Audio Engineering Society

Engineering Brief

Presented at the 130th Convention
2011 May 13–16 London, UK

A Free Database of Head-Related Impulse Response Measurements in the Horizontal Plane with Multiple Distances

Hagen Wierstorf, Matthias Geier, Alexander Raake, Sascha Spors

Deutsche Telekom Laboratories, Technische Universität Berlin, Ernst-Reuter-Platz 7, 10587 Berlin, Germany

Correspondence should be addressed to Hagen Wierstorf (`Hagen.Wierstorf@telekom.de`)

ABSTRACT

A freely available collection of Head-Related Impulse Response (HRIR) measurements is introduced. The impulse responses were acquired in an anechoic chamber using a KEMAR manikin at four different loudspeaker distances – 3 m, 2 m, 1 m and 0.5 m – reaching from the far field to the near field. The loudspeaker was positioned at ear height and the manikin was rotated with a high-precision stepper motor in one degree increments. Besides the raw HRIRs also datasets are available which have been compensated for the use with specific headphone models.

1. INTRODUCTION

In order to localize a sound source, the human auditory system takes advantage of the time lag and level difference between the two ear signals as well as cues added by reflection and diffraction by torso, head and external ears [1]. Under the assumption of a linear and time-invariant system, a pair of impulse responses can be used to describe all of these cues. If such an impulse response is acquired under free-field conditions, only the influences of the abovementioned cues are present and the impulse response is called head-related impulse response (HRIR). Its Fourier transform is termed as head-related transfer function (HRTF). Several databases of HRIRs are freely available [2, 3, 4], but their angular resolution in the horizontal plane is limited and only [4] contains measurements for more than one source distance (0.8 m and 3 m).

For distances larger than 1 m the distance between the manikin and the source is assumed to have no influence on the binaural cues, only the amplitude decreases for larger distances [5]. If the source is positioned closer to the listener, changes in the binaural cues have to be considered. Hence, specific HRIR measurements for nearby sources were made [5, 6]. To the knowledge of the authors these databases are currently not publicly available. This study will provide HRIR databases measured in the horizontal plane with a resolution of 1° for sources positioned at a distance of 0.5 m, 1 m, 2 m and 3 m, respectively. The location for downloading the database is provided in the last section of the paper.

2. SETUP AND MEASUREMENTS

The impulse response measurements were carried out in the anechoic chamber of the Institut für Tech-

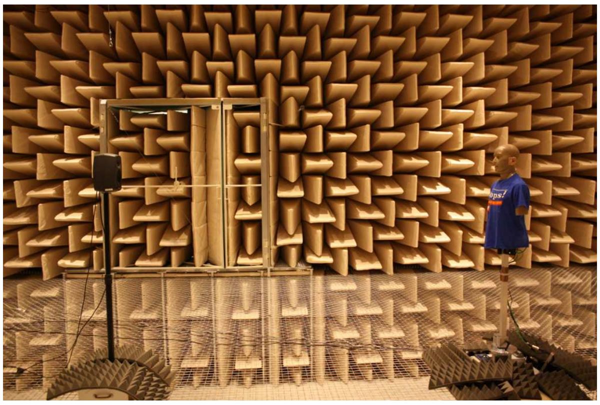

Fig. 1: Setup of manikin and loudspeaker in the anechoic chamber.

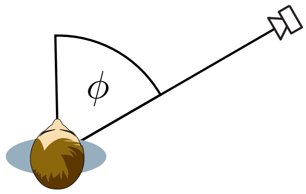

Fig. 2: Coordinate system used in this paper. The azimuth angle ϕ describes the movement of the loudspeaker counter-clockwise around the manikin. Here $\phi = -60°$.

nische Akustik at Technische Universität Berlin. The lower frequency limit of the anechoic chamber is 63 Hz. An active two-way loudspeaker (Genelec 8030A) was used as a sound source. The excitation signal was a linear sine sweep with a length of 5.3 s. It had a 6 dB per octave low-shelf emphasis below 1000 Hz causing an amplification of 20 dB for low frequencies.

The responses were measured using a Knowles Electronics Manikin for Acoustic Research (KEMAR, type 45BA) with the corresponding large ears (type KB0065 and KB0066). The manikin was mounted on the turntable of the VariSphear measurement system [7] to be able to rotate it automatically with high mechanical precision. All measurements were performed for a full rotation (360°) in 1° increments. The distance between the center of the manikin and the center of the loudspeaker was set to 3 m, 2 m, 1 m and 0.5 m, respectively. The loudspeaker was positioned at ear height of the manikin.

Ear signals were recorded with G.R.A.S. 40AO pressure microphones using a RME QuadMic preamplifier and a RME Multiface II audio interface. Playback, recording and control of the measurements was done with the VariSphear software [7]. All data was recorded with a sampling rate of 44.1 kHz and stored as single precision floating point values.

In addition to the HRIRs, headphone and loudspeaker impulse responses were measured. Several headphones (AKG K 601, AKG K 271 MK II, Sennheiser HD25-1) were placed on the manikin and their impulse responses were measured in order to provide compensation filters. Each headphone measurement was repeated several times including removing and placing the headphones on the manikin.

The impulse response of the loudspeaker was measured for all distances at the center of the artificial head by replacing it with a single pressure microphone (Brüel & Kjær type 4189 using a B&K Nexus conditioning amplifier type 2690-OS1).

3. DATA PROCESSING

Post-processing was applied to the HRIRs in order to fix measurement inaccuracies and to compensate the influence of the loudspeaker. To reduce the amount of data, the length of the HRIRs was truncated. Depending on the distance of the source to the manikin there are leading zeros in the impulse responses. These were removed for the HRIRs with a distance larger than 0.5 m to achieve the same amount of leading zeros as in the 0.5 m dataset. After this step the length of each impulse response was truncated to 2048 samples.

The setup of the manikin in the anechoic chamber suffered from slight inaccuracies regarding the placement of the loudspeaker with respect to the head orientation. In order to compensate for these, the mini-

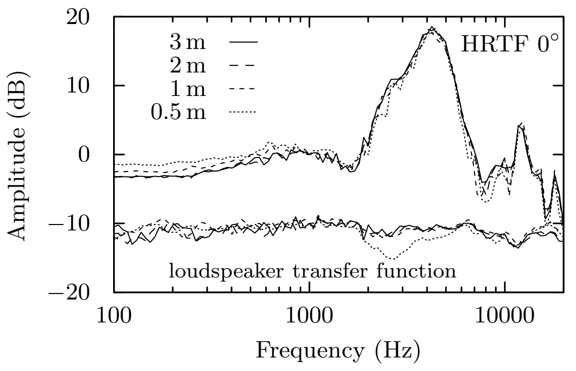

Fig. 3: *Above:* HRTFs for a source position of $\phi = 0°$ including loudspeaker compensation but without headphone compensation. *Below:* loudspeaker transfer functions.

mum value of the interaural time difference (ITD) for angles between $-90°$ and $90°$ was determined and the datasets were rotated by 2°-3° in order to align the minimum to an angle of 0°. In a second step, the interaural level difference (ILD) between the two ears was corrected by applying a gain to the left and right HRIR in order to achieve an ILD of 0 dB for an azimuth angle of 0°.

The distance-dependent frequency response of the loudspeaker was considered by designing and applying inverse finite impulse response (FIR) filters. The filters were designed as linear-phase filters with a length of 128 coefficients using frequency-dependent regularization of the measured loudspeaker responses. The loudspeaker responses were only equalized between 100 Hz and 10 kHz due to the chosen regularization parameters. Note that the such compensated HRIR datasets conform to the definition of HRIRs used in [5].

Furthermore, compensation filters were calculated for three different headphone models. The procedure followed the work presented in [8]. In a first step, one complex transfer function per headphone model was derived by averaging the repeated measurements for the left and right channel. A frequency domain least-squares error inversion technique with frequency-dependent regularization was used to calculate the headphone compensation filters. The chosen regularization allows to equalize only those parts of the frequency response which did not vary too much for the repeated measurements. The length of the resulting filters is 4096 coefficients.

4. DATA ANALYSIS

In order to evaluate the measured HRIRs, the signal-to-noise ratio (SNR) has been calculated for the different distances at an azimuth of $\phi = 0°$. This has been done by subtracting the peak level of the noise from the peak level of the entire measured signal for the left and right ear separately. As noise signal the part from 0.25 s to the end of the impulse response was used. The SNR values for the right and left ear are presented in Table 1.

Table 1: SNR values for the left and right channel of the impulse response for different distances and an azimuth angle of $\phi = 0°$.

distance	left SNR (dB)	right SNR (dB)
0.5 m	77.3	79.0
1 m	84.2	84.3
2 m	83.2	82.7
3 m	81.4	81.0

Diffraction due to the head leads to an amplification at low frequencies in the HRIRs [5], which is shown in the upper part of Figure 3 for an azimuth angle of $\phi = 0°$. As can be seen for the distances of 3 m and 2 m, the HRTFs are very similar, but for 1 m and 0.5 m, frequencies below 800 Hz are slightly more prominent.

In Figure 4, the ILDs for the different HRIR datasets are presented. It can be seen that the magnitude of the ILD is greater for closer sources. This effect was also reported in [5]. It is most prominent for a location of the source at the side of the right or left ear ($-90°$ or $90°$). This can be explained by the fact that the head shadow effect and the decrease in magnitude proportional to $1/r$ are more pronounced for nearby sources.

As mentioned in Section 2, a two-way loudspeaker was used. This potentially violates the assumption of a point source for the measurement of the impulse response for close distances. The lower part of Figure 3 shows the transfer function of the loudspeaker measured at the center of the head for the different distances. For a distance of 0.5 m deviations can be

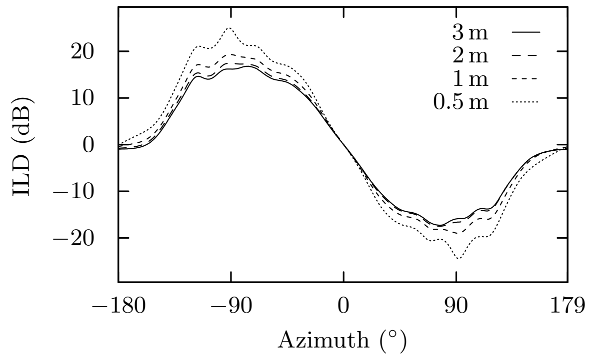

Fig. 4: Interaural level differences (ILDs) of the HRIR datasets for different source distances.

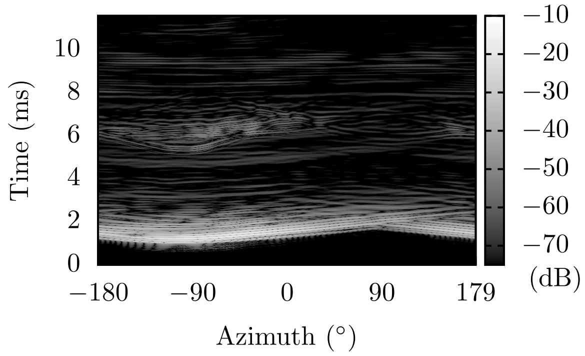

Fig. 5: Amplitude of the measured HRIRs for a distance of 1 m.

observed with respect to the other distances. These deviations are compensated by the filter described in Section 3.

The spatial extent of the loudspeaker leads to further deviations in comparison with an ideal point source. It should be noted that there are additional reflections from the artificial head back to the loudspeaker cabinet and from there back to the ears. This can be observed in the HRIR datasets for 1 m (see Fig. 5) and 0.5 m. For larger distances these reflexions are negligible.

5. HRIR DATABASES

All measurements described in this paper are freely available for download[1] under a *Creative Commons Attribution-NonCommercial-ShareAlike 3.0* license[2]. The datasets are available in different formats suitable for a variety of applications: `.mat` files for use in *GNU octave* and *Matlab*, `.daff` files in the *Open Directional Audio File Format* (OpenDAFF)[3], an emerging open standard for spatial audio data, and `.wav` files for use in the *SoundScape Renderer* (SSR)[4].

[1] `http://dev.qu.tu-berlin.de/projects/measurements/`
[2] `http://creativecommons.org/licenses/by-nc-sa/3.0/`
[3] `http://www.opendaff.org`
[4] `http://tu-berlin.de/?id=ssr`

6. REFERENCES

[1] J. Blauert. *Spatial Hearing.* The MIT Press, 1997.

[2] B. Gardner and M. Keith. HRTF measurements of a KEMAR dummy-head microphone. Technical Report #280, MIT Media Lab Perceptual Computing, 1994.

[3] V. R. Algazi, R. O. Duda, D. M. Thompson and C. Avendo. The CIPIC HRTF database. In *WASPAA*. 2011.

[4] H. Kayser et al. Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses. *JASP*, 2009.

[5] D. S. Brungart and W. M. Rabinowitz. Auditory localization of nearby sources. Head-related transfer functions. *JASA*, 106(3):1465–1479, 1999.

[6] T. Lentz. Near-Field HRTFs. In *DAGA*. 2007.

[7] B. Bernschütz, C. Pörschmann, S. Spors and S. Weinzierl. Entwurf und Aufbau eines variablen sphärischen Mikrofonarrays für Forschungsanwendungen in Raumakustik und virtual Audio. In *DAGA*. 2010.

[8] Z. Schärer and A. Lindau. Evaluation of equalization methods for binaural signals. In *126th AES Conv.* 2009.