The TUM-LDV HRTF Database

M. Rothbucher [1] * P. Paukner [1] M. Stimpfl [1] K. Diepold

In this article a high resolution database of Head-Related Impulse Responses (HRIRs) and anthropometric parameters for 35 human subjects and two KEMAR manequins is introduced. The HRIRs were computed using Normalized Least Mean Square (NLMS) adaptive filtering.

The NLMS method has no audible drawbacks compared to well-established HRIR estimation approaches (MLS and Exponetnial Sweep method) while allowing quasi-infinite azimuth resolution by continuously recording of the excitation signal of a rotating subject (Rothbucher et al., 2013).

The obtained LDV database provides HRIRs of six different elevation planes with an azimuth resolution of 1° or 0.2° and corresponding anthropometric data of 35 human subjects. In this article, the measurement process and the post processing of the HRIRs is described in detail.

1 Introduction

Head-Related Transfer Function (HRTF) based approaches for 3D-sound synthesis have become promising techniques to enable immersive sound playback for various applications. The spectral changes of a sound waves when they enter the ear, caused by refections and diffractions of the human body, e.g. the torso head and shoulders are individual for each person due to the unique body geometry of each user.

Unfortunately, HRIRs, the time domain representation of the HRTFs, are usually measured by a cumbersome procedure with expensive equipment, including an anechoic chamber (Algazi et al., 2001), which is not commonly accessible to researchers that are working on immersive 3D-sound synthesis. Consequently there is increasing number of researchers that seek to customize a set of HRTFs based on the anthropometric data without acoustic measurement (Nishino et al., 2007).

*[1] Institute for Data Processing, Technische Universität München

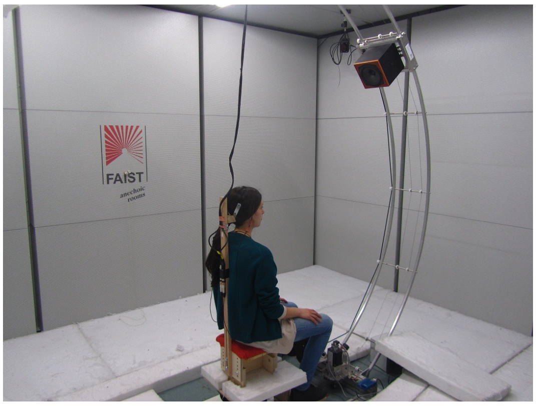

Figure 1: Measurement setup in the semi-anechoic chamber

Our aim is to provide a high resolution database of measured HRIRs and the corresponding anthropometric data, that enables artefact-free HRTF-based 3D sound synthesis without interpolation. The database consists of the acoustic measured HRIRs of 35 human subjects and two KEMAR manequins with the corresponding anthropometric data and could therefore be a basis for HRTF customization approaches.

2 Setup

The measurements were carried out in the semi-anechoic chamber of the Institute for Data Processing at the Technische Universität München. A detailed description of the measurement environment can be found in (Rothbucher and Volk, 2013) and an overview of the measurement equipment is given in Table 1.

The subjects were placed on a custom-build chair with adjustable seat hight and headrest, mounted on a turntable with a stepper motor (Nanotec PD6-N turntable position control).

The chair was positioned in the center of a metallic arc, designed to allow an automatic elevation adjustment of the mounted speaker. The distance from the center of the arc to the membrane of the speaker was 1.3 m. Ear signals were recorded with two microphones (KnowlesAcoustics), placed in blocked-meatus conditions (Møller, 1992). To ensure the correct position of the microphone and the reproducibility of the results, the ear canals of the subjects were blocked with silicon imprints that were constructed for each test subject individually.

Sound sources	KS digital C5 tiny
Microphones	Knowles FG-23629
Microphone amplifier	SNR: 68dB, high-pass: 15-240Hz 6dB/octave
Sound card	RME Multiface II
Audiolab dimensions	4.7 m x 3.7 m x 2.84 m
Audiolab noise level A-weighted	17 dB
Audiolab reverberation time t_{60}	0.08 s

Table 1: Information about the equipment and the environment used in our localization experiments.

3 Measurement Procedure

The complete procedure of obtaining the LDV-database can be summarized as follows: the manufacturing of the ear imprints, the HRIR-measurement and the measurement of the anthropometric parameters. All parts of the procedure were carried out by instructed supervisors. In this section, the three parts of the measurement process are described in detail.

3.1 Imprints

The imprints are made by injecting a liquid self-curing silicon into the ear canals of the subject. After hardening, the imprints are cut such that the imprint provides a plane surface at the entrance of the meatus. Finally, we fix the microphones to the imprints by drilling a microphone hole into the centroid of the entrance surface, illustrated in Figure 2.

3.2 HRIR Measurement

In order to align the interaural axis to the center of the hoop, the seat height is adjusted in advance of the acoustic measurement with respect to the height of the test subject. Moreover, the person has to take a predefined posture (upright, hands on thighs) as illustrated in Figure 1. An adjustable headrest further prevents head movements during the measurement process.

According to (Rothbucher et al., 2013) the acoustic measurement per elevation was achieved by a continuous counter-clockwise turn from -30° to 370° while the speakers presented an equally distributed noise with a sound pressure of 70 dB(A) at each subjects head. The rotation speed of the turntable was set to $v_{rot} = 3.8^\circ \frac{1}{s}$, which leads to a recording duration of $t_{rot} = 105$ s per elevation. The turntables encoder continuously logged the actual azimuth position in order to map the angle with the actual sample of the sound files for the calculation of the HRIRs.

Figure 2: Imprints with microphones

The whole procedure was controlled with a webcam that was placed in the lab. Moreover, the sound pressure levels of the recorded signals and the microphone signals were continuously controlled during the measurement process. If the subject moved during measurement or other events disrupted the recordings, the measurement was repeated.

3.3 Anthropometric Measurement

After the audio recordings, the anthropometric parameters were measured using a caliper and a slide gauge. The choice of the parameters to be measured is in accordance with (Hu et al., 2006): head width, head depth, shoulder width, cavum concha height, cavum concha width, fossa height, pinna height and pinna width as illustrated in Figure 3.

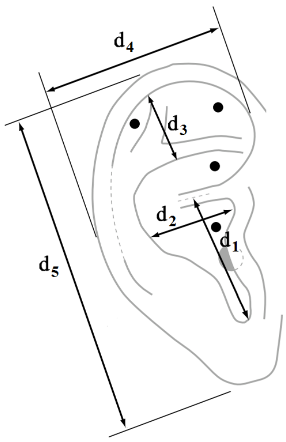

Figure 3: Anthropometric parameters for the pinnae (Algazi et al., 2001). (d1: cavum concha height, d2: cavum concha width, d3: fossa height, d5: pinna height and d6: pinna width)

3.4 Compensation Measurement

After every measurement-day an extra measurement for the compensation was conducted, resulting in a compensation impulse response (cIR). The microphones were fixed at the center of the hoop and the impulse response was measured by using the MLS method described in (Schroeder, 1979).

4 Compensation Procedure

Using Blauert's definition for the HRTF (Blauert, 1997), one has to divide the subjects HRTF by the frequency domain representation of the cIR. HRIR and cIR are modeled as Finite Impulse Response (FIR) filters.

Inverting the cIRs in frequency domain results in an unstable Infinite Impulse Response Filter (IIR), due to the non-minimum phase characteristics of the compensation Transfer Function (cTF). Therefore, a FIR filter approximation of the inverse cTF was used by computation of the Moore-Penrose-Inverse.

To overcome inversion instabilities, a frequency-dependent regularization was used. Only small filter coefficients were regularized while stable filter coefficients remain unaffected (Choueiri, 2008).

5 Public Database

The public database consists of four different MAT-files for each subject containing raw and compensated HRIRs with an azimuth resolution of 1° or 0.2°, respectively. The variables of the MAT-files are:

- 'HRIR': matrix with the impulse responses. The dimensions of the matrix correspond to (samples, [left channel, right channel], azimuth, elevation).
- 'azimuth': array of the provided azimuths.
- 'elevation': array of the provided elevations.
- 'anthropometricParameters': cell-array with the measured parameters. Parameters concerning the ears are provided as arrays which dimensions correspond to [left ear, right ear].

References

1. V. Algazi, R. Duda, D. Thompson, and C. Avendano. The CIPIC HRTF database. In *IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics*, pp. 99–102. 2001.

2. J. Blauert. *Spatial hearing: the psychophysics of human sound localization*. MIT press, Massachusetts, 1997.

3. E.Y. Choueiri. *Optimal Crosstalk Cancellation for Binaural Audio with Two Loudspeakers*. Technical Report, 3D3A Laboratory, Princeton University, 2008.

4. H. Hu, L. Zhou, J. Zhang, H. Ma, and Z. Wu. Head related transfer function personalization based on multiple regression analysis. In *International Conference on Computational Intelligence and Security*, volume 2, pp. 1829–1832. IEEE, 2006.

5. KnowlesAcoustics. Datasheet knowles acoustics fg-3329. *http://datasheet.octopart.com/FG-3329-Knowles-Acoustics-datasheet-98123.pdf*. [accessed october-2013].

6. H. Møller. Fundamentals of binaural technology. In *Applied acoustics*, 36(3), pp. 171–218, 1992.

7. T. Nishino, N. Inoue, K. Takeda, and F. Itakura. Estimation of HRTFs on the horizontal plane using physical features. In *Applied Acoustics*, 68(8), pp. 897–908, 2007.

8. M. Rothbucher, K. Veprek, P. Paukner, T. Habigt, and K. Diepold. Comparison of head-related impulse response measurement approaches. In *The Journal of the Acoustical Society of America*, 134(2), pp. EL223–EL229, 2013.

9. M. Rothbucher and T. Volk. *HRTF Measurement in the Audiolab at LDV*. Technical Report, Institute for Data Processing, Technische Universität München, 2013.

10. M.R. Schroeder. Integrated-impulse method measuring sound decay without using impulses. In *The Journal of the Acoustical Society of America*, 66(2), pp. 497–500, 1979.