Design and implementation of a 3D sound system based on the HRTF

Ivana Rossell
Robert Barti
Department d'Acústica, Enginyeria La Salle
ivana@salleURL.edu http://www..salleURL.edu

Abstract

Virtual Reality is one of the hottest buzzwords today in electronics industry. Until recently, the focus of these efforts was in providing stereoscopic three-dimensional graphics to stimulate our sense of vision. However, to create believable worlds, visual cues are not sufficient, the field of auralization, -three dimensional sound- is necessary. There are a number of motivations for developing auralization systems in addition to the advent of virtual reality. The three dimensional processing adds an auxiliary creative element to be manipulated by the commercial musician or record producer, providing a new realm of entertainment to explore. In this paper, a tridimensional sound system for performing static sound source placement with headphone playback using the filtering by HRTF is developed.

1 Introduction

Research on the acoustical and psychological bases of human sound localization suggests that the primary acoustical cues used to determine sound source position are the interaural differences in time of arrival (ΔT) of a sound wave at a listener's two ears, the interaural differences in overall intensity (ΔI) of the sound, and the position-dependent filtering caused by the interaction of an incoming sound wave with the folds of the pinnae. In the past quarter century, many studies have shown that cues provided by pinnae filtering are more important than previously believed, especially for localizing sounds on the median plane (where ΔT and ΔI cues are minimized), and for establishing the "externalized " (out of head) character of sounds in the natural environment.

In this paper we will describe how we obtain this position dependent filter called HRTF (Head Related Transfer Function). This filter will be used, convolved with a anechoically signal, to create the impression (auditive event) of receiving a this signal from a specific direction.

2 The HRTF

The first phase of our synthesis procedure involves measurement of the Head Related Transfer Function, HRTF. We can understand this transfer function as the filter introduced by a subject to the sound incomming to the eardrum in each ear. So there is a pair of filters for every position of a sound source in the space.

2.1 Method

Our approach is based on linear filtering principles. Let $x_1(t)$ represent an electrical signal that drives a loudspeaker in free field, and let $y_1(t)$ represent the resultant electrical signal from a probe microphone positioned at a listener's eardrum. Similary, let $x_2(t)$ represent an electrical signal that drives a headphone, with $y_2(t)$ the resultant microphone response.

The design of the appropiate filter is best described in the frequency domain.Thus X_1 is the Fourier Transform of the x_1 (t) and so forth. The signal coming from the loudspeaker at a specific position at the eardrum can be written as:

$$Y_1 = X_1 \cdot H_s \cdot HRTF \quad (1)$$

And the signal at the probe microphone from the headphone can be written as:

$$Y_2 = X_2 \cdot H_h \quad (2)$$

Where Hs is the loudspeaker transfer function. And Hh is the headphone transfer function.

Thus, if $x_1(t)$ is equal to $x_2(t)$, using both equations we obtain the expression that we will give us the HRTF:

$$HRTF = Y_1 \cdot H_h / Y_2 \cdot H_s \quad (3)$$

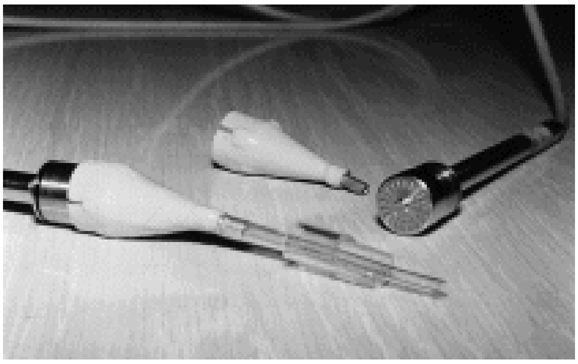
This transfer function will be measured for a large number of sound source positions with different subjects.

3 HRTF measurement

A wideband noise (MLS sequence) was presented (either by loadspeaker or headphones) repetitively, and the response at the listener's eardrum was obtained by averaging the output of the probe microphone.

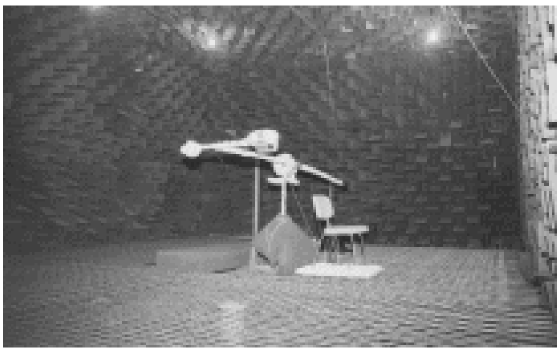
The acoustical response at the eardrum was measured with a condenser microphone (1/2 inch) with a conical copper cavity coupled to a flexible probe tube with an outer diameter of 2.5mm. The microphones used were identical in sensibility and phase response.

The photography shows the system used to collect the signal.



For free-field measurements, the wideband signal was transduced by one loudspeaker positioned at 1.5m from the subject in the anechoic chamber. Next picture shows the system:

The elevation swept was made by the rotatory axis in wich the bar was subjected. The azimuthal swept was made by moving the chair where the subject was sited to make the measurement. The accuracy of the positions was not very good but enough to obtain our first results.



4 Receiver model

Fist measurements were made on human subjects. In a second phase of our work we decided to "built" a head and torso simulator (HATS).

We use a comercial dummy and we made an intent to make it suitable for our experiments: The face and head of the dummy were covered by latex, the interior was filled with absorbent material and the ears were perforated to introduce the microphone.

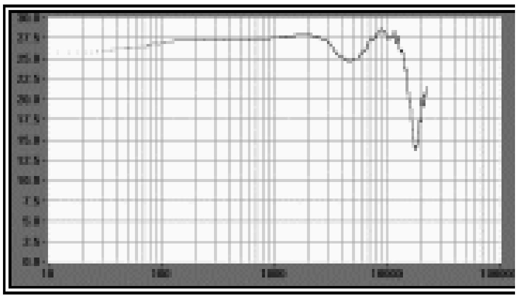
Next pictures show our HATS:

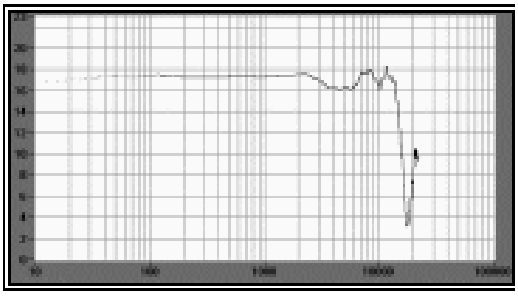
The receiver model allowed us to obtain standard HRTFs in spite of the individual transfer functions calculated before.

5 Results and discussion

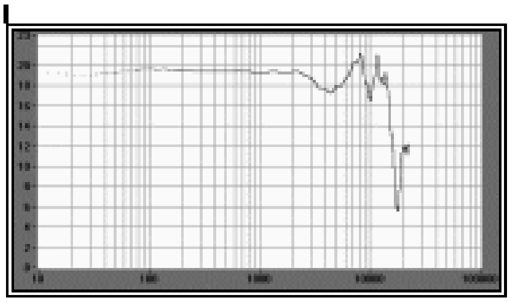
The HRTFs obtained were satisfactory and provided a quite good localization performance.

Next figure shows a pair of –40 degree elevation and 135 degree azimuth HRTF mesured.





Next pictures show a comparison of a 0 degree azimuth and 72 degree elevation HRTF measured using different subjects. The localization performance was better when listening through one’s own HRTF measurement.



The results of filtering these pairs of HRTFs with anechoically recorded mono signal and presenting the resulting sound over headphones, create an impression of perceiving the sound from the position measurement.

The accuracy of the auditory event depended on the nature of the signal to convolve. Broadband signals, impulsive and complex signals provided more satisfactory results.

6 New system

At present, we are finishing a new system to create a 3D Sound space in the anechoic chamber in order to minimize measurement variability.

In our system we have 19 loudspeakers mounted on a semicircular arc (3m diameter) the ends of wich were attached directly above and directly below the subject. We have a sound source every 10º. The subject or the receiver model (HATS) is seated on a rotary and adjustable chair. It has controlled movements in 10º steps. Thus, we can obtain in further works all the HRTFs of the whole sphere with a 10º accuracy.

Nexts photographies show the whole system in the anechoic room.

The tracking control, to avoid head movements during the measurements, is made with a laser beam positioned and fixed on the head of the subject. The laser beam has to aim at a target for any azimuth position.

7 Conclusions

The system has given satisfactory results and good 3D sound simulations. The nature of the signal to be convolved with the HRTF’s has been very important

in the impression created when it was presented under headphones. Presenting some signals continuosly has given movement sensation creating more spectacular simulations. The *inside the head* effect produced by the headphones was a problem observed in the sound presented to the subjects.

Future lines to develop in further works could be:

- An implementation of a crosstalk canceller to play the 3D sound trhough speakears. We are working in this issue just now with the simple crosstalk canceller and,after this first step, next will be the implementation of the room effect canceller.

- Another work we are implementing in our department is the obtention of the BRIR (Binaural Room Impulse Response). We use in this project the same HATS we have shown in this paper.

- As we have mentioned before, we will work with any sound localization experience, in order to improve reseach in binaural effects, in human sound perception.

8 Acknowledgments

This work has been the first step to develop a new workline in our department. We are now finishing our system to obtain better HRTFs and to experiment in all kind of binaural experiences. We are going to work with human sound localization and to achieve experiments to examine auditory perceptions (following Blauert's steps). We are working too in designing an auralization system using BRIRs (Binaural Room Impulse Response) obtained with our HATS.

References

[1] Frederic L. Wightman i Doris J. Kistler, "Headphone simulation of free-field listening. I: Stimulus synthesis". J. Acoust. Soc. Am. 85 (2), febrer de 1989.

[2] Frederic L. Wightman i Doris J. Kistler, "Headphone simulation of free-field listening. II: Psychophysical validation". J. Acoust. Soc. Am. 85 (2), febrer de 1989.

[3] Henrik Moller, Michael Friis Sorensen, Clemen Boje Jensen i Dorte Hammershoi, " Transfer Characteristics of Headphones Measurements on Human Ears".J. Audio Eng. Soc., Vol 43, Num.4, abril de 1995.

[4] Henrik Moller, Michael Friis Sorensen, Clemen Boje Jensen i Dorte Hammershoi, "Head-Related Transfer Functions of Human Subjects". J. Audio Eng. Soc., Vol 43, Num.5, maig de 1995.

[5] Frank Filipanits Jr, "Design and Implementation of an Auralization System with a Spectrum-Based Temporal Processing Optimization".

[6] Keith Dana Martin, "A Computational Model of Spatial Hearing". Master of Science in Electrical Engineering at the Massatchusets Institute of Technology.

[7] Douglas D.Rife i John Vanderkgoy, "Transfer-function measurement with maximum_length sequences".J. Audio Eng. Soc., Vol 37, Num.6, juny de 1989.

[8] Bill Cardner and Keith Martin, "HRTF Measurements of a KEMAR Dummy-Head Microphone".MITMedia Lab, maig de 1994.