Precise measurement of HRTFs for better auralization

Møller, Henrik; Hammershøi, Dorte

Published in:
Proceedings of International Symposium on Simulation, Visualization and Auralization for Acoustic Research and Education, ASVA 97, April 2-4, 1997, Tokyo, Japan

Publication date:
1997

International Symposium on Simulation, Visualization and Auralization for Acoustic Research and Education

ASVA 97

2-4 April 1997 Tokyo, Japan

PRECISE MEASUREMENT OF HRTFs FOR BETTER AURALIZATION

Henrik Møller, Dorte Hammershøi
Acoustics Laboratory, Aalborg University
Fredrik Bajers Vej 7 B4
DK-9220 Aalborg Ø, Denmark
e-mail: acoustics@kom.auc.dk

This paper presents the method of binaural auralization. Special attention is given to the determination of the head-related transfer functions, and how this affects the further processing.

A brief introduction to spatial hearing is given, followed by a division of the sound transmission to the eardrum. It is then described how the binaural auralization is carried out (using binaural synthesis). Finally examples of headphone characteristics are given, and it is discussed how the correct eardrum signals are obtained by headphone reproduction.

The material presented in this paper was obtained in investigations, which are reported more thoroughly in [1]-[4].

1. INTRODUCTION

The term binaural auralization describes the process of creating eardrum signals for the listener, which provides the audible impression of the sound transmission in a room. Based on numerical descriptions (see e. g. [5]) the eardrum signals are artificially generated, typically by means of digital signal processing.

This can be done only if the head-related transfer functions (HRTFs) of the listener are known, and appropriately represented by digital filters. HRTFs can be determined by measurements in the anechoic chamber, where small microphones are mounted in the ears of the listener.

The geometry of humans vary considerably, and ideally the HRTFs used in the binaural auralization should originate from the particular listener [6]. It is however possible to find subjects, whose characteristics better represent the majority of listeners [7] than do others. A tempting shortcut is to use an artificial head for the measurements of HRTFs, but present artificial heads seem to replicate human listeners rather poorly [8].

The measurement of HRTFs is not a straightforward procedure, although carried out worldwide nowadays. A decision of principle matter is how the microphone shall be mounted in the ear canal and thereby at which physical point the signal is picked up. It is the objective to control the eardrum signals in the final reproduction, but it is more convenient to mount the microphone distal to the eardrum.

The computer generated signals are typically reproduced by headphones, because these readily provide the necessary channel separation. Headphones do not have a 1:1 reproduction (a flat frequency response) since they have been designed for other purposes, but supplemented by a correction filter, the correct eardrum signals can be obtained.

The binaural auralization as presented here has proven competitive, especially in auditory virtual environments for Virtual Reality (see also [9]).

2. SPATIAL HEARING - BINAURAL TECHNOLOGY

The hearing has two inputs: sound pressure at the two eardrums. From these inputs the hearing creates an image of the acoustical surroundings, determines direction and distance to sound sources, extracts sound from a single source in noisy surroundings etc. The sound transmission to each of the eardrums depends on direction to the sound source. Various effects like reflection, diffraction, shadowing, dispersion, interference and resonance are involved in a complex acoustical system formed by the body, head, pinna, ear canal and eardrum. The hearing localizes using interaural differences in level and time as well as coloration of the sound signal.

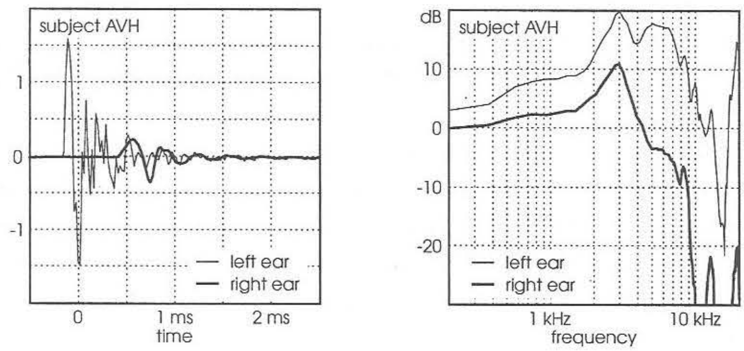

Figure 1

Figure 1 shows the sound transmission to the eardrums of one subject for sound coming from the left side, given in the time domain (impulse responses, left frame), and the frequency domain (amplitude responses, right frame). The sound reaches the left ear approximately 0,6 ms before it reaches the right ear. The sound level in the right ear is lower than in the left ear, and especially the high frequencies are attenuated. If the sound arrives from a source in the median plane, the sound transmission is nearly identical to the two ears. Only the coloration serves as a cue in localization, and it may be difficult to hear the exact direction.

If a listener is given the correct sound at the two eardrums, he may be given an auditory perception that differs from the perception corresponding to his actual physical situation. This is the concept of *binaural technology*. The artificial head recording technique is an example of binaural technology. Sound is recorded in the ears of an acoustical mannequin and played back through headphones. The three dimensional effect is overwhelming. The acoustics of the recording room is precisely reproduced, and sources can be perceived in all directions, including up and down, and they can come close to the listener, down to a few centimeters from the ear.

3. SOUND TRANSMISSION TO THE EARDRUM

The sound transmission to the eardrum can be modelled by the diagram in Figure 2. The transmission outside the ear canal is represented by a Thevenin equivalent circuit, consisting of the impedance seen from the ear canal into the free air $Z_{radiation}$ and the generator $P_{blocked\ ear\ canal}$. "Blocked" refers to the "open circuit" situation, which is obtained by blocking the ear canal, thereby rendering the volume velocity zero.

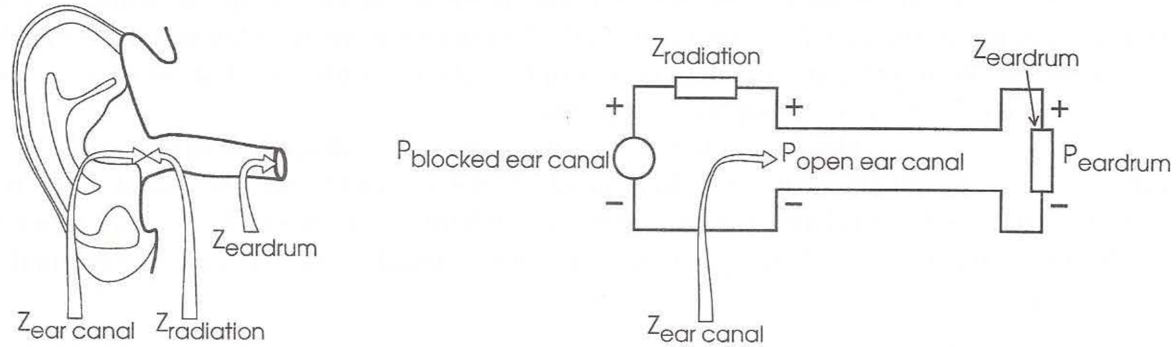

Figure 2

If $P_{reference}$ denotes sound pressure at the center position of the head, but with the subject absent, then the sound transmission can be divided into three parts in the following way:

$$\frac{P_{eardrum}}{P_{reference}} = \frac{P_{blocked\ ear\ canal}}{P_{reference}} \cdot \frac{P_{open\ ear\ canal}}{P_{blocked\ ear\ canal}} \cdot \frac{P_{eardrum}}{P_{open\ ear\ canal}} \qquad (1)$$

The first term is a *head-related transfer function* (HRTF). The second term is the *pressure division* $\frac{Z_{ear\ canal}}{Z_{ear\ canal} + Z_{radiation}}$, while the last term describes the *transmission along the ear canal*. The three terms are shown for one subject and three angles of sound incidence in Figure 3. The way the sound reaches the ear canal does not affect the transmission within the ear canal. Thus only the first term, the HRTF, depends on direction of sound incidence.

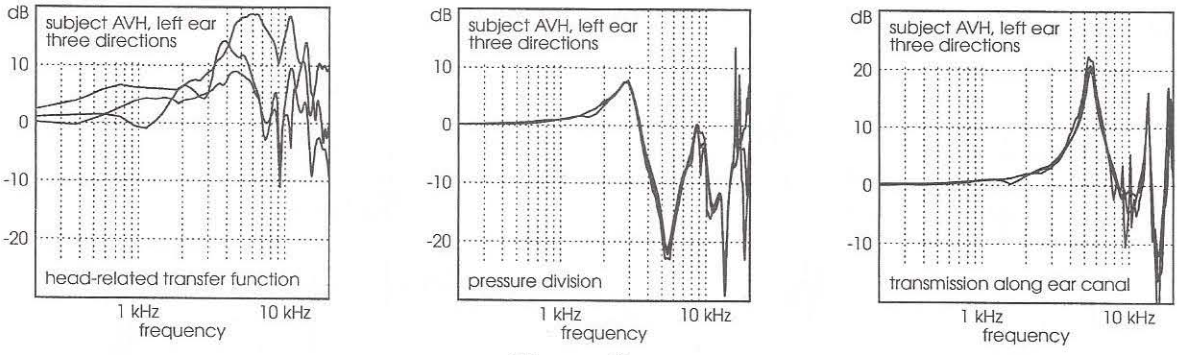

Figure 3

Due to anatomical differences, all elements of the sound transmission to the eardrum are highly individual. The term head-related transfer function is sometimes defined in such a manner that it includes the pressure division and possibly also the transmission along the ear canal (or a part of it). However, full spatial information is included at the blocked ear canal, and the smallest effect of interindividual variation is seen here, since the inclusion of any additional transmission will add variation.

4. BINAURAL SYNTHESIS

When considering sound transmission, the role of the head is to transform each sound wave into two sound pressures, one for each ear. If sufficient knowledge is available about the transmission to the ears for sound from "all" directions (as many as we can discriminate), then it is possible to program a computer to simulate the transmission. The art of artificially creating binaural signals is called *binaural synthesis*.

The directional dependent part of the transmission is described by the HRTF, the transmission to the blocked ear canal. In binaural synthesis, HRTFs are often used in the time domain, that is as head-related impulse responses (HRIRs). The blocked ear canal pressures for the two ears $p_{left}(t)$ and $p_{right}(t)$ resulting from a sound wave s(t) can be obtained by convolution:

$$\begin{aligned} p_{left}(t) &= HRIR_{left}(t) * s(t) \\ p_{right}(t) &= HRIR_{right}(t) * s(t) \end{aligned} \tag{2}$$

Normally, a sound field contains many sound waves, including the direct sound from one or more sound sources, and a number of wall reflections. If the sound field consists of N sound waves, each described by the direction i and the time signal $s_i(t)$, then the resulting sound pressures at the blocked ear canals can be found by summation:

$$\begin{aligned} p_{left}(t) &= \sum_{i=1}^{N} HRIR_{left,\,i}(t) * s_i(t) \\ p_{right}(t) &= \sum_{i=1}^{N} HRIR_{right,\,i}(t) * s_i(t) \end{aligned} \tag{3}$$

For binaural synthesis, the computer should hold a database of HRIRs. Each HRIR has a duration of 1-2 ms. For a 48 kHz sampling frequency this corresponds to 48-96 taps in an FIR filter. Figure 4 shows examples of HRIRs with the left part given as thin lines and the right part as thick lines.

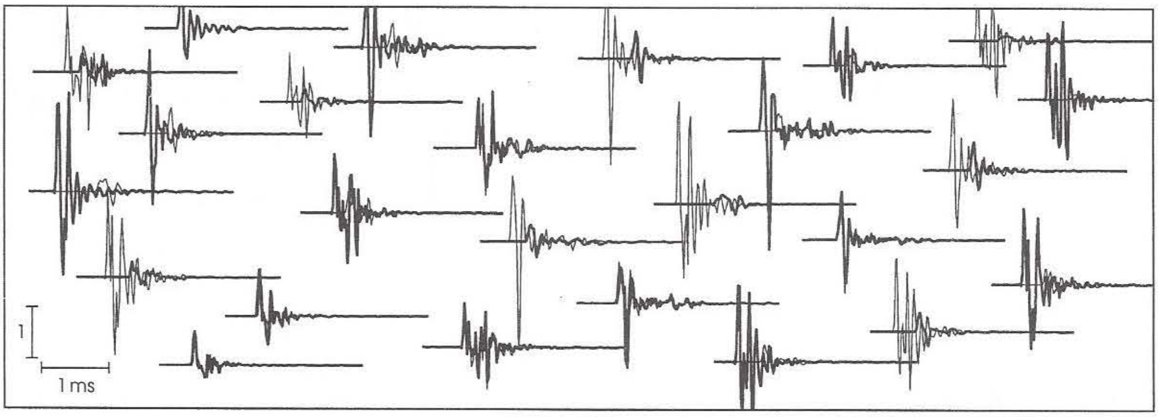

Figure 4

Head-related transfer functions can be split into a minimum phase part, a linear phase part (that is a delay), and an all-pass phase part. In binaural synthesis it is essential to include the minimum phase part and the linear phase part. At present it is unknown, what effect the all-pass phase part has on the sound quality.

5. REPRODUCTION WITH HEADPHONES

The binaural synthesis described in the preceding section simulates the real life sound transmission to the blocked ear canal. Assuming that the remaining transmission to the eardrum is the same during headphone listening and in real life, then the headphone should have a flat frequency response measured at the blocked ear canal. Figure 5 shows frequency responses of three headphones, each measured at the blocked ear canal of 40 subjects. The responses are far from being flat, and equalization is needed. Individual variations are clearly seen, and individual equalization may be relevant.

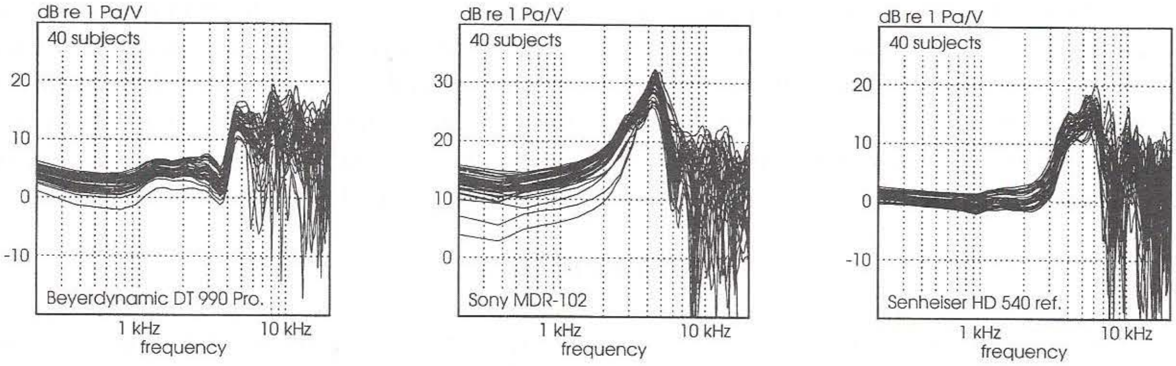

Figure 5

The above assumption about the remaining transmission can be verified by splitting up the transmission into the pressure division and the transmission along the ear canal. The transmission along the ear canal in the headphone situation is identical to that of the real life situation. Only the pressure division is changed to $\frac{Z_{\text{ear canal}}}{Z_{\text{ear canal}} + Z_{\text{headphone}}}$. It differs slightly from that of the real life listening situation, since the radiation impedance is replaced by $Z_{\text{headphone}}$, the acoustical impedance of the headphone as seen from the ear canal. If $Z_{\text{headphone}}$ and $Z_{\text{radiation}}$ are identical, or if they are both small compared to $Z_{\text{ear canal}}$, then the pressure division will be the same in the two situations.

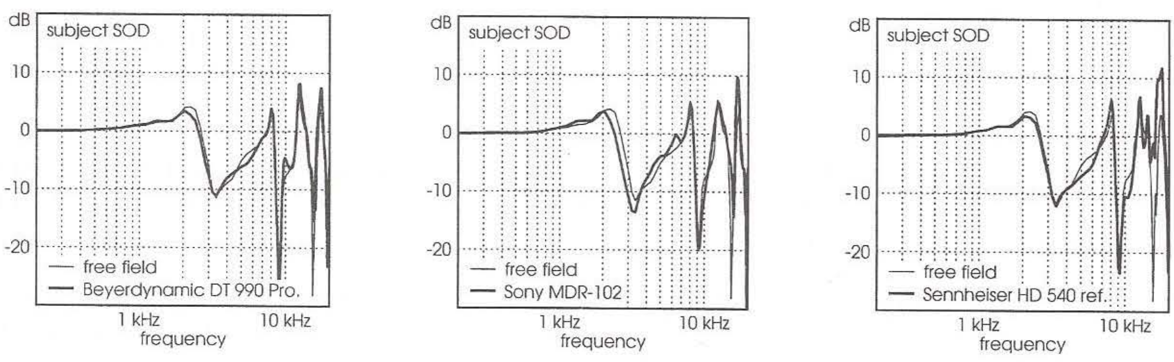

Figure 6

Figure 6 shows pressure divisions measured on one subject in real life and when listening to three commercial headphones. Only minor differences are seen.

6. DISCUSSION

It was mentioned in Section 3 that the sound pressures at the blocked ear canal contains full spatial information, and it was further argued that it had the smallest effect of interindividual variation. The consequence of this is that blocked ear canal sound pressures have a more general applicability, with respect to the use for listeners other than the subject, which the HRTFs were determined for.

Apart from that the blocked ear canal measurements also offer a number of practical advantages. The most immediate is that the microphone mounting becomes relatively easy, especially compared to the critical and often unpleasant measurements at the eardrum. Since the microphone can be mounted in the blockage, it can be larger and provide a better sensitivity than the probe microphones typically used for eardrum measurements.

One advantage which should also be stressed is that using the blocked ear canal measurements, no measurements need ever be made at the eardrum. This has been overlooked by some, who believe that the transmission from blocked entrance to eardrum must be determined once for all. This is not correct. As stated in Section 5, the filters for correcting the headphone characteristics shall also be determined by measurements at the blocked entrance, and the measurements at the eardrum are thus completely avoided.

7. REFERENCES

[1] H. Møller "Fundamentals of binaural technology", Applied Acoustics **36**(3/4) 1992, pp. 171-218.

[2] D. Hammershøi, H. Møller "Sound transmission to and within the human ear canal" J. Acoust. Soc. Am. **100**(1), July 1996, pp. 408-427.

[3] H. Møller, M. F. Sørensen, D. Hammershøi, C. B. Jensen "Head-related transfer functions of human subjects", J. Audio Eng. Soc. **43**(5) May 1995, pp. 300-321.

[4] H. Møller, D. Hammershøi, C. B. Jensen, M. F. Sørensen "Transfer characteristics of headphones measured on human ears", J. Audio Eng. Soc. **43**(4) April 1995, pp. 218-232.

[5] H. Møller "Interfacing Room Simulation Programs and Auralisation Systems", Applied Acoustics **38** 1993, pp. 333-347.

[6] H. Møller, M. F. Sørensen, C. B. Jensen, D. Hammershøi "Binaural Technique: Do We Need Individual Recordings ?", J. Audio Eng. Soc. **44**(6) June 1996, pp. 451-469.

[7] H. Møller, C. B. Jensen, D. Hammershøi, M. F. Sørensen "Using a Typical Human Subject for Binaural Recording", 100th Audio Eng. Soc. Conv. Copenhagen, May 11-14, 1996, preprint 4157, pp. 1-18.

[8] H. Møller, C. B. Jensen, D. Hammmershøi, M. F. Sørensen "Evaluation of artificial head recording systems", To be presented at the 102nd Audio Engineering Society Convention, March 22-27 1997, Munich, Germany.

[9] D. Hammershøi, J. Sandvad "Application of Binaural Synthesis for Auditory Virtual Environments", to be presented at the International Symposium on Simulation, Visualization and Auralization for Acoustic Research and Education, ASVA 97, April 2-4 1997, Tokyo, Japan.