THE GRAPHIC ANALYSIS AND PRE-PROCESSING OF DIFFERENTIAL HEAD RELATED TRANSFER FUNCTION

D. Štorek, F. Rund

Department of Radioelectronics, Faculty of Electrical Engineering

Czech Technical University in Prague, Prague, Czech Republic

Abstract

This paper focuses on examination of features of Differential Head Related Transfer Function (DHRTF) and it's issues of its processing. DHRTF is being developed and is considered as a new method in virtual sound source positioning with advantages of reduction of data storage and data processing requirements. Processing by DHRTF is based on affecting only one of the pair of the common stereo channels. Anyway, this approach carries several limitations and issues to be solved. The appropriate solutions requires a good comprehension of DHRTF behavior and exploring and visualization of important spatial information which it carries. Matlab offers a sufficient tool for exploring this area. According to the preliminary results a method for increasing DHRTF processing efficiency is introduced.

1 Introduction

Differential Head Related Transfer Function (DHRTF) introduced in this paper is derived from common Head Related Transfer Function (HRTF). HRTF can be considered as a pair of unique filters (impulse responses), which carries the information about changes of sound signal during propagation form a sound source to the listener's ears [1]. With a knowledge of the HRTF filter characteristics an arbitrary sound can be processed in order to create an illusion the sound is coming from a certain location in space. This is the essential effect for many multimedia application and for virtual reality. Nowadays, the principles of virtual sound source positioning based on standard HRTF processing are well known, see e.g. [1-3].

The issue is that HRTF is strongly (but not only) directionally-dependent, therefore a huge amount of data is required for processing. The first authors' analysis dealing with exploration of the main idea of DHRTF sound processing can be found in [4]. As DHRTF is derived from HRTF (see below), it is also strongly frequency-dependent. This article tries to discover the dependence of HRTF behavior on spatial settings.

2 Principles of DHRTF

For better understanding of DHRTF function a few mathematical definitions shall be introduced [4]. HRTF is a complex function which can be defined for both ear channels as shown in Equation 1.

$$HRTF_n = \left|HRTF_n\right| \cdot e^{j\psi_n} = \frac{FT\{p_n(t)\}}{FT\{p_S(t)\}} \quad n = L, R \tag{1}$$

where $p_n(t)$ is sound pressure in time at the place of left or right ear canal entrance and $p_s(t)$ is sound pressure at the place of the sound source, FT represents Fourier transform.

Processing by HRTF provides unique amplitude and phase spectral features to each ear channel. The phase information is important for preserving time delay of the signal coming to contralateral ear. Human brain is able to evaluate the sound source position according to the mutual time and level differences in both channels. Common terms used in this area are Interaural Time Difference (ITD) and Interaural Level Difference (ILD). The effect of spatial impression can be created either by time, or level shift, or by combination of both. Classic stereo panning (e.g. in music mixing) is achieved by applying ILD. The fact, that as ILD, as ITD is frequency-dependent is usually neglected.

The idea of DHRTF deals with an extraction of the inter-channel time and level differences from a pair of HRTF filters, as shown in Equation 2.

$$DHRTF = \frac{|HRTF_L|}{|HRTF_R|} \cdot e^{j(\psi_L - \psi_R)} = \frac{FT\{p_L(t)\} \cdot FT\{p_S(t)\}}{FT\{p_R(t)\} \cdot FT\{p_S(t)\}} \cdot e^{j(\psi_L - \psi_R)} = ILD(\omega) \cdot e^{j \cdot ITD(\omega)} \qquad (2)$$

where ψ is the phase spectrum of standard HRTF for left and right ear. Therefore DHRTF can be considered as a transfer function, whose magnitude is actually frequency-dependent ILD and the phase is frequency-dependent ITD. The term $\psi_L - \psi_R$ stands for differential phase, defined as $d\psi$ in Eq. 3, where k stands for discrete time.

$$d\psi[k] = \arg\left(FT\{HRIR_L[k]\}\right) - \arg\left(FT\{HRIR_R[k]\}\right) \qquad (3)$$

where HRIR is inverse Fourier transform (so-called Head Related Impulse Response). Inverse Fourier transform of a complex DHRTF results in Differential HRIR (dHRIR), as shown in Eq. 4.

$$dHRIR[k] = FT^{-1}\left\{ \left| DHRTF[k] \right| \cdot e^{jd\psi[k]} \right\} \qquad (4)$$

As mentioned above, an advantage of DHRTF method lies on a fact, that only one of the stereo channel is processed, as described by Eq. 5 and 6, where *x[k]* stands for the initial processed signal. The second one keeps its original parameters, but the mutual time-shift and amplitude ratio follows the same values as in the case of classic virtual positioning.

$$y_L[k] = dHRIR[k] * x[k] = \sum_{l=0}^{M-1} dHRIR[l] \cdot x[k-l] \qquad (5)$$

$$y_R[k] = x[k] \qquad (6)$$

3 Analysis of DHRTF behavior

In order to understand DHRTF behavior, it is important to use some visualization tool. Matlab offers a wide range of functions for three-dimensional representation of DHRTF. The data to be analyzed are available in [5]. For the purpose of this article a set of 'special_kemar_hrir' was used. HRTF is also subject-dependent, therefore the data corresponding to the normalized artificial head-and-torso model was chosen. The first analysis concerns a visualization of HRIR for one ear (left), as shown in Fig. 1.

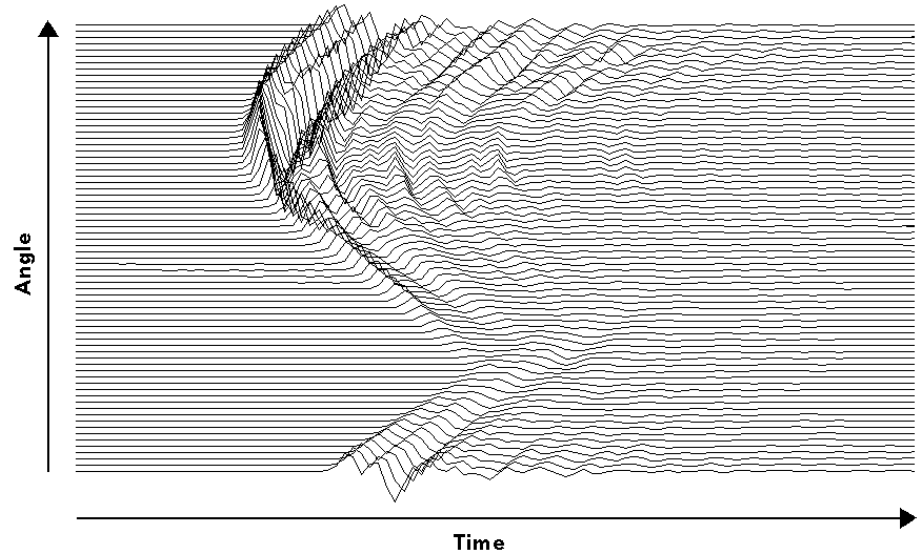

Figure 1: The 3D visualization of pure HRIR performed by 2D *plot* function The distance between each step corresponds to the 5° change in azimuth plane. The 72 measured positions from [5] refers to whole 360 degree radius round the user, while the head turns clockwise.

The time of incidence of the signal is determined by the angle between the head center and the sound source. The starting point (down) refers to position φ = 0° (straight ahead). As the head changes its position (clockwise turn), the time of incidence increases. The smallest HRIR corresponds to position φ = 90°, when the source is directly in front of the right (opposite) ear. The head structure causes strong attenuation. For the position φ = 180° (source behind) the HRIR resembles position 0° due to head symmetry. The shortest time of incidence corresponds to position φ = 270°, where the source is directly in front of the left ear. This visualization was created by common *plot* function and in *for* cycle each curve was depicted straight to above the previous HRIR. 72 values of HRIR are able to create very impressive notch-and-peaks structure.

The same approach can be applied in order to visualize the DHRTF, as shown in Fig. 2. mentioned above, the curves refer to the spectral difference in stimuli coming to the left ear from different directions.

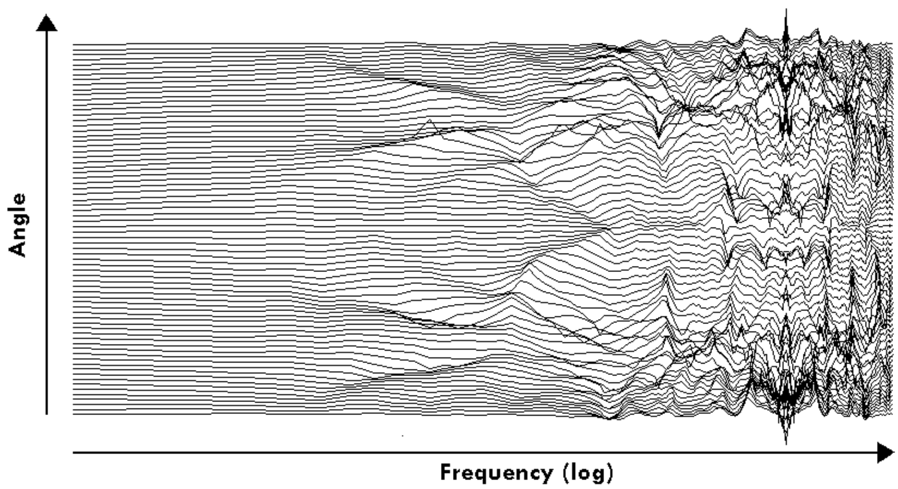

Figure 2: DHRTF behavior on the spectral-azimuth plane. The conditions are the same as in Figure 1.

The structure of spectral changes in higher frequency band is now quite poorly recognizable. When the data is visualized by *meshc* function, the result of spectral changes is much obvious, as can be seen in Fig. 3. The higher the frequency is, the easier can be attenuated by head as an obstacle due to the wavelength. Therefore, DHRTF shows wider range of features at higher frequencies.

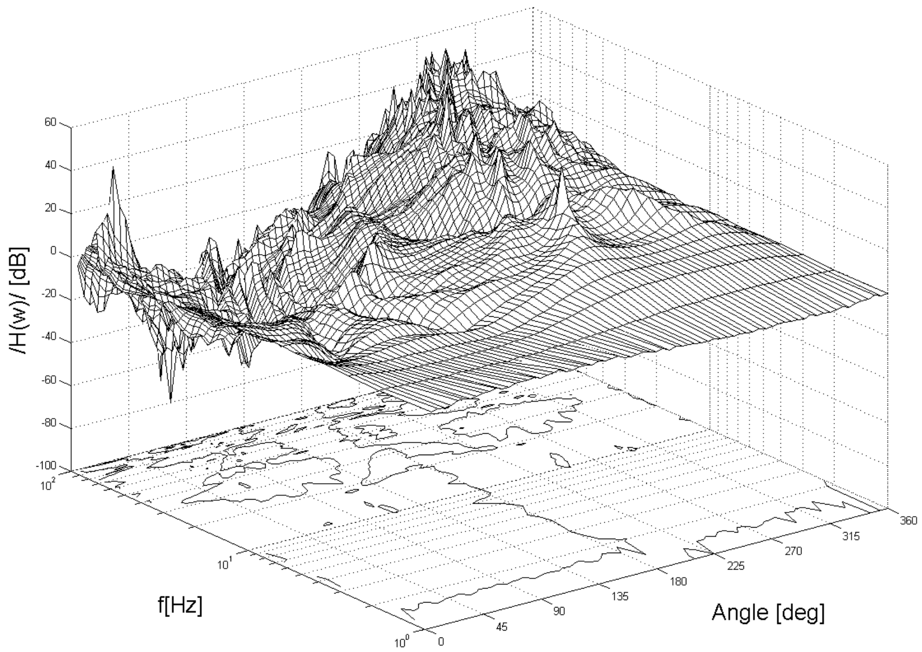

Figure 3: DHRTF visualized by *meshc* function provides much better information about trend of the curves. The peaks and the notches are much better recognizable than in Fig. 3

4 The Inter-channel Energy Balance

The first issue to be solved in this area concerns energy balance of different DHRTF positions. As DHRTF is defined by Eq. 2, the final value of DHRTF either falls under 0 dB level, when φ < 180°, and exceeds this level when φ > 180°. For better clarity of the final effect the Eq. 2 can be turned to Eq. 7, as follows.

$$|DHRTF| = 20 \cdot \log_{10}|FT\{HRIR_L\}| - 20 \cdot \log_{10}|FT\{HRIR_R\}| \tag{7}$$

According to Fig. 3, the gain in the processed channel either increases or decreases with dependence on the azimuth angle. This fact causes fluctuations of loudness of the final perceived positioned signal. Anyway, this feature does not correspond to real perception, therefore DHRTF set has to be pre-processed in advance. The weighting finally enables perception of the virtually positioned sound in an appropriate position with correct loudness in relation to the sound positioned by standard HRTF pair. Instantaneous signal energy (meant as the square of the signal) of the impulse responses HRIR from Fig. 1 is shown in Figure 4 below.

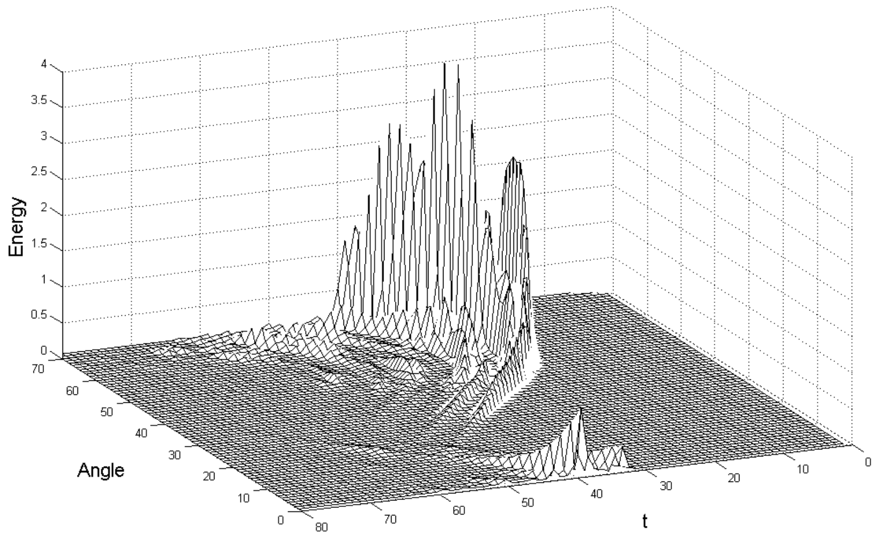

Figure 4: Instantaneous energy of HRIR set for left ear in 3D visualization.

Distribution of signal energy from various directions enables better upgrade of the channel weighting in DHRTF processing. The whole energy of HRIR set can be obtained by Equation 7, where *n* represents the n-th HRIR of 72 sample set.

$$E_n = \int_{-\infty}^{\infty} |HRIR_n(t)|^2 dt \tag{7}$$

The final graphic representation of the HRIR energy distribution for polar coordinates is summarized in Figure 5 as in standard graph, as in polar graph.

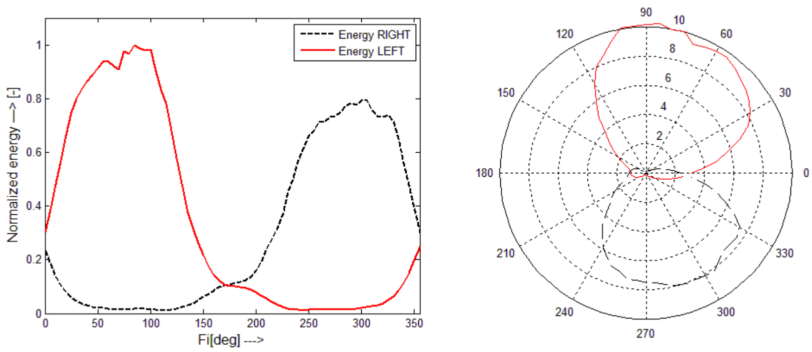

Figure 5: Distribution of HRIR energy in polar coordinates.

As expected, the energy curve for left and right ear are symmetrical due to the symmetry of the acoustic manikin on which the HRIR set has been measured.. The surprise is the gains in left and right ear are different (even though the symmetry). Interesting positions are 0° (in front)and 180° (behind). There can be demonstrated the influence of the pinna. The pinna structure can cause up to 4.5 dB attenuation of the signal. For signal balance during the virtual positioning the channels are supposed to be weighted by normalized energy index in order to preserve the same loudness distribution over the spatial source positions. This can be achieve by utilizing the normalized energy curve. The aspect of final loudness perception has to be focused in further experiments.

5 Results

The graphical analyses introduced in this article were great assistants during solving the issue of energy balance of virtual sound source positioning performed by new method of utilizing Differential Transfer Function. A proper signal energy distribution in both channels is essential for final spatial impression of the virtually positioned stimuli while being presented by headphones.

An appropriate design of energy weighting is going to be tested further by series of psychoacoustic listening test in order to verify the suitability of this method. The energy distribution curves will have to be created over a large number of live subjects. This approach can be also reliable method for a quick check of HRTF data in order to avoid occurrence of offset in virtual positioning. This is now the goal for further work in this area.

Acknowledgement

The project „The Graphic Analysis and Pre-processing of Differential Head Related Transfer Function“ was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS11/159/OHK3/3T/13.

References

[1] U. Zölzer, *DAFX – Digital Audio Effects,* John Wiley & Sons, 2002
[2] Algazi V. R.; Duda R. O., *"Headphone-Based Spatial Sound"*, IEEE SIGNAL PROCESSING MAGAZINE, Vol. 28, pp. 33-42, 2011
[3] Rumsey F., *"Whose head is it anyway? Optimizing binaural audio"*, J. Audio Eng. Soc., Vol. 59, pp. 672-677, 2011
[4] Štorek, D. - Rund, F.: *Differential Head Related Transfer Function as a New Approach to Virtual Sound Source Positioning*. In Proceedings of 22nd International Conference Radioelektronika 2012, p. 71-74, ISBN 978-80-214-4468-3.
[5] Algazi, R., Aveando, C., Thompson, D., Avendano, C., "The *CIPIC HRTF Database,*" in Proc. IEEE WASPAA´01, New Platz, NY, 2001, http://interface.cipic.ucdavis.edu/

Ing. Dominik Štorek
Department of Radioelectronics, FEE, CTU in Prague, Technická 2, 166 27 Praha 6, Czech Republic,
email: storedom@fel.cvut.cz

Ing. František Rund, Ph.D.
Department of Radioelectronics, FEE, CTU in Prague, Technická 2, 166 27 Praha 6, Czech Republic,
email: xrund@fel.cvut.cz