Proceedings of the IMProVe 2011
International conference on Innovative Methods in Product Design
June 15th – 17th, 2011, Venice, Italy

Virtual acoustic environment interface for sound source localization training

L. Dunai [(a)], B. Defez [(a)], V. Santiago Praderas [(a)], N. Ortigosa Araque [(a)], G. Peris-Fajarnés [(a)]

[(a)] Centro de Investigación en Tecnologías Gráficas, Universidad Politécnica de Valencia, Spain

Article Information

Keywords:
Virtual sounds

Corresponding author:
Larisa Dunai
e-mail: ladu@upv.es
Address: Camino de Vera s/n, 8L, Valencia, 46022, Spain

Abstract

In the paper, a virtual acoustic environment interface for sound source localization training is presented. It is shown that, by means of a 3D virtual environment and 3D sound sources, the subject is able to interact and learn to localize virtual sound sources. This can constitute the basis of a potential navigation device able to be used in the future by blind people.

Many different interfaces have been developed for sound source localization; however, there are no virtual interfaces where the user can learn to localize spatial sound sources with great precision, through interactive virtual environments and stereo headphones. During these last decades, researchers worked hardly in the development of virtual auditory spaces. The idea of a virtual acoustic space is based on the fact that the acoustic sounds, which are presented to the listener through headphones, are perceived by him as coming from the free field. This phenomenon is also known as sound source lateralization. Since its beginnings, the development of virtual sounds has been mainly based on the linear system analysis in the frequency domain, with the Fourier Transform, named Head Related Transfer Function (HRTF). The sounds reproduced by headphones appear to be originated within the head. Due to the human hearing system, humans are able to externalize the head-originated sounds as if they were coming from the surrounding environment.

1 Introduction

With the development of new technologies and introduction of computers in the experimental trials, the demand of developing new and wearable Graphical User Interfaces as well as new evaluation methods continuously increases [1]. Thanks to the use of Graphical User Interfaces and software packages during the experimental trials, psychoacoustic experiments have led to better and more precise results in terms of analysis and representation. The present paper describes the development of a simple Graphical User Interface (GUI) for acoustical experiments, which enables the subjects to visually interpret the location of the virtual sounds [2]. Usually each psychoacoustic experiment has its own GUI and software, but some experiments use software packages enabling the researcher to set up a large variety of experiments under a common platform.

It is well known that humans have abilities to perceive the surrounding environment through the vision and the hearing [3]. But what happens when one of these human senses is lost? When humans lose their vision, then they are not able to perceive their surrounding, are not able to communicate or travel without help, they lose their independence [4]. In this case, it becomes necessary to make use of other senses such as hearing or feeling. The same situation happens when the objective is to learn to localize virtual sounds. All normal hearing people may perceive when a car or machine is getting near or will hit him. They can also perceive with certain precision the direction and distance of the noise or surrounding sounds. But are humans able to perceive with great precision (about centimetres) the position of a virtual sound? This task is difficult; a long training period is necessary so that the human gets used to the sound, associating it to a special spatial position. This paper describes the application of the GUI in the experiments for virtual sound source localization through headphones. An important factor in the design of a graphical interface for acoustical experiments is the resolution of the image, its precision and accuracy. Other important factor is the accuracy of sound reproduction. This factor depends on the selected software, which must deliver and process virtual sounds without modifying sound features, so that users can obtain a clear externalized sound. Another consideration in the design and use of a graphical interface for the acoustic experiments is the method employed for spatial sound reproduction, which must imply a simple solution for representing the sound scene and answer method. To find and represent the sound centre and region, in the paper, cells are used. This method was previously used in order to eliminate some problems related to sound localization accuracy. The interface can be used for testing adults and children abilities. The interface was developed at the Research Centre in Graphic Technology at the Universitat Politecnica de Valencia based upon software Macromedia Flash 8.0.

2 Design of the virtual acoustic environment interface

After defining the process and the requirements of the study, the process of designing the virtual acoustic environment interface can be divided into 5 steps: step 1: sound generation; step 2: definition and design of the experimental procedure; step 3: running experiment; step

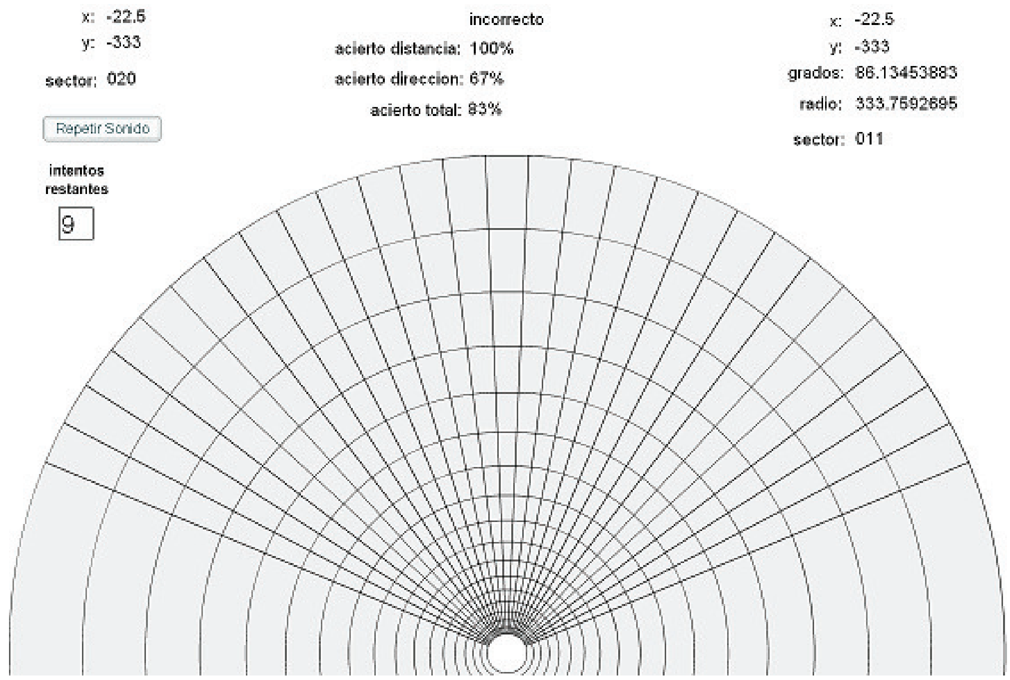

Fig. 2 Running experiment interface.

4: results presentation and step 5: analysis of the collected responses.

2.1 Interface design

A virtual acoustic environment interface is designed to provide a simple and intuitive tool for representing the sounds both in azimuth and in depth. The interface allows the user to perceive spatial judgements. At the same time, the interface is designed to be easy to be used. The actual virtual acoustic interface is capable of performing two sets of experiments: first, the anechoic sound experiments and, second, experiments with reverberation. Each set of experiments is based on three, in which three different sounds are used: wood, bongo and delta sound.

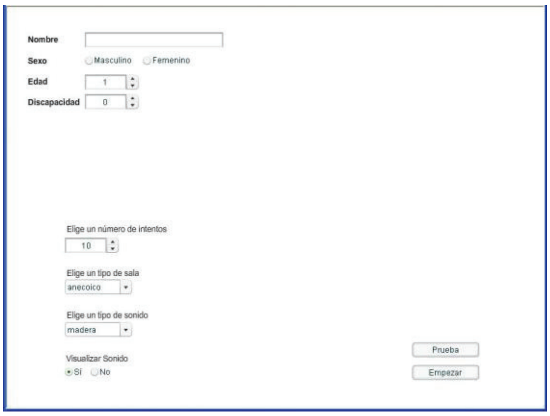

Fig. 1. Data introduction window.

The experimental interface consists of three interfaces: data introduction, running experiment and results presentation. Data introduction window (see Fig.1), allows to introduce subject details, to select experimental mode (anechoic sounds or with reverberation), the type of sound or the number of delivered sounds. Also the interface is designed for two types of experiments: training experiment and localization experiment. When selecting training experiment, the user is able to see the position of the listened sound. That is, when a sound is reproduced by the system, then the cell with the coordinates of the sound is coloured in green. In this way, the user is able to hear the sound and associate it with a location in the space.

In case of the localization experiment, the user listens the sound through the headphones, clicking on the cell of the interface in which he thinks that the delivered sound is located.

The running experimental interface includes a data presentation block and a graphical image of the environment (Fig.2). In the data presentation section, the coordinates of the mouse are presented at the upper left hand as well as a button for sound repetition and a window informing on the number of the remaining sounds. In the column in the middle, the accuracy results regarding the location of the precedent sound are shown. Finally, in the column on the right hand the coordinates selected are specified both in azimuth and in distance.

The Graphical image of the environment represents a horizontal plane of the frontal view, where the small circle represents the human eyes. The distance is presented in 17 divisions between 0,5m and 5m increasing in an exponential way and 26 divisions in azimuth between 55º at the left side and 55º at the right side of the human ears. Each division has 5º in azimuth. In this way, the environment is divided in cells.

Finally, results presentation interface is based on a table, in which sound localization accuracy in azimuth, distance and mean accuracy is presented in three columns for each individual trial.

2.2 Sound

Two different sounds are firstly recorded in order to represent real sounds. These sounds come from everyday life objects made from impacted wooden and bongo beam. A delta sound is generated with Adobe Audition software. Then, each sound is simplified based on additive synthesis techniques to resize these recorded sounds at a 44,1kHz sampling frequency and sound level of 72dB. All sounds are recorded and generated for both anechoic and with reverberation environments, by using binaural cues. Finally, each sound is convolved with a non-individual Head-Related Transfer Function (HRTF) using (1) (See Fig.3) [5].

$$y(n)=x(n)*h(n) \qquad (1)$$

where *x(n)* represents the sound with *n* samples and *h(n)* represents the Head-Related Transfer Function of *n* samples.

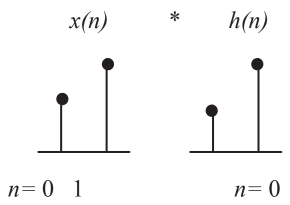

Fig. 3 Convolution representation of two sequences.

The non-individual HRTF's are generated with the CATT Acoustic application, which represents a computer model for room acoustic generation, convolution, auralization, etc. [6.]. In order to generate the HRTFs the Maximum Length Binary Sequences MLBS was applied and later cross correlation method was used between the system answer and MLBS input.

Each sound was processed using Matlab and saved as a Microsoft ".wav" file sampled at 44.1 kHz with 16-bit resolution.

3 Experimental procedure

After the sound has been created and the experimental procedure has been defined, subjects can perform the experiment. The users were seated in front of a computer where a graphical user interface is presented as in Fig.1.

Stimuli were presented through SENNHEISER HD201 headphones connected to a laptop using the experimental program developed in Flash 8.0. The program allows the reproduction of qualitative sounds. The subjects interacted with the experiment using the GUI represented in Fig.2.

This experiment was carried out in a conventional laboratory where external noises, such as human speech, traffic on the street, equipment noises, etc... are present.

The experiment is divided in two parts: the anechoic material sounds and material sounds with reverberation. Each part consisted of the three aforementioned sounds: wood, bongo and delta sounds.

In each test, the user listened a train of sounds in a specific spatial position, randomly selected by the program; he should estimate their position according to his perception. After the user had selected the response, the next train of sounds was reproduced. Each test was based on a set of ten sounds reproduced at different positions. The method employed was based on the forced selection; this is, the sounds are presented to the listener in different locations and the listener is required to answer whether the sound appears at the left, right, centre, near or far, by clicking on one cell of the user interface. Even if he does not know, he is required to provide an answer. In the experiment the duration of the sound was not taken into account. Participants could repeat the sound as many times as they wanted by pressing the repetition button. Then the listener responded by clicking with the mouse on one of the cells of the interface. The volume control was constant for all subjects. There was no feedback. A typical run lasted 5-8 min, depending on the listener.

All listeners performed one trial for each six types of sound consecutively. Initially, the listeners performed the experiments for each of the three types of sound individually for the anechoic environment and, later, for the reverberation environment. Provided that the program delivered only one material sound type per trial, it was assumed that this method did not affect participants' response. The sounds were delivered one by one in a randomized order for each stimuli and trial. After hearing the sound through the headphones, the listeners were asked to click on the cell from where they considered that the sound was coming from. Before answering, the subjects were allowed to repeat the stimulus as many times as they need by pressing the Repeat Sound button.

4 Results

For a better analysis of the results, the program specifies for each response, the azimuth, distance, mean azimuth and mean distance and gives the average accuracy for sound position, saving all these data into a '.txt' file. The obtained results are analysed offline. The results are presented for each sound and trial.

Value, %	anechoic			reverberation		
	WOOD	BONGO	DELTA	WOOD	BONGO	DELTA
Max.	97,1	96,6	95	96,7	95,4	94,9
Mean	89,87	93,38	88,61	91,63	92,83	91,11
Min.	71,5	88,4	80,5	83	87,7	83,1
Mean sd.	8,8	3,1	5,8	4	2,5	3,5

Tab. 1 Mean and minimum values in percentage for all three rounds for 1) anechoic and 2) with reverberation sounds.

Tab.1 presents subject results for 1) anechoic sounds and 2) sounds with reverberation, for wood, bongo and delta sounds. In Tab.1 it can be observed that localization accuracy of all sounds is higher than 70%, and maximum standard deviation (N-1) on localization is 8.8% for wood anechoic sound, which is indeed a very low deviation.

The development of the experiments showed that the GUI plays an important role in virtual sound localization. Also, it was found that, even if a simple and poor graphical interface is used, subjects are able to associate this graphical interface with the real environment.

5 Conclusion

In the present paper an easy-to-use research virtual acoustic interface, which enables the development of acoustical experiments, has been developed. This interface allows researchers to test localization ability of the users using hearing and visual cues. The tests provide data which are consistent and acceptable in virtual sound localization task. A further study will investigate the design of a new three-dimensional interface, able to be adapted to sound localization task: It will allow subjects to localize sounds through a simulated real environment in which the objects will reproduce sounds, having the subjects to decide which object sounds and where the sound is coming from. The idea of the new interface is based on the idea of interactivity with the user, just as if the subject was playing a videogame.

References

[1] J. Laneau, B. Boets, M. Moonen, A. van Wieringen, J. Wouters. A flexible auditory research platform using acoustic or electric stimuly for adults and young children. Journal of Neuroscience Methods, 142, pp. 131-136, 2005

[2] J. Usher, W. Woszczyk. Design and testing of a graphical mapping tool for analysing spatial audio scenes. AES 24th International Conference on Multichannel Audio.

[3] Brungart D.S Nathaniel I., W. R. Raibiowitz. Auditory localization of nearby sources II. Localization of a broadband source. Journal of Acoustical Society of America 106 (4), pp.1956-1968

[4] World Blind Union, “White cane safety day,” World Blind Union Press release, October, Canada, 2009.

[5] L. Dunai, G. Peris Fajarnés, B. Defez Garcia, N. Ortigosa Araque, F. Brusola Simon, 2009, “Perception of the sound source position”, Acoustical Physics Journal, Vol. 55 (3), pp. 448-450

[6] L. Dunai, G. Peris Fajarnés, B. Defez Garcia, V. Santiago Praderas, 2010, “The influence of the Inter-click intervalo in moving sound source localization for navigation systems”, Acoustical Physics Journal, Vol. 56 (3), pp. 384-353.