DESIGNING 3D AUDIO AND HAPTIC INTERFACES FOR TRAINING THE SOUND LOCALIZATION ABILITY OF THE VISUALLY IMPAIRED PEOPLE

BY

OANA BĂLAN* and ALEXANDRU BUTEAN

University POLITEHNICA of Bucharest
Faculty of Automatic Control and Computers

Received: December 2, 2014
Accepted for publication: December 20, 2014

Abstract. An assistive device, based on an alternate sensory modality, such as hearing or touch, can help the blind individuals to gain spatial awareness, to enhance their navigational skills and to improve their life quality. The purpose of our research is to develop a navigational system that would enable the visually impaired people to travel and orient in space by substituting the visual sense with relevant auditory information and haptic cues. As the proposed visual-substitution system will employ 3D binaural sounds synthesized from non-individualized Head Related Transfer Functions which offer an ambiguous spatial acoustic perception, we identified the need for training the visually impaired subjects’ sound localization abilities through a perceptual feedback based learning approach. This paper outlines the design and implementation of the 3D audio and haptic interfaces that we used for training and testing the spatial acoustic perception of the visually-impaired subjects who participated in our experiment.

Key words: 3D sound; training; HRTF; virtual auditory display; haptic; blind people.

2010 Mathematics Subject Classification: 00A06.

*Corresponding author; e-mail: oanab_2005@yahoo.com
1. Introduction

Worldwide, more than 285 million people suffer from a certain degree of visual impairment, of which 40 million are legally blind (World Health Organization Statistics, 2014). As a result, they need an assistive device that would help them to navigate in unfamiliar settings by improving their orientation and mobility skills (Bălan et al., 2013). This study is part of a more complex research project which has as purpose the development of an assistive system for the visually impaired people that would replace sight with an alternative sensory modality, such as hearing and touch. The final prototype will employ 3D binaural sounds delivered through headphones, vibrations and other auditory or haptic cues. The 3D binaural sounds incorporate directional information that defines the position of the sound source both in the horizontal and in the vertical planes, offering the same acoustic perception as under free-field listening conditions. The 3D sounds are synthesized using the Head Related Transfer Functions (HRTF), a measure of the sound transformation from the source to the listener’s ears (Wallach, 1940; Wenzel et al., 1993; Wersényi, 2007; Wersényi, 2009). However, the HRTFs are highly dependent on the anatomical characteristics of the listener’s body (size and shape of the pinna, head and torso) (Ahveninen et al., 2014), so that the use of the same HRTFs for different listeners will conduct to an ambiguous acoustic perception and high localization errors (Wersényi, 2012). Due to the demanding and laborious process of recording individualized transfer functions for each listener apart, the majority of virtual auditory displays employ 3D sounds generated from non-individualized HRTFs (Parseihian & Katz, 2012), that lead to a less accurate sound localization performance which is reflected in a higher incidence of precision and reversal errors (Bălan et al., 2014). Our approach is oriented towards improving the blind people’s spatial hearing resolution through perceptual feedback based training, crossmodal sensorial adaptation and procedural learning (Mendonça, 2014; Ohuchi et al., 2006; Towers et al., 2012). This paper presents the design and implementation of the 3D audio interfaces that have been used for training and testing the visually impaired subjects’ sound localization performance in a virtual auditory environment.

2. The 3D Audio and Haptic Interface for the Visually Impaired People

Our 3D audio application, called Binaural Navigation, consists of two modules: first, Binaural Navigation Test, an auditory interface that allows the user to move freely (using the mouse movement/touchpad as the main method of interaction) in order to locate the position of the incoming 3D sound source and the second module, Binaural Navigation Analyzer, a visualization and statistical tool designed with the aim of interpreting and evaluating the results of the navigational tests (Bălan et al., 2015). For the training session, we designed
an auditory and haptic interface that offers crossmodal perceptual feedback (auditory, haptic and visual – for the users with a higher degree of residual vision) in what concerns the correct direction of the sound source. All our applications have been developed using the C# programming language and the CSound programming language for sound processing (Csound, 2015).

2.1. The Binaural Navigation Test

In the Binaural Navigation Test interface (Fig. 1), the user is required to introduce his name, age, sex and the number of years since he is suffering from a visual impairment. Consequently, he starts searching for the 3D sound source from a randomly generated position on the margin of a circle of 150 pixels radius. The source is located in the center of the circle, so that the distance the listener needs to travel is relatively short. The sonification approach is based on using 3D binaural sounds that have been synthesized with non-individualized HRTFs (from the MIT HRTF database) and on the inverse proportional sound intensity encoding of distance, so that the sound intensity increases as the user gets nearer to the source and decreases as he gets farther, until complete silence (outside the auditory area of 200 pixels). The sound localization test procedure consisted of 20 rounds, in which the sound stimuli were white and pink noise and a “ding”-like sound with short breaks between two consecutive bursts (Csapó & Wersényi, 2013). The personal information of the user, the date and time of the test and the results of each session were recorded and saved in a “log” file, in order to be further processed and analyzed by the Binaural Navigation Analyzer tool.

Fig. 1 – Binaural Navigation Test.
2.2. The Binaural Analyzer Interface

For evaluating the sound localization performance of the visually impaired subjects, we calculated the following parameters:

- The ratio of the distance travelled by the listener (from the randomly generated starting position on the margin of the circle to the location of the sound source in the center of the circle) to the minimum possible distance of 150 pixels (the radius of the circle).
- The percentage of good movements (movements that the user performs towards the sound source, minimizing the distance between his current position and the location of the sound source).
- The number of mouse movements.
- The round completion time (in seconds).

By pressing the “Browse log file” button from the Binaural Analyzer Interface (Fig. 2), the evaluator can load the “log” files containing the results of the sound localization test sessions for each subject apart. The application allows real-time visualization and audio playback of the performance of the users for each round. The segments which are considered as “good moves” are colored in green, while those which correspond to “wrong moves” are colored in red; an alternative option paints the path travelled by the user in progressive grayscale colors, from the starting position to the target. The interface also displays the mean values for all the 4 studied parameters, for the 10 rounds that used white/pink noise and the 10 rounds that used the “ding” sound, as well as for all the 20 rounds. Finally, the “Export all data” button allows the data to be saved to an Excel file in order to be assessed for further analysis (Bălan et al., 2015).

Fig. 2 – Binaural Navigation Analyzer.
2.3. The Training Interface

The training interface has the following three modules:

a) A free listening module, where the user can move the mouse cursor inside a circle and listen to the 3D sound that corresponds to the angle between the center of the circle (considered as his fixed position) and the location pointed inside the circle, for both types of sounds (Moldoveanu & Bălan, 2014).

b) A sound discrimination module, where the listener is presented 4 sound sources (respectively, 8 in the second round) that are displayed on the screen simultaneously, as small circles of 5 pixels radius. The 3D sounds corresponding to the target sources are played sequentially and the listener is required to indicate the direction of the current audio stimulus by clicking on the small circle assigned to that position. When he makes a correct choice, the source disappears, narrowing thus the searching range for the other stimuli. The sound discrimination procedure is, nevertheless, applicable only to the visually impaired subjects who benefit from a higher degree of residual vision (more than 20%).

c) A sound localization module, where the subjects with residual vision are required to indicate the perceived direction of the incoming 3D audio stimulus by clicking inside a circle (considering that their position is in the center) and where the blind participants are asked to indicate the location of the sound source by using the conventional notation of the hour hand of the clock (for instance, 90 degrees to the right corresponds to 3 o’clock or 180 degrees in the back corresponds to 6 o’clock). The perceptual feedback consists in a series of vibrations that the subjects receives on a haptic belt placed on the scalp (consisting of 12 vibration motors placed at 30 degrees distance around the head, which communicate with the computer through a Bluetooth interface), graphical feedback that is displayed on the screen (the user’s choice is colored in red, while the correct direction of the sound is painted in green) and a continuous replay of the 3D sound stimulus (delivered through headphones) (Fig. 3), in order to facilitate a crossmodal sensory association that would promote an effective adaptation to the perception of 3D sounds synthesized from non-individualized HRTFs (Fig. 4).

![Fig. 3 – Visual feedback.](image-url)
2.4. Results of the Sound Localization Experiment

We performed a sound localization experiment in which we tested the spatial acoustic resolution of nine visually impaired subjects, before and after the training procedure based on perceptual feedback learning. The results of our tests demonstrated that the visually impaired people improved their sound localization performance, enhanced their navigational skills and their ability to orient in space based on listening to 3D sounds synthesized from non-individualized HRTFs as the main auditory modality (Figs. 5 and 6). Moreover, they recorded a decrease in the angular precision error and a lower incidence of front-back confusions (situation when the listener perceives a sound coming from the front as originating from the back and vice-versa). In a usability study that took place after the experiment, the subjects unanimously considered that the training procedure helped them to improve their spatial acoustic perception and that the presented approach is a useful strategy for the future development of an assistive device for the visually impaired people. Furthermore, they suggested the integration of the sonification technique (based on 3D binaural sounds and on the inverse proportional encoding of distance) into navigational audio games and serious games as a novel, entertaining and challenging modality for improving the spatial acoustic performance in a virtual auditory environment.
Fig. 5 – Evolution of the percent of front-back confusions between the 2 days of training.

Fig. 6 – Evolution of the sound localization error between the 2 days of training.
3. Conclusions

In conclusion, our 3D audio interfaces (the Binaural Navigation Test, Binaural Navigation Analyzer, the sound discrimination and training applications) and the haptic tool (the haptic belt) have been successfully used for testing, training and evaluating the sound localization performance and general spatial perception of the visually impaired people. The results we obtained demonstrate the efficiency of our method for training the acoustic resolution of the blind subjects, considering the future development of an assistive device aimed to provide a rich and complex auditory representation of the environment.

Our research will continue by designing a navigational audio game (with hierarchical levels of difficulty), using the same sonification method as in our previous applications. In addition to this, we will employ other auditory dimensions (such as pitch or timbre), auditory icons, earcons and speech, in order to encode the visual information into acoustic cues and to ensure a higher level of immersion and flow in the game. Moreover, we intend to increase the duration of the training session and to make some non-invasive BCI EEG measurements for studying the subjects’ neuroplasticity during the practice of the game. The results that will be obtained will be integrated into a training procedure for using the auditory sensory substitution device which will be developed in the next 2 years.

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.

REFERENCES

O problemă deosebită care afectează viața persoanelor cu deficiențe de vedere este sedentarismul, lipsa de activitate și de implicare socială. Un dispozitiv asistiv, bazat pe alte simțuri, cum ar fi auzul și simțul tactil, poate ajuta persoanele nevăzătoare să își dezvolte abilitățile de navigare prin spațiu și să își îmbunătățească calitatea vieții. Scopul acestui studiu este de a dezvolta un sistem de navigare care să le permită acestora să se orienteze în spațiu, prin substituirea stimulilor vizuali cu informații transmise prin intermediul canalelor auditiv și tactil. Pentru a dezvolta un astfel de sistem, trebuie să luăm în considerare capacitatea de localizare a sunetelor și performanța spațial-acoustică a utilizatorilor. Deoarece sistemul asistiv propus de noi se va baza pe sunete binaurale 3D obținute prin filtrarea cu funcții HRTF neindividualizate.
(care oferă o percepție acustică ambiguă), am identificat nevoia de a antrena capacitatea de localizare a sunetelor printr-o metodă de training bazată pe feedback perceptual, auditiv, vizual și tactil. Eficiența metodei propuse a fost demonstrată într-un experiment pentru testarea și antrenarea abilităților de localizare a sunetelor, la care au participat 9 persoane cu deficiențe de vedere. Subiecții au primit feedback auditiv (sunetul 3D era percut în căști stereo) și haptic (simțeau o serie de vibrații pe cap, provenite de la benta haptică pe care o purtau pe scalp - aceste vibrații erau resimțite din direcția corectă din care provenea sunetul). În acest fel, nevăzătorii au putut să creeze o asociere eficientă între percepția stimulului auditiv și direcția din care acesta provine, care a condus la rezultate foarte bune în ceea ce privește capacitatea de localizare a sunetului, reflectate într-o precizie de localizare mai mare și o reducere a incidenței erorilor de tipul front-back confusion (confuzii față-spate).