ANALELE UNIVERSITĂȚII
"EFTIMIE MURGU" REȘIȚA
ANUL XVIII, NR. 2, 2011, ISSN 1453 - 7397

Tudor Marcel Culda, Dumitru Stanomir, Cristian Negrescu

The Influence of Nonindividual Hrtf in Binaural Reproduction Using Headphones

A 3D virtual surround sound environment can be created for a listener using headphone reproduction. For this purpose, a set of head related transfer functions are needed. These functions are convolved with audio signals of a sound source resulting in a 3D virtual spatialization of the source. Two of the most common problems in headphone reproduction are in-head-localization and front-back confusion. In the present paper we are concentrating on analysing the influence of a personal or not personal Hrtf use. For this task it has been measured and built a database with individual Hrtfs for more than one person and performed subjective audio listening tests. In these listening tests we have been evaluated the perfomance of the reduction of reversal confusion in localization when different nonindividual Hrtf in as compared to using individual Hrtf for different types of audio signals.

Keywords: *Hrtf, front-back confusion, 3D audio*

1. Introduction

The binaural technique states that if the acoustic pressure signals at the point of the eardrums of a human person are faithfully recorded and then correctly reproduced in the same place for another moment of time, then the whole audio experience of that person, from the moment of recording, is exactly reproduced for that person, both aspects regarding space and localization and also aspects regarding the tone color [1]. The sound signals on which the human auditory system will form the perception of the audio image, are the two sound pressures signals located at the eardrum position (binaural signals), inside each ear possessed by a human person.

For a natural hearing case under a free field condition, the sound signals arriving to the listener from a given direction will be exposed in their propagation path to a unique filtering. This phenomenon can be described through the next statement: "A sound wave coming from a given direction and distance, results in two sound pressures, one at each eardrum. The transmissions are described in terms of two transfer functions that include any linear distortions, such as coloration and interaural time and spectral differences [1]". The mentioned transfer functions are widely called, in the literature, "Hrtf – head related transfer functions" or "Hrir – head related impulse response" if we consider time domain description. These Hrtfs are dependent upon the angle of incidence and the distance to the sound source. If the distance is reasonably large (more than 2m [1]), the incidence wave can be approximated with a plane wave and the Hrtfs become independent of distance. Practical, the Hrtfs will describe the physical effects of the diffraction of sound waves determined by human torso, shoulders, head and ear's pinnae.

The knowledge of human Hrtfs is very important in developing applications that will synthesis binaural signals with the purpose of creating audio virtual spaces as a replica of reality [10], or for auralization in room modeling systems. If the reproduction chain is correct and recording and synthesis is carried out carefully, then true reproduction can be achieved [2].

One of the main problems in binaural reproduction is that there are anatomical differences between people that will lead to individual differences between Hrtfs, and in most applications all listeners must listen to the same binaural signals. It is expected that not all listeners will have the intended audio perception. During the last two decades there were many investigations regarding this topic [3]-[9]. Looking into the specific problems when using headphones for reproduction, we will find two widely spread issues, namely in-head-localization (IHL) and reversal errors (front-back, back-front) in localization. As a general tendency, using nonindividual Hrtfs will cause a worsening in localization accuracy with an increase in IHL and reversal errors. It was found that reverberation improves the externalization factor for a virtual sound source, and correlating head movements with the sound signals will improve as well the reversal error. Equalization of the playback chain is also very important in diminishing these errors [9].

2. Goal of the experiment

In this paper an evaluation of the reversal error is performed, based on a series of listening tests, correlating the dependence for the rate of error with the use of individual or nonindividual Hrtfs in sound source synthesis. Because a device for head tracking is mostly unknown to a home audio consumer, plus the fact that the application that integrate the head movement are not widely spread and can consume great processing power for real time synthesis, reversal errors remain one of the common errors in headphone reproduction.

It is intended in the present study research not to use correlation with head movements and as well not to use standard unnatural audio test signals like white noise burst or static harmonic sounds, but rather sounds produced by common musical instruments. Having this kind of signals, it was observed the behavior of human auditory system regarding reversal errors when the proprieties of the signals are not the fact that they are static in long term, but that they have a varying energy in a given frequency band.

For this purpose, audio samples from four categories of instruments were taken. The length of a sample was 3 seconds long, but the listener was able to repeat the sample until he could take a decision, but he couldn't compare one situation with another. Every location synthesized was judged separate. The listener was unaware of the true position until the whole experiment was finished. The reproduction chain was calibrated for each person individually. This process is ensuring that the error rate difference is just due a Hrtf difference between individuals.

The four categories of instruments were: a bass guitar category with the main energy varying under 300 Hz, an acoustic guitar strumming filtered on 300-8000Hz band, a shaker having the main energy variation above 8000 Hz and a percution category that simulate the impulse like signals.

3. Experimental procedure

For the proposed investigation issue, first of all, a human Hrtf database needed to be measured, so we could compare the reversal confusion error rate when individual Hrtfs versus nonindividual Hrtfs are used for reproduction. Six otologically normal persons were measured in whole horizontal plane with a 10° resolution step. Before playback, the reproduction chain was calibrated as described in section 3.2. For the reproduction of the signals, a pair of Sennheiser Hd 485 headphones were used. All the processing of the audio signals was done on a computer using mainly a Matlab environment. After the acquisition of the Hrtfs, a virtual audio space was simulated using the measured Hrtfs. For this purpose a "clean" sound signal was convolved with a pair of Hrtfs and a pair of headphone calibration filters (depending on the listener) and therefore was simulated a natural sound arriving from a given direction in a free field condition. Random positions unknown to the listener were generated for each signal type and listening test were performed using individual and nonindividual Hrtf synthesis. Each listener had to mark on a chart the perceived location for the simulated sound source. The training with the binaural signals was minimal and none of the interviewed persons had any experience of this kind, except one of the authors who was involved in the experiments. Even so, the simulated sources were played in random order so the judgment for the localization was based only on the personal auditory system. For each type of signals, from one of the proposed categories, random direction were simulated using personal and not personal Hrtfs, and the reversal errors from the listening test were marked and counted.

3.1. Hrtf measurement

Ideally, to measure the impulse response of a system, an impulse signal should be provided to the input and then to measure the output. Because creating an impulse like signal with the electroacoustic transductors is a difficult task, other methods must be investigated. An exponential sine sweep method [11] was used in these present research measurements. The swept band was 80-20000 Hz. The sound source (loudspeaker) was placed at the distance of 2 m from the human person and a pair of signals were recorded, one inside each ear. The measurements were done in an empty room with reflection due the walls, but the reflections were truncated out from the recordings, leaving only the direct path as shown for an example in Figure 1.

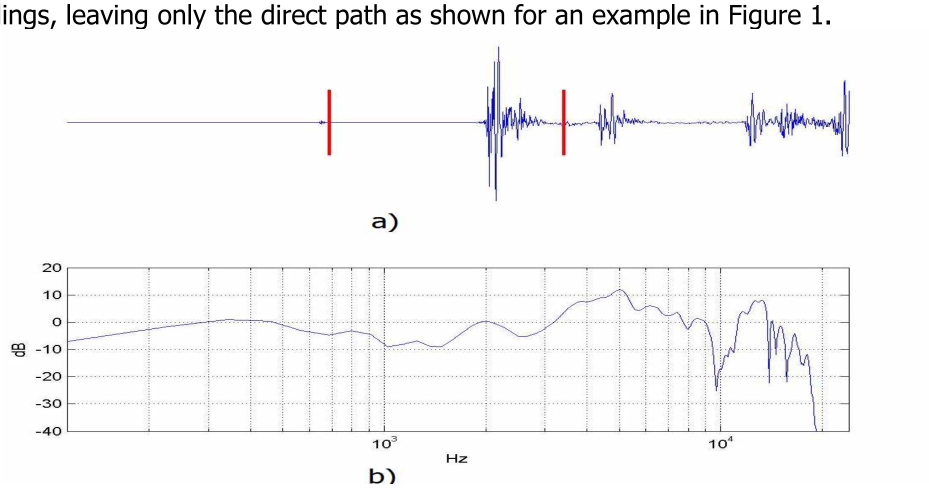

Figure 1. a) Example of removing the reflection from a room Hrtf measurement, the useful signal is located between the red lines b) example of a measured Hrtf (frequency domain)

The person was sitting on a rotating chair and after the measurement for one direction was made, the chair was rotated with a 10° step resolution. The microphones for capturing the sound field from the outer were placed at the blocked entrance of the ear canal. Some authors preferred measurements near the eardrum [5][12] but it rater complicated to fix the microphones in this positions and implies as well an injury risk. Hammershøi and Møller [13] [1] performed a thorough research on different location were on can place the recording microphones inside the ear canal and still to be able to acquire all the spatial information. Their research concludes that any point from the ear canal, including the entrance, even the blocked entrance will provide all the spatial information and correct eardrums signals reproduction with the condition to calibrate the headphones using the same position for the microphones as the position in which the Hrtf were measured. The blocked entrance of the canal position has some advantages: the injury risk is smaller, the microphones are not disturbing relevant acoustic field, they will be placed in the blockage, these recordings contain the fewest inter-individual differences comparing to other recording positions.

3.2. Calibration of the headphones

As clearly showed by Møller [1], for correct reproducing of eardrum signals we need to calibrate the headphones, and this is done by inserting a correction function. Analyzing in frequency domain, this function will be the inverse transfer function of the transmission from the voltage of the headphones terminals to the sound pressure recorded in the same position as the position of Hrtf measurements. If we use the same microphone for Hrtf measurements and for calibration, then the transfer function of the microphone will be canceled out. For the equalizing filter we used a Wiener filtering approach [9] using orders of N=200-300. Also a frequency dependent regularization was used [14] so by inverting the transfer function, the spectral dips will not generate high amplification bands.

4. Results and discussion

The reversal error rate was noted for each individual and comparing charts were established. Looking at the overall reversal error chart (Figure 2) we can say that there is no noticeable difference between different kinds of instruments and also there is a very small difference between personal and not personal Hrtf use. This may due to the equalization of the reproduction chain.

More conclusions can be taken if we analyze the chars for each person separate (Figure 3). Here we can observe two groups of people. The first group seems to follow a pattern and the second group seems to have no pattern.

Looking into the first group in the section of personal Hrtf use, it can be noticed that the shaker instrument has the greatest error rate. This will represent the high frequency energy distribution instruments. Asked about externalization, both groups recognized sound been well externalized for personal Hrtf use.

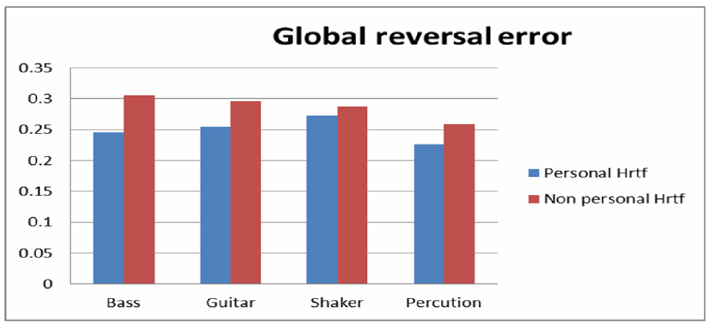

Figure 2. Global reversal error rates

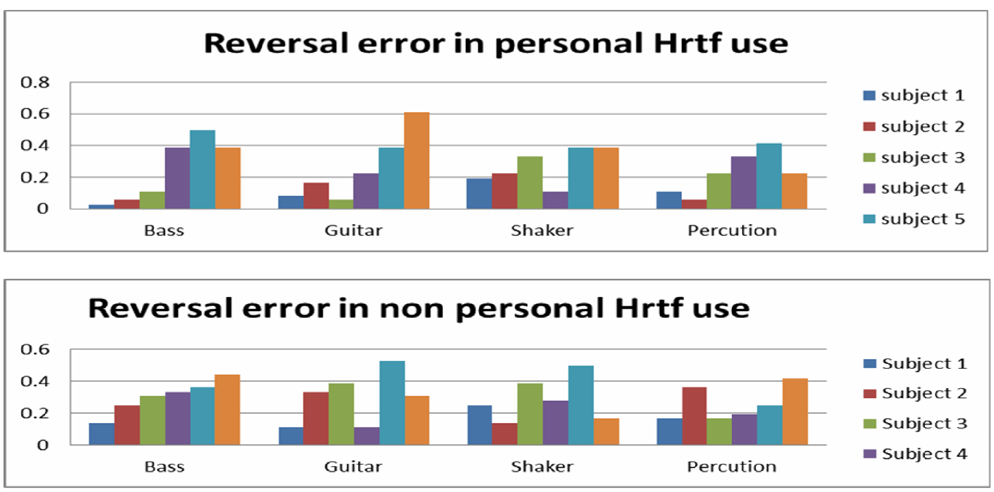

Figure 3. Individual reversal error rates

The overall conclusion was that the bass signals were clearly externalized even if the distance of perception was short (0.5-1 m), but the shaker was localized closer to the head and IHL sometimes occurred. It would be interesting to expand this experiment and search, if exists, for a correlation between the degree of externalization and the rate of reversal errors. For nonindividual Hrtf use the error rates were similar for all instruments and no tendency can be observed.

Looking into the next group it is hard to find any pattern for the error rates. These persons could form an exception due a mistake in reproduction chain or an unfocused condition during the experiments, but to be able to conclude this fact, we need to repeat the same experiment using a larger group of persons and observe again the size of the no pattern group.

5. Conclusions

Understanding and evaluating the errors in localization is very important in designing many types of applications that use an audio virtual reality and are destined to more than one consumer.

The reversal error for different types of musical instrument signals was evaluated using the criteria of individual or nonindividual Hrtf use. For this goal, a human Hrtf database was measured for more than one person. Individual calibration for the headphones was used and localization listening test were performed with previously measured persons. Different results were discussed.

References

[1] H. Møller, *Fundamentals of Binaural Technology*, Applied Acoustics, vol. 36, 1992, pp. 171-218.

[2] D. Hammershøi, *Fundamental Aspects of the Binaural Recording and Synthesis Techniques*, Denmark: AES 100th Convention, 1996.

[3] H. Møller et al., *Binaural Technique: Do We Need Individual Recordings?*, J. Audio Eng. Soc., vol. 44, Iunie. 1996, pp. 451-469.

[4] D. Hammershoi,J. Sandvad, *Binaural Auralization, Simulating Free Field Conditions by Headphones*, AES Convention:96 (February 1994), Amsterdam

[5] E.M. Wenzel, M. Arruda, D.J. Kistler, and F.L. Wightman, *Localization using nonindividualized head-related transfer functions*, J. Acoust. Soc. Am., vol. 94, 1993, pp. 111-123.

[6] Hartmann WM, Wittenberg A. *On the externalization of sound images*. J Acoust Soc Am. 1996 Jun; vol 99(6),pp:3678-88.
[7] N.Durlach et al. *On the externalization of auditory images*, Journal Presence: Teleoperators and Virtual Environments archive, Volume 1 Issue 2, Spring 1992
[8] D.Begault; A. Lee, E. Wenzel, M. Anderson, *Direct Comparison of the Impact of Head Tracking, Reverberation, and Individualized Head-Related Transfer Functions on the Spatial Perception of a Virtual Speech Source*, AES Convention:108 (February 2000), Paris
[9] Sang-Myeong Kim and Wonjae Choi, *On the externalization of virtual sound images in headphone reproduction: A Wiener filter approach*, J. Acoust. Soc. Am. Volume 117, Issue 6, pp. 3657-3665 (2005)
[10] Pulkki V. *Efficient spatial sound synthesis for virtual worlds,* Aes 35th International Conference, London, 2009.
[11] A. Farina, *Simultaneous measurement of impulse response and distorsion with a swept-sine technique,* Aes Convention 108, Feb 2000
[12] F.L. Wightman and D.J. Kistler, *Headphone simulation of free-field listening. I: Stimulus synthesis*, J. Acoust. Soc. Am., vol. 85, Feb. 1989, pp. 858-867.
[13] D. Hammershøi and H. Møller, *Sound transmission to and within the human ear canal*, J. Acoust. Soc. Am., vol. 100, iulie. 1996, pp. 408-427
[14] O. Kirkeby şi P. A. Nelson, *Digital Filter Design for Inversion Problems in Sound Reproduction*, J. Audio Eng. Soc, vol. 47,(7/8), p. 583--595, 1999.

Acknowledgement: The work has been funded by the Sectoral Operational Programme HumanResources Development 2007-2013 of the Romanian Ministry of Labour, Familyand Social Protection through the Financial Agreement POSDRU/6/1.5/S/19.

Addresses:

- Eng. Tudor Culda, "ETTI", Politehnica University of Bucharest, B-dul Iuliu Maniu 1-3, Sector 6, Bucharest, tudor_culda@yahoo.com
- Prof. Dr. Eng. Dumitru Stanomir "ETTI", Politehnica University of Bucharest, B-dul Iuliu Maniu 1-3, Sector 6, Bucharest
- Prof. Dr. Eng. Cristian Negrescu "ETTI", Politehnica University of Bucharest, B-dul Iuliu Maniu 1-3, Sector 6, Bucharest