

Evaluation of HRTF Data using the Head-Related Transfer Function Differences

András Illényi

University of Technology and Economics, Department of Telecommunications and Telematics, H-1111 Budapest, Stoczek u. 2, Hungary,
illenyi@tmit.bme.hu,

György Wersényi

Széchenyi István University, Department of Telecommunications, H-9026 Győr, Egyetem tér 1, Hungary, wersenyi@sparc.core.hu

A simple spectral computation method is introduced for the evaluation of differences in dummy-head Head-Related Transfer Functions (HRTFs). These transfer functions describe the directional-dependent filtering of the outer ears and are responsible for basic localization cues. Head-Related Transfer Function Differences (HRTFDs) are defined as the quotient of measured HRTFs in the same direction, comparing the cases of modified environmental conditions. The HRTFDs are relative measures and thus free from systematic errors of the measurement system and individual parameters. Our experience shows that HRTFDs are well suited to analyze the fine structure of measured HRTF data and for representation of the changes in the acoustical environment near or on the head. HRTFDs of a BK dummy-head will be shown representing differences from about 1 dB. This measurement accuracy was determined by replacing the dummy-head with an omnidirectional microphone.

1 Introduction

A number of recent and prospected applications of binaural technology use binaural signals. In all these cases the input consists two sound pressure signals at each of the eardrums, including timbre and spatial aspects. The term 'binaural recording' refers to the fact that the two inputs of the hearing system are reproduced correctly.

For recordings a replica of the human head (the so-called dummy-head) is often used. It is shaped of an average human head: it has a nose, pinnae, and ear canals, enlarged with an attached torso. It is ensured that the sound waves reaching the head undergo the same transmission on their way to the ear channels, as if they were reaching the real listener.

The binaural technique is superior to other recording techniques: it gives a very realistic impression of being present during the recording.

For human listeners seems to be important that small head movements result in perceptible coloration in timbre, they can be a guide to distinguish the incidence of the sound from front and back.

The unambiguous transmission from the one pressure to the other pressure at eardrum means that the sound transmission from that other point to the eardrum is independent of the direction and the distance. In normal cases the listener gets both the direct sound and the reflections. These are responsible for the directional transfer functions (TF). The interaural time differences at low frequencies, the interaural phase differences in the medium frequency ranges, and the interaural level

differences at high frequencies are responsible for the sensation of directions in the full 3D space [1-6].

For investigating the acoustical environment near the head relative measures are called for. The typical indicates here are as following: P_1 the sound pressure in the middle of the head with the listener absent, P_2 the sound pressure at the entrance to the blocked ear canal, P_3 the sound pressure at the entrance to the open ear canal, and P_4 the sound pressure at the eardrum.

The rate of these quantities referred to P_1 , e.g. the P_i/P_1 ratios, are the FTFs, (the free field transfer functions, introduced by Blauert [2]), depending on angle of incidence (ϕ and θ), and distance to the sound source (r).

The P_4/P_1 form of TFs are the directional transfer functions (DTFs), the head-related transfer functions (HRTFs) and the head TFs (HTFs) treated by Møller [4]. Furthermore, we handle monaural and interaural transfer functions as well. These tell us how the transmission deviates in a given direction from the frontal incidence.

Møller introduced the monaural transfer function referenced to diffuse field: P_i/P_1 (average of all angles) [4]. The TFs give a good description of the sound transmission including the diffraction around the head, reflections from shoulders, etc. Physically the explanation can be treated as an impedance-depending behaviour of Z_{earcanal} and $Z_{\text{radiation}}$.

The filtering of the Head-Related Transfer Functions is one of the most important steps during the evaluation of directional information [4, 5]. Virtual simulation

through equalized headphones requires the “accurate” simulation of HRTF filtering. Changes and deviations may lead to confused sensation or localization errors [3-5]. On the other hand, free-field listening in real life allows major variations of the HRTFs without really influencing the localization performance.

To investigate the importance of the fine structure of the HRTFs, we definitely need an accurate measurement system. Our goal was to install a system with increased SNR compared to the general used systems by keeping a good spatial resolution [7-10]. Furthermore, after measuring and collecting a database of dummy-head HRTFs a simple mathematical tool is applied to represent changes in the HRTFs measured from the same direction but under modified environmental conditions.

2 The measurement

2.1 System setup

The HRTFs were measured using a BK dummy-head Type 4128 placed on a turntable in the anechoic room. Spatial resolution is 1° in the horizontal plane and 10° in the elevation from -10° to +60°. Pseudo-random broadband noise signal is used as stimulus, and two channel responses are collected and averaged in a reference measurement. 50 kHz sampling frequency, 16 bit resolution and 4096-point FFT is used [11].

The effects of the undesired transfer characteristics in the measurement chain were eliminated by the reference signal and by calculating the HRTFs as usual:

$$HRTF(j\omega) = \frac{H_{outerears}(j\omega)}{H_{reference}(j\omega)} \quad (1)$$

The validity of the measured HRTFs is above 200 Hz.

2.2 Repeatability of the measurement

In order to determine the repeatability property of the entire measurement system the dummy-head was replaced with an omnidirectional microphone. Repeated transfer function measurements show that deviations according to the uncertainties are about 0,5 dB independent of azimuth and elevation (see Fig.1). Thus, any further measurement result higher than about 1 dB corresponds to real measurement data and is due to physical background. Fig 1 shows the difference between repeated measurements.

2.3 Head-Related Transfer Function Differences

This precision allows us to evaluate differences in a range of about 1 dB in the fine structure of the HRTFs and in the so called HRTFDs (HRTF Differences). Assuming that the complex HRTFs are divisible mathematically, the free-field HRTF Difference (HRTFD) is defined as a quotient of HRTFs from the same direction but under modified conditions:

$$HRTFD = \frac{HRTF_{C_2}}{HRTF_{C_1}} \quad (2)$$

where C_1 identifies the reference and C_2 the modified condition. We plot the $20\log/HRTFD/$ magnitude response as the function of frequency or as 2D polar histogram as function of frequency and azimuth (see Fig. 2-3). In case of $C_1=C_2$ we got the measurement repeatability and the directional properties of the microphone being measured (repeatability property of the bare dummy-head itself).

The division refers to subtraction of two logarithmic magnitude responses. This difference gives us the deviation in dB between two HRTFs measured in the same direction but under modified conditions at all frequencies. For analyzing the HRTFDs we do not need individual recordings on real human heads because the dividing will eliminate and cancel the individual differences. So the dummy-head HRTFs can be regarded as a particular individual set of HRTFs.

We are only interested in changes and deviations caused by modifications of the acoustical environment near the head. But first of all, we have to measure the repeatability property of the bare dummy-head.

3 Results

It was already shown that differences among repeated transfer function measurements are limited below 1 dB with an omnidirectional microphone. By using the dummy-head the differences of the re-measured transfer functions are the HRTFDs. Fig.2 shows uncertainties of the “bare” dummy-head in a 2D polar diagram as function of azimuth and frequency. White domains indicate “accurate” re-measured HRTFs, mostly as the closer ear is directly radiated by the sound source. On the other hand, as the sound source is moving to the contralateral side, differences greater than 9 dB occur in repeated measurements. Black filled areas show the “head-shadow area” where only low frequency evaluation is available.

The polar histogram shows the deviations between HRTFs from repeated measurements in the horizontal plane for the right ear. The natural deviations of the HRTFs caused by the filtering and shadowing effects are shown as unsigned absolute values in dB. The circles correspond to frequency domains with 1 kHz bandwidth marked with the center frequency (linear scale).

The domain filled black is the head-shadow area. Here HRTFs vary more than 9 dB after independent measurements from the same direction. We know from the literature that losing high frequency information of a sound source result in decreased localization performance [2]. Within the head-shadow area (about between 210-330°) on the contralateral side only low frequency components can be evaluated by the hearing system because the variation of the HRTFs can only deliver confusing information. It is impossible to get real directional information of a sound source where the filtering effect of the HRTFs can be more than 9 dB greater in a measurement only seconds later. We assume that this kind of low-pass filtering with a cut-off frequency of about 3-4 kHz can lead to localization errors and the spatial extension of this area by any other shadowing effect in the acoustical environment is the most disturbing effect for the localization.

Note the domain caused by the pinnae between 60 and 90 degrees at 11 kHz and same time at the contralateral side at 2 kHz. In this narrow frequency range and spatial domain HRTFs vary strongly in repeated measurements: differences up to 7-9 dB were observed from the same direction. This phenomenon is due to the pinna only and maybe it can be handled better in the time-domain (reflections). In order to validate this finding we tried to cover and/or remove some parts of the head and torso (shoulders, head, even the complete torso) but this effect only disappeared after the removal

of the artificial pinna off the torso's head. Furthermore, the artificial pinna was coupled to a BK 4166 microphone and placed in a spherical lamp-shell in order to model another "head" and this kind of disturbance was observed again. We have to point on this phenomenon because only this precision of the settings of the azimuth allowed us to discover it. We also suggest making more measurements at this point.

The dummy-head measurement method and results may differ from measurements obtained with real humans due to the deviation in stiffness or shape of the artificial pinna from human pinnae (if the sound source is too close maybe touching the pinna) [12]. Our measurement cancels this kind of individual uncertainties due to the division. The repeatability of 0,50 dB in the standard for L_{Aeq} measurements with a manikin up to 10 kHz can be achieved according to our measurement data up to 8-9 kHz.

This method is well suitable for further analysis as not only the "bare" torso, but the "dressed" can be measured [13]. Fig.3 shows ten HRTFDs from the horizontal plane representing significant influence of hair on the head of the dummy-head. Note that this spatial and frequency domain (150-195 degrees, 3-4 kHz) that is affected by the toupee on the head by 2-10 dB is only affected by 1-3 dB without hair on Figure 2. The representation method of Fig.2 can be also applied for repeated HRTFD measurements of the dressed torso [14].

The HRTFDs are useful tools to investigate differences between sets of HRTFs. The physical explanation can be that the resultant impedance - depending on various components at the head surface such as Z_{glasses} , and Z_{hair} - has frequency depending amplitude and phase values in a wide frequency range that can be responsible for the coloration and interaural differences of measured HRTFs.

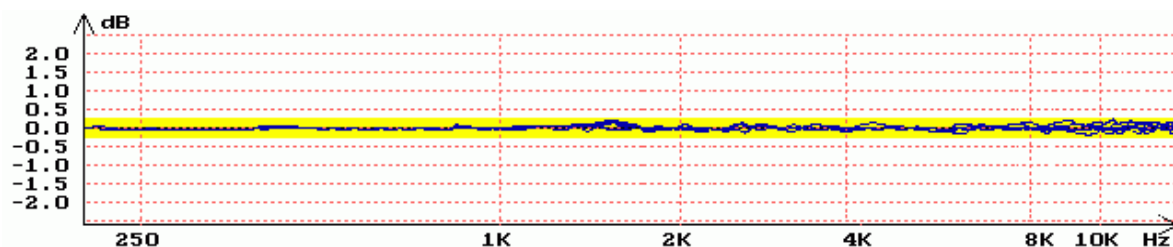


Figure 1: Uncertainties of repeated measurements using the BK 4166 omnidirectional microphone placed on the turntable. TFs are within a 0,5 dB domain (highlighted).

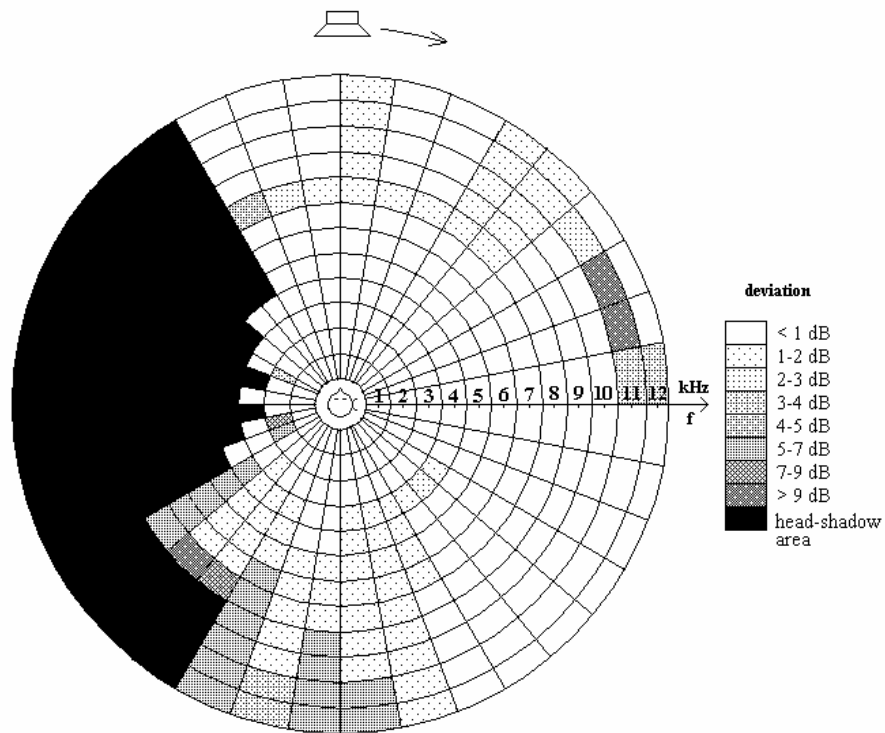


Figure 2: 2D spatial representation of the magnitude of HRTF data for a fixed elevation as a function of azimuth and frequency. The polar histogram shows the deviations between HRTFs from repeated measurements in the horizontal plane for the right ear ($\theta=0^\circ$). The natural deviations of the HRTFs caused by the filtering and shadowing effects are shown as unsigned absolute values in dB. The circles correspond to frequency domains with 1 kHz bandwidth marked with the center frequency (linear scale).

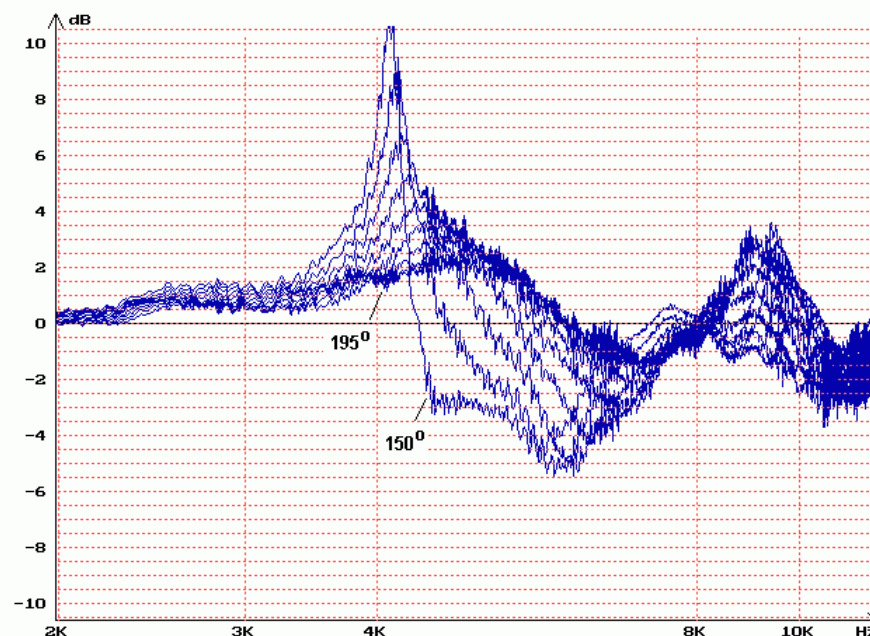


Figure 3: Example from the horizontal plane: HRTFDs as the function of frequency between $\varphi=150^\circ$ - 195° in 5° steps using hair on the dummy-head referring to the bare torso. Note that on Fig.2 there is only a difference of about 1-3 dB in the same spatial domain at 4 kHz (without hair in repeated bare torso measurement) instead of 2-10 dB by using hair.

4 Summary

The measurement system allowed us to evaluate a database of recorded dummy-head HRTFs and HRTFDs. The definition of HRTFD and its 2D representation by unsigned deviations in a polar histogram are well suited for the evaluation of small differences of about 1 dB in the fine structure of the spectra. These HRTFDs

- can be easily calculated (complex division),
- contain no individuality (it will be eliminated by the dividing),
- and they can be measured with the system accurately in a huge amount.
- They can determine the measurement accuracy (using an omnidirectional microphone)
- In simple cases they are able to detect primary reflections, their distance and show the affected spectral regions by the reflections without any time-domain measurement [15].

Effects of clothing and everyday life objects were found to be significant and typical in some frequency and spatial domains

With the HRTFDs it has been proved that small differences in the acoustical environment near the head influences the HRTFs and thus, the SPL at the eardrums without strongly affecting the localization performance and the transmission of the acoustical information in real life environments.

5 Future works

Future works include listening tests using measured HRTF data in order to determine the significance of these changes of the HRTFs caused by the environment near the head in virtual audio synthesis.

References

- [1] P. Minnaar, S.K. Olsen, F. Christiansen, H. Møller, 'Localization with Binaural Recordings from Artificial and Human Heads'. *J. AES.*, Vol. 49(5). pp. 323-336 (2001)
- [2] J. Blauert, 'Spatial Hearing', The MIT Press, MA, (1983).
- [3] H. Møller, M.F. Sorensen, D. Hammershøi, C.B. Jensen, 'Head-Related Transfer Functions of human subjects'. *J. AES.*, Vol. 43(5). pp. 300-321 (1995)
- [4] H. Møller, 'Fundamentals of binaural technology'. *Applied Acoustics.*, Vol. 36. pp. 171-218 (1992)
- [5] D. Hammershøi, H. Møller, 'Free-field sound transmission to the external ear; a model and some measurement', *Proc. of DAGA 91*, Bochum, pp. 473-476. (1991)
- [6] C. B. Jensen, M.F. Sorensen, D. Hammershøi, H. Møller, 'Head-Related Transfer Functions: Measurements on 40 human subjects'. *Proc. of 6th Int. FASE Conference, Zürich*. pp. 225-228 (1992)
- [7] P. Berényi, A. Illényi, 'What does it mean for an HRTF not to have the minimal phase property?'. *Proc. of InterNoise 96, Liverpool*. pp. 2121-2130 (1996)
- [8] J. V. Hudebøll, K.A. Larsen, H. Møller, D. Hammershøi, 'Transfer characteristics of headphones'. *Proc. of 6th Int. FASE Conference, Zürich*. pp. 161-164 (1992)
- [9] K. Genuit, H. Sibinger, 'Kalibrierung einer Kunstkopf-Übertragungskette'. *Proc. of DAGA 94*, Dresden, pp. 685-688. (1994)
- [10] T. Behrens, H. Prante, C. Maschke, 'Untersuchungen zur Summenlokalisierung in der Medianebene'. *Proc. of DAGA 94*, Dresden, pp. 1157-1160. (1994)
- [11] Gy. Wersényi, 'HRTFs in Human Localization: Measurement, Spectral Evaluation and Practical Use in Virtual Audio Environment', PhD thesis, BTU Cottbus, ISBN-963-430-059-6 (2002)
- [12] Acoustics – Determination of sound immersion from sound sources placed close to the ear – Part 2: Technique using a manikin. ISO 11904-2:2004(E), International Standard
- [13] Gy. Wersényi, A. Illényi, 'Differences in Dummy-Head HRTFs Caused by the Acoustical Environment Near the Head'. *Electronic Journal of Technical Acoustics (EJTA)*, <http://webcenter.ru/~eeaa/ejta/eng/abstracts2005eng/wersenyi1eng.shtml>, pp. 1-15 (2005)
- [14] A. Illényi, Gy. Wersényi 'Environmental Influence on the Fine Structure of Dummy-Head HRTFs', *Proc. of Forum Acusticum*, Budapest. (2005)
- [15] Gy. Wersényi, P. Tatai, 'Detection of reflections in free-field directional hearing by waveform analysis of accurate dummy-head HRTFs', *Proc. of IEEE Instrumentation and Measurement Technology Conference, Budapest*, pp. 606-609. (2001)