the pink noise. In the right hemifield, the amount of white noise decreases (and that of pink noise increases), reaching equal levels at 90 degrees to the right. In the left hemifield, the relationship between the white and the pink noise is complementary to that from the right side, so that both noises record identical perceptive levels at 90 degrees to the left.

The aim of this encoding is to the help the listeners to differentiate the direction of the sound sources located to the front and to the rear and to avoid the very common reversal errors that occur in virtual auditory displays, generally known as front-back confusions [11, 20, 22] (situation when the listener encounters difficulties in determining whether the sound source is originating from the front or from the back). In this way, the perception of the spectral patterns of the noises, together with the directional cues provided by the 3D sounds convey a more accurate image of the location of the sound source in space. On the other hand, the obstacles have been encoded using a broadband alarm sound. Both types of sounds (the combination of white and pink noise and the alarm sound) have been processed using the Csound programming language [9] and synthesized with the non-individualized HRTFs from the MIT database [17].

The HRTFs are the basis of binaural rendering, as they represent the modifications that occur in the spectrum of the sound from the position of the source en route to the human subject's ears. These complex transfer functions describe the magnitude of the frequency response in correlation to the source position, defined by azimuth and elevation angle [3]. Non-individualized sets of HRTFs are generally recorded in anechoic chambers, using dummy head mannequins. However, as the HRTFs strongly depend on the morphological characteristics of the external auditory system (the pinna and concha have a non-linear frequency response [12]), the non-individualized (or generic) HRTFs introduce severe perception artifacts and localization errors [19]. On the other hand, recording personalized sets of HRTFs for each listener apart is a time-consuming and impractical task, so that most of the 3D audio based applications employ non-individualized HRTFs on a large scale. The simulation of the virtual position of a particular sound source in space can be achieved by filtering a given signal with the pairs of HRTFs corresponding to the left and to the right ear, creating thus left and right ear signals that are delivered through headphones [15, 19].

The HRTFs from the MIT database have been recorded using a dummy-head mannequin and a set of 7 loudspeakers placed at 1.4 meters distance from it. The elevation positions in the median plane range from -40 degrees (below the interaural axis) to +90 degrees (directly overhead). The azimuth measurement locations were disposed at 5 degrees angle increment all around the listener's head [17].

Csound is the first sound processing programming language, firstly developed in 1984 at the Massachusetts Institute of Technology and written entirely in C. Since then, it improved considerably, becoming today one of the most powerful and reliable music generating instruments. Moreover, it is available for the Linux, Macintosh and Windows operating systems, as well as for the new mobile platforms, such as Android [8].

In order to obtain 3D binaural stimuli for our game, we convolved the stereo sounds (the combination of white and pink noise in varying proportions and the alarm sound encoding the obstacle position) using the hrtfmove2 opcode (orchestra function, typical for the Csound programming language) [8]. The hrtfmove2 opcode generates 3D binaural sounds to be rendered over headphones for any location in the 3D space using the magnitude interpolation algorithm and a phase model based on the Woodworth's formula [13] for Interaural Time Difference (ITD) [13,22]:

$$\mathrm{ITD} = \frac{r}{c}(\theta + \sin\theta) \tag{1}$$

Where r is radius of the subject's head, θ is the angle between the listener and the sound direction and c is the speed of sound, of approximately 340 m/s at sea level.

Magnitude interpolation means that the four nearest HRTF values (left, right, below and above) are linearly interpolated in order to obtain the most appropriate magnitude level for the given angular position [12]. The piece of code that processes the given audio input into a 3D binaural sound is the following:

aleft, aright hrtfmove2 aSignal, kAzimuth, kElevation, $HRTF_LEFT, $HRTF_RIGHT

where aleft and aright are the output signals for the left and for the right channels, aSignal represents the input stimulus, kAzimuth is the azimuth angle, kElevation corresponds to the elevation angle (set by default to 0 in our game), while $HRTF_LEFT and $HRTF_RIGHT represent the left, respectively the right HRTF spectral data files.

In addition to 3D sounds, we used the modifications in the perceived sound intensity to convey relevant information concerning the distance between the current virtual position of the player and the location of either target objects or obstacles. The inversely proportional relationship between the sound intensity level (SI) and the distance is the following:

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

Where d is the current Euclidean distance between the virtual position of the player and the target object/obstacle, dmax is the maximum recorded distance to the active target (in the case of obstacles, it has a fixed value of 150 pixels), SIMIN=0.05 represents the minimum perceptible sound intensity and SIMAX=1 corresponds to the maximum perceivable auditory level.

The auditory icons are “short, icon-like sound events that have semantic connections to the physical events they represent” [7]. They are largely employed in audio games, as they provide the player an easily understandable connection between their significance and the situations and events with which they are associated. Examples of auditory icons are the sound of footsteps that give clues about someone approaching, the sound of a door opening, a dog barking etc. In our game, we used the sound of a crash to indicate the collision with an obstacle.

The earcons are short, simple, abstract sounds (sometimes composed only of a few notes) that have associated a relationship with the events they represent [7]. As they do not have instantly recognizable characteristics that could be achieved from previous experience, the listener needs to learn first the meaning of the earcons before using them in auditory environments. The earcons we used in our game are a sparkling sound that can be heard when the player identifies the location of the hidden target sound source and the sound of a bell ringing that announces the end of the game.

THE BINAURAL NAVIGATION GAME

The Binaural Navigation Game (Figure 1) has been written in the C# programming language [5] and designed using the Microsoft Visual Studio Integrated Development Environment [16].

At the beginning of each game session, the player is required to introduce some personal information: name, age, sex, the number of years since he is suffering from the visual impairment and the number of the set level that he wants to play. The player can choose from different sets of levels that are created using the Game Editor application.

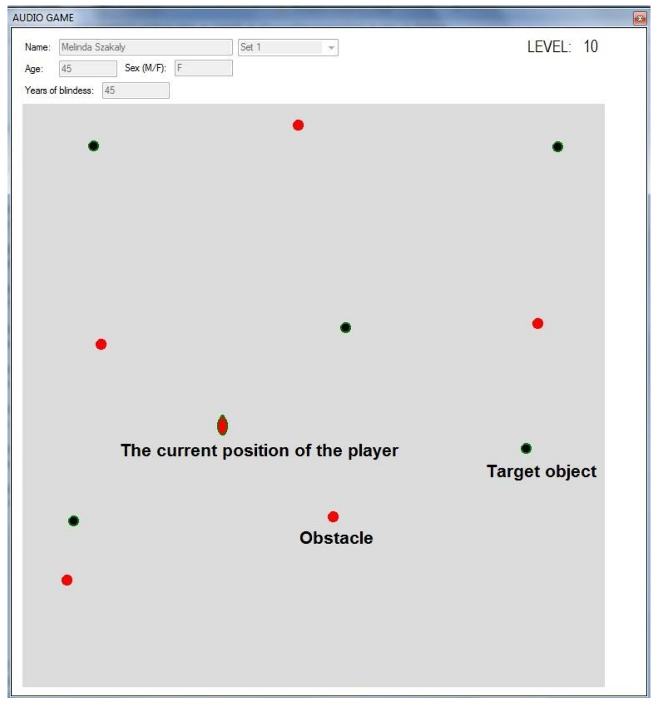

Figure 1 The Binaural Navigation Game

At each level, the player is required to navigate freely, using the mouse (or touchpad) movement as interaction method, in order to identify the location of the hidden target sound sources (only one target 3D sound can be heard at a moment of time), while avoiding blocking obstacles (the obstacles are audible within a 150 pixels range). After the location of the current target has been identified, it consequently turns silent and the next sound object becomes audible. If two or more obstacles are positioned at a distance smaller than 150 pixels, they are perceived simultaneously, although their corresponding 3D sounds vary in intensity and direction.

THE GAME EDITOR

The experimenter can design the layout of the game (the distribution of target auditory objects and obstacles) in a well-defined and interactive manner. Thus, he is required to introduce the level number and to place items (either targets or obstacles) on the canvas (Figure 2). The positions of the target objects need to be set first (by pressing the “Set

Targets” button from the user interface application), in ascending order, from the bottom side of the playing canvas to the top of it. Subsequently, the user is required to press the “Set Obstacles” button in order to start adding obstacles to the current level. In this case, no more than 5 obstacles per level can be introduced. The locations of the target objects or obstacles are set by clicking on the playing window in the positions where the items need to be placed. The game layout components are encoded by green (the target objects), respectively red circles (in the case of the obstacles).

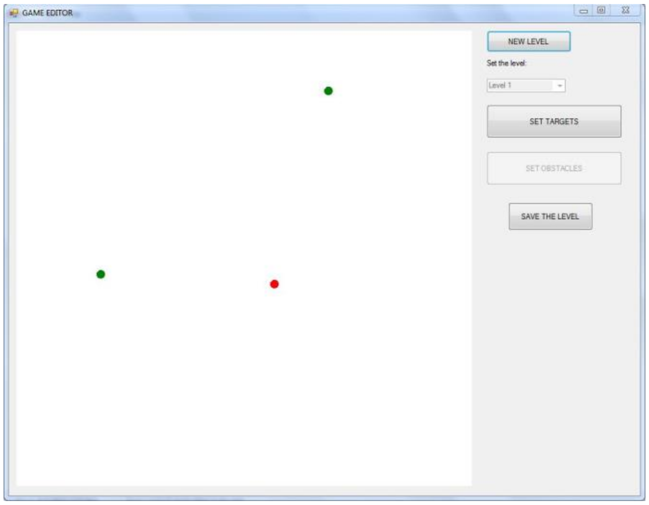

Figure 2 The Game Editor

The data for each level is stored in separate .txt files that comprise on distinct lines the items type (“t” for targets and “o” for obstacles) and the 2D coordinates in the horizontal plane for each of them. For example, the content of the “Level 3.txt” file, where we have 2 target sound sources and 2 obstacles is the following:

Level 3
t 127 481
t 627 36
o 201 559
o 630 125

THE BINAURAL GAME ANALYZER

In order to evaluate the game playing performance for each target object and for each level apart, we defined the following parameters:

P1 – the ratio of the distance travelled by the player from the starting position until he reaches the target sound source to the minimum possible distance between the starting position and the location of the target object (which is actually the Euclidean distance between these two points). For the first round of every level, the starting position is the center of the bottom side of the playing window, while for the others it is represented by the location of the previously identified target object.

P2 – the percent of correct travel decisions (movements performed towards the sound source, minimizing the distance to the target object).

P3 – the round completion time in seconds, i.e. the time needed to identify the location of the target source.

P4 – the number of obstacle collisions per round.

The Game Analyzer tool (Figure 3) allows real-time visualization and audio playback of the users’ performances for each round (for each target object identification) and level.

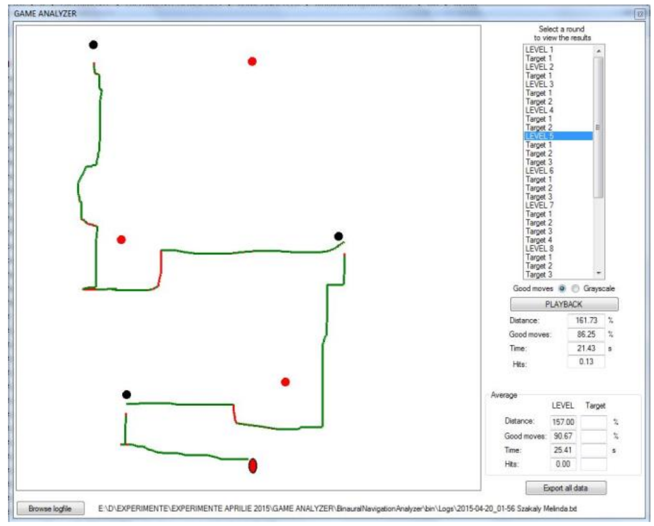

Figure 3 The Binaural Game Analyzer

This figure presents the player’s performance for level 5 (which has 3 sound targets and 3 obstacles), using the “Good moves” visualization option.

By pressing the “Browse log file” button, the experimenter can upload the log file that has been created and updated during the gameplay session and that stores all the data concerning the user’s performance. The experimenter can select a round or a level from the list box control on the right side of the interface. The “Good moves” visualization option displays the player’s efficient displacements (mouse movements effectuated towards the sound source) as green colored segments, while the wrong movements that maximize the distance to the sound source are represented by red segments. The “Grayscale” visualization option presents the path travelled by the listener from the starting position until he reaches the target object of each round in progressive grayscale color tones, from light gray to black (Figure 4). By pressing the “Playback” button, the experimenter can see in real-time the performance of the player and simultaneously hear in the headphones the 3D sounds that the user perceived while navigating from the starting position to the target sound source. Furthermore, the Game Analyzer interface displays the mean values of parameters P1-P4 for each round and level, as well as for all the levels of the game. The “Export all data” button allows the statistical information to be exported to an Excel file, together with the player's personal information, such as name, age, sex and the extent of the visual impairment condition.

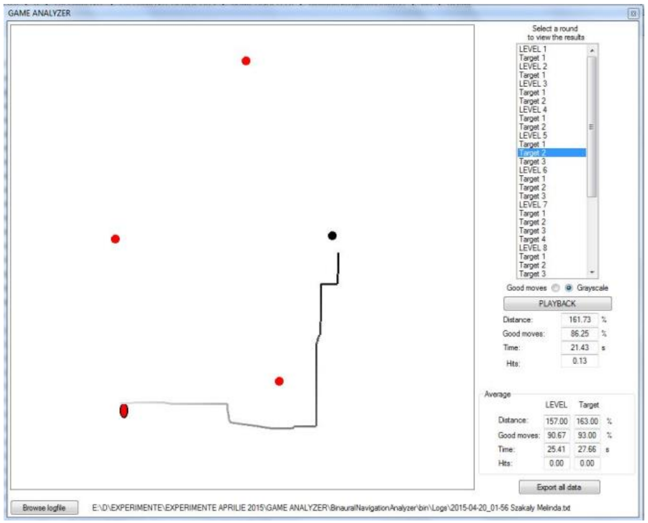

Figure 4 The "Grayscale" visualization option

This figure presents the player's performance for the second round of the 5th level of the game, using the "Grayscale" visualization option which paints the path travelled by the user in progressive grayscale tones

The purpose of the Game Analyzer application is to evaluate the game playing performances of the users through a complete statistical analysis (the mean values of the four studied parameters, distributed for each round, level and for all the levels of the game), as well as to provide a deeper understanding of the issues the subjects are confronted with during gameplay – the perception of 3D sounds, sound localization accuracy, front-back confusions, auditory-based spatial navigation and the interaction modality.

The Analyzer has been used as evaluation tool in another sound localization experiment, where the users were asked to identify the location of a hidden sound source by navigating freely from the starting point to the location of the target (randomly generated on the margin of a circle of fixed radius). In that case, the sound stimulus was a train of continuous white noise filtered with non-individualized HRTFs. The Analyzer instrument helped us to thoroughly investigate the occurrence of front-back confusions and the difficulties the players encounter when they have to discriminate the direction of the sources located in the front and in the back (Figure 5). These observations conducted to the idea of combining two types of noises (white and pink noise) with different spectral profiles in varying proportions, according to the direction of the sound source in space. The purpose of this sonification approach is to offer the listeners a more accurate perception of the sound sources originating from the front and from the rear hemifield.

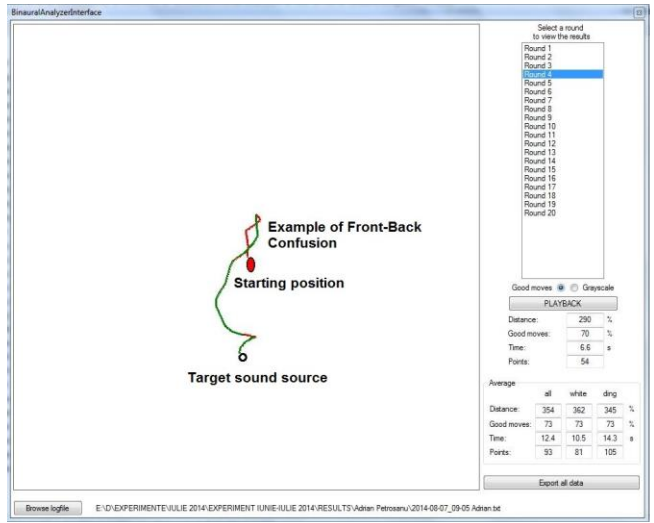

Figure 5 Example of front-back confusion

This figure shows that the listener, instead of going down to the target sound source, makes a front-back confusion and navigates initially up and then returns back on the correct path to the target

In order to play the Binaural Navigation Game or to operate with the Game Editor and with the Binaural Navigation Analyzer, the user is required to previously install the Csound kit [9] and the Microsoft .NET Framework.

CONCLUSIONS

To sum up, this paper provides a contribution to the design of navigational 3D audio games by presenting the development of the Binaural Navigation Game and that of its adjacent tools – the Game Editor, which allows the experimenter to simply and interactively define new sets of levels and the Binaural Game Analyzer, which offers statistical and real-time visual and auditory playback of the players' performances. The sound stimuli used were the combination of white and pink noise in varying proportions, according to the direction of the source in space that encodes the presence of target objects and the broadband alarm sound corresponding to the obstacles' positions. Both types of sounds have been convolved with the pairs of non-individualized HRTFs for the left and for the right ear, creating 3D binaural sounds delivered via headphones. The directional cues offered by the 3D binaural stimuli, together with the inversely proportional sound intensity encoding of distance aim to help the listener to easier identify the location of the auditory items (to reach the target obstacles and to avoid the blocking obstacles), to improve navigation and to offer the player a complete perception of the auditory environment. Other important sound cues are the auditory icons and the earcons that create a solid mapping between the acoustic perception and the events and situations from the game, enhancing thus the sense of presence and the level of player immersion.

Future plans involve the development of a portable version of the game, available on mobile platforms, that would have an additional haptic and vibro-tactile component.

ACKNOWLEDGMENTS

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

REFERENCES

1. Bălan, O., Moldoveanu, A., Moldoveanu, F. Navigational Audio Games: An Effective Approach Towards Improving Spatial Contextual Learning for Blind People. The International Journal on Disability and Human Development, 14, 2 (2015), 109-118.
2. Bălan, O., Moldoveanu, A., Moldoveanu, F., Dascălu, M.I. Audio Games- A Novel Approach Towards Effective Learning in the Case of Visually-Impaired People. In Proceedings of the ICERI Conference, Seville, Spain, (2014), 6542-6548.
3. Bilinski, P., Ahrens, J., Thomas, M. R. P., Tashev, I., Plata, J. C. HRTF magnitude synthesis via sparse representation of anthropometric features. In Proc. IEEE ICASSP, Florence, Italy, (2014), 4501-4505.
4. Brieger, S.W. Sound Hunter developing a navigational HRTF-based audio game for people with visual impairments. In Proceedings of the Sound and Music Computing Conference 2013, Stockholm, Sweden (2013), 245-252
5. C# Programming Language, https://msdn.microsoft.com/en-us/library/67ef8sbd.aspx
6. Connors, E.C., Chrastil, R.E., Sanchez, J., Merabet, L.B. Action Video Game Play and Transfer of Navigation and Spatial Cognition Skills in Adolescents Who Are Blind. Frontiers in Human Neuroscience (Internet) (2014).
7. Csapó, Á., Wersényi, Gy. Overview of auditory representations in human-machine interfaces. Journal ACM Computing Surveys (CSUR), 46, 2(2013).
8. Csound Manual, http://www.csounds.com/manual/html/hrtfmove2.html
9. Csound Programming Language, http://www.csounds.com/downloads/
10. Friberg, J., Gärdefors, D. Audio games: New perspectives on game audio. Advances in Computer Entertainment Technology, 5, 3 (2004), 1-7.
11. Hill, P.A., Nelson, P.A., Kirkeby, O. Resolution of Front-Back Confusion in Virtual Acoustic Imaging Systems. Journal of the Acoustical Society of America, 108, 6 (2000).
12. Lazzarini, V., Carty, B. New Csound Opcodes for Binaural Processing. In Proc. 6th Linux Audio Conference, Cologne, Germany, (2008), 28-35
13. Letowski, T., Letowski, S. Localization Error: Accuracy and Precision of Auditory Localization (Chapter 4) in Advances in Sound Localization, 2011.
14. Merabet, L.B., Sanchez, J. Audio-Based Navigation Using Virtual Environments: Combining Technology and Neuroscience. AER Journal: Research and Practice in Visual Impairment and Blindness, 2, 3(2009),128-137
15. Meshram, A., Mehra, R., Yang, H., Dunn, E., Frahm, J.M., Manocha, D. P-HRTF: Efficient personalized HRTF computation for high-fidelity spatial sound. In Proceedings of the International Symposium on Mixed and Augmented Reality, (2014).
16. Microsoft Visual Studio, https://www.visualstudio.com/
17. MIT HRTF Database, http://sound.media.mit.edu/resources/KEMAR.html.
18. Sanchez, J., Lumbreras, M. Usability and Cognitive Impact of the Interaction with 3D Virtual Interactive Acoustic Environments by Blind Children. In Proc. 3rd Intl Conf. Disability, Virtual Reality & Assoc. Tech., Alghero, Italy (2000).
19. Spagnol, S., Geronazzo, M., Rocchesso, D., Avanzini, F. Extraction of Pinna Features for Customized Binaural Audio Delivery on Mobile Devices. In Proceedings of International Conference on Advances in Mobile Computing, (2013).
20. Wallach, H. The Role of Head Movements and Vestibular and Visual Cues in Sound Localization. Journal of Experimental Psychology, 27, 4(1940), 339-368.
21. Westin, T. Game Accessibility Case Study: Terraformers- a real-time 3D Graphics Game. In Proc. of the Fifth International Conference on Disability, Virtual Reality and Associated Technologies, Oxford, UK, (2004).
22. Wightman, F.L. Resolution of Front-Back Ambiguity in Spatial Hearing by Listener and Source Movement, J Acoust Soc Am, 105, 5 (1999), 2841-53.