Data set: Eigenmike-DRIRs, KEMAR 45BA-BRIRs, RIRs and 360° pictures captured at five positions of a small conference room

Christian Schneiderwind[1], Annika Neidhardt[1], Stefan Fichna [2], Florian Klein[1]
[1] *Technische Universität Ilmenau, 98694 Ilmenau, Deutschland, Email: christian.schneiderwind@tu-ilmenau.de*
[2] *Carl von Ossietzky Universität , 26129 Oldenburg, Deutschland, Email: stefan.fichna@uni-oldenburg.de*

Abstract

A data set was created to study the position dependent perception of room acoustics in case of a small conference room. Binaural room impulse responses (BRIRs) were measured with a KEMAR 45BA head-and-torso-simulator at 5 different potential listening positions. To achieve comparable direct sound conditions the Genelec 1030A two-way loudspeaker was always placed at a distance of 2.5 m to the front. In one scenario, the loudspeaker was directed towards the listening position. In the other scenario it was turned by 180° to realize an indirect reproduction with low direct sound energy. Furthermore, an mh acoustics Eigenmike was placed at each of the five listening positions to record 32-channel directional room impulses responses (DRIRs). Omnidirectional room impulse responses (RIR) were measured where the center of the head was placed before.
Additionally, 360° visual footages were captured with a GoPro Omni spherical camera array to provide audiovisual impressions of the listening situation at the five positions. The data set is documented in detail and is freely available for download.

Introduction

In order study the perception of the room acoustical deviations between various listening positions within one room, a data set providing equal direct sound but differences in the early reflection pattern is of interest.

Shinn-Cunningham and Ram [1] conducted a study with static binaural synthesis using BRIRs captured at different positions in the room for equal relative source positions. The results indicate a limited sensitivity to local acoustical differences and the corresponding temporal and spatial arrangement of the early reflections. The room chosen for that study was an ordinary class room with the size of 5 m × 9 m × 3.5 m with a reverberation time of T_{60} = 700 ms.

The room selected for the presented data set was chosen because its acoustical properties were very similar to that used by Shinn-Cunningham and Ram. The data set aims at providing the required acoustical information for psychoacoustic experiments using dynamic binaural synthesis with interactive head rotation and a detailed physical comparison of the five positions. Furthermore, the visual perspective was captured to enable an audiovisual 360° reproduction.

The data set includes:

- BRIRs measured with a Kemar 45BA with an azimuth resolution of 5°
- DRIR measured with an mh acoustics Eigenmike
- Omnidirectional RIRs measured at the center of the head/ center of the Eigenmike position
- Visual 360° footage captured with a GoPro Omni spherical camera array

Small Conference Room

A small conference room with the size of 10.3 m × 5.8 m × 3.1 m, was chosen for the study. It has a volume of V = 182.3 m^3 and the reverberation time is T_{60}= 0.63 s (broadband). Tab. 1 provides an overview of the T_{60} in the different frequency bands. The critical distance for omnidirectional sound sources is about r_{crit} = 0.95 m. Fig. 1 shows the room with the setup for the BRIR measurement at one of the positions.

Figure 1: Measurement of BRIRs at position 3 according to Fig. 2.

Measurement Positions

Five potential listening positions in the room were chosen with the goal to consider different local room acoustical situations. In order to keep the direct sound as similar as possible, the loudspeaker was placed in exactly the same relative position with a distance of 2.5 m in front of each of the listening positions. The lower of the two loudspeaker chassis (center) was located at a height of

Table 1: Reverberation time T_{60} at different frequencies calculated from the omnidirectional room impulse responses at the five positions for both loudspeaker orientations

Frequency	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz
$\mathbf{T_{60}}$ for LS towards HATS	567 ms	520 ms	449 ms	596 ms	676 ms	616 ms
$\mathbf{T_{60}}$ for LS turned by 180°	588 ms	523 ms	454 ms	622 ms	748 ms	700 ms
$\mathbf{T_{60}}$ mean value	**577 ms**	**522 ms**	**451 ms**	**609 ms**	**712 ms**	**658 ms**

1.47 m, the upper chassis at a height of 1.64 m. Fig. 2 illustrates the arrangement of measurement and source positions in the room. They were measured subsequently using a laser distance meter. A more detailed description of the measurement position is depicted in Fig. 6.

Furthermore, two acoustic scenarios were considered. In one direct scenario the cone of the loudspeaker is facing towards the listener position to provide a strong direct sound. Additionally a second indirect scenario with low direct sound energy was chosen to resemble an acoustic situation where the listener is not located in the main beam direction of a sound source. This was realized with the speaker being turned about 180°. With the rotation axis of the speaker being in the center of loudspeaker a slight deviation in the time of arrival of the direct between both irradiation scenarios has to be mentioned. Fig. 3 shows the omnidirectional impulse responses at position 2 for the high and low direct sound energy scenarios, respectively.

Measurement of BRIRs

The loudspeakers were positioned at each of the five intended locations subsequently. A Kemar 45BA head-and-torso-simulator was placed on an Outline ET 250-3D electronic turn table which allowed for accurate rotation. To realize an accurate orientation of the HATS, it was turned until the ITD was exactly 0 samples for the frontal position. The ears were located at a height of 1.59 m above the floor and therefore slightly below the upper chassis of the loudspeaker.

The measurement was conducted using the Swept Sine Method [2] with logarithmic sine sweeps from 60 Hz to 20 kHz over a duration of 3 s. The measurement setup was driven by a MATLAB script. The BRIRs were determined with a 5° azimuth resolution for 360°.

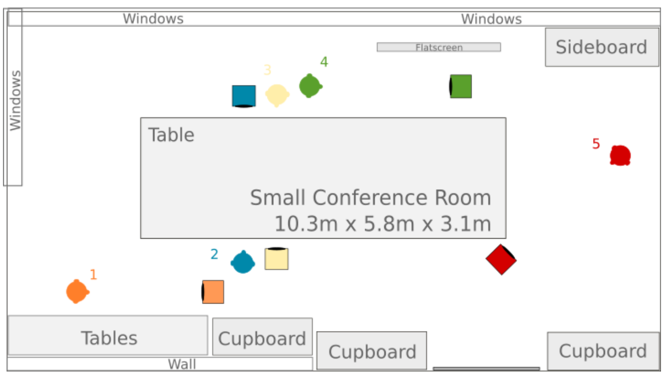

Figure 2: Illustration of the five measurement positions with their corresponding loudspeaker position. Auditory and visual perspectives were captured.

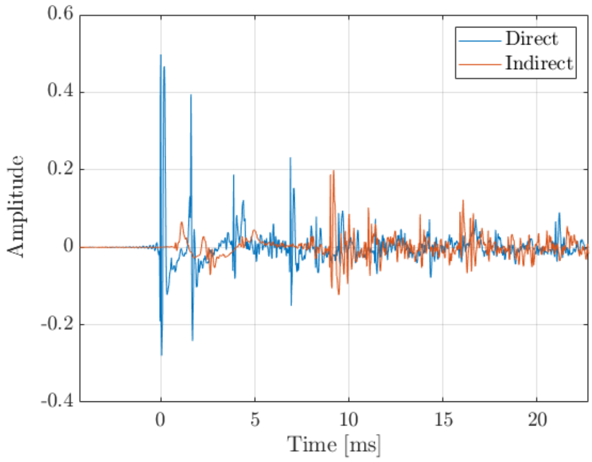

Figure 3: A plot of the omnidirectional impulse responses of position 2 with strong direct sound energy (blue) and low direct sound energy (red).

Measurement of DRIRs and omnidirectional RIRs

In order to be able to analyze the early reflection patterns, DRIRs with an spherical microphone were measured. For those measurements the mh acoustics Eigenmike was placed at the center of the head at each listening position. Concerning the calibration weights, the Eigenmike's internally stored calibration values were used.
The power maps in Fig. 4 show the energy distribution of the early reflections power maps for position 1 and 2, respectively. The energy was calculated over a time window of 20 ms shortly after the arrival of the direct sound. A detailed description of this procedure will be provided in [3].

Omnidirectional RIRs were captured at the center of the head position with a Microtech Gefell MK221 microphone capsule and a MV203 amplifier unit.

The underlying parameters regarding the excitation signal and the measurement process were chosen to be the same for the DRIR and RIR measurements in this data set with a sampling rate of 48 kHz and a logarithmic swept sine from 40 Hz to 20 kHz over a duration of 11 s as an excitation signal.

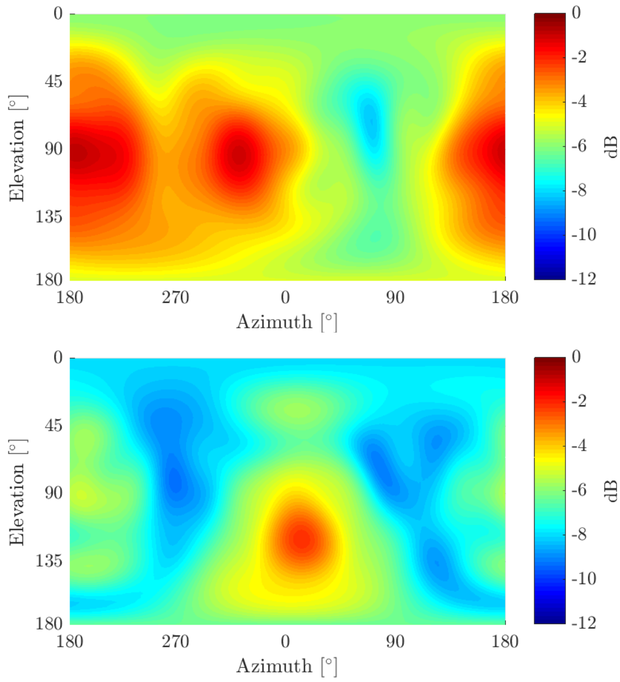

Figure 4: Powermaps calculated via plane wave decompositions for position 1 (top) and 2 (bottom) for the strong direct sound scenario.

Postprocessing

All impulse responses were shortened to 1 s (44100 samples for BRIRs and 48000 samples for DRIRs and RIRs) and multiplied with half a tukey window to create a smooth fade-in and multiplied with a window signal to create a smooth fade-in. The window signal consisted of half a tukey window (40 samples) with a zero-padding of 20 samples in front. A fade-out was not performed as there were no audible artifacts perceived as the data was auralized for testing purposes. Nonetheless, this can be done in hindsight by the user if it is desired.
In order to compensate for errors in the low frequency information a minimum-order FIR highpass filter was employed. The filters passband frequency was set to 60 Hz with a stopband attenuation of 60 dB at 30 Hz.

Visual Footage

In addition to the capturing of the acoustics at these positions visual footages were created in order to allow for audio visual testing scenarios. At each of the five listening positions the visual perspective was captured with a GoPro Omni spherical camera array. The images recorded with each of the six GoPro cameras were combined to 360° representation by stitching with the software Autopano from Kolor. Fig. 5 shows the visual perspective from listening position 3.

Related Psychoacoustic Studies

So far, two studies investigated, whether listeners are able to identify the position only by acoustical information presented over dynamic binaural synthesis using the measured BRIRs.

Neidhardt [4] provided photographs of the different positions and the illustration shown in Fig. 2 for the participant to get an idea of the room and the potential listening perspectives.

Klein et al. [5] investigated the training effect on the perception of position dependent room acoustics. For this purpose, subjects were trained to distinguish the acoustics at different listening positions and to detect mismatches between the visual and the acoustical representations.

An additional paper investigating the physical and perceptual differences between the five positions will be published [3].

Download

The data set can be downloaded following this link [6]:

zenodo.org/record/2593714

Acknowledgement

We want to thank Matthias Döring and Torsten Demmler for their technical support. This work was funded by the German Research Council (DFG BR 1333/18-1). During the course of this investigation Stefan Fichna wrote his bachelor thesis at Technische Universiät Ilmenau.

References

[1] B. Shinn-Cunningham and S. Ram. Identifying where you are in the room: Sensitivity to room acoustics. In *Proceedings of the 2003 International Conference on Auditory Display, Boston, MA, USA, July 6-9*, 2003.

[2] A. Farina. Simultaneous measurement of impulse response and distortion with a swept-sine technique. In *Audio Engineering Society Convention 108, Paris, France, February 19-22*, 2000.

[3] C. Schneiderwind and A. Neidhardt. Perceptual differences of position dependent room acoustics in a small conference room. submitted at the International Symposium on Room Acoustics, Netherlands, Amsterdam, 2019.

[4] A. Neidhardt. Perception of the reverberation captured in a real room, depending on position and direction. In *22nd Int. Congress on Acoustics (ICA), Buenos Aires, Argentina*, 2016.

[5] F. Klein, A. Neidhardt, M. Seipel, and T. Sporer. Training on the acoustical identification of the listening position in a virtual environment. In *143th Int. Convention of the Audio Eng. Soc., New York, USA*, 2017.

[6] A. Neidhardt, C. Schneiderwind, F. Klein, and S. Fichna. Dataset: Eigenmike-DRIRs, KEMAR 45BA-BRIRs, RIRs and 360 pictures captured at five positions of a small conference room (Version V1.0) [Data set]. Technical report, 2019. Zenodo. http://doi.org/10.5281/zenodo.2593714.

Figure 5: 360° visual footage captured at position 3 according to Fig. 2.

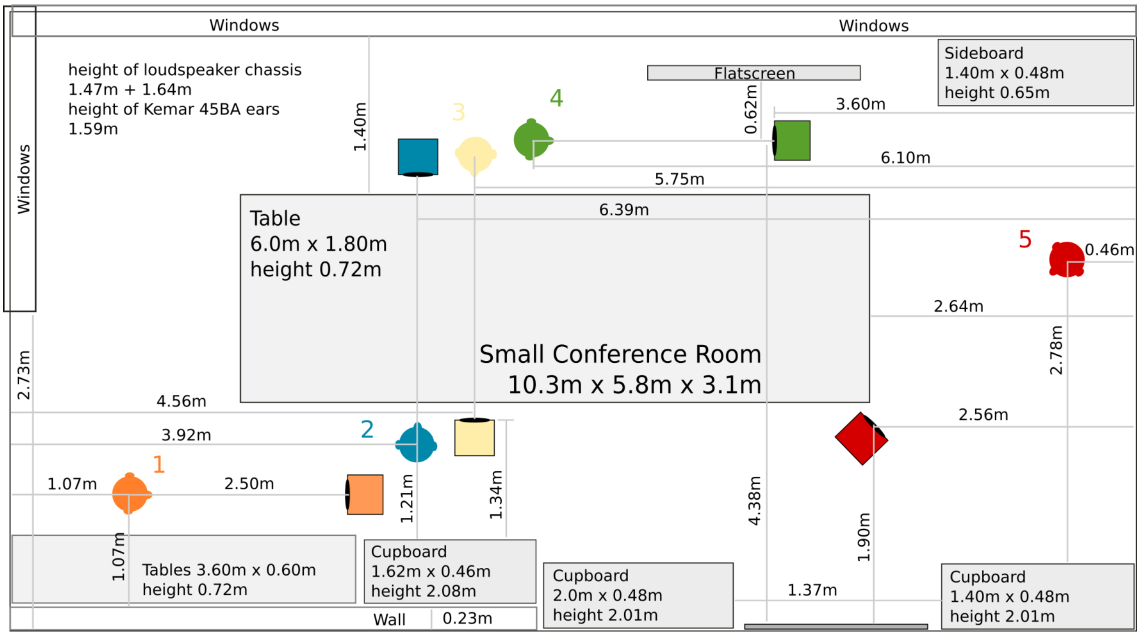

Figure 6: Exact measures of the room, relevant furniture, the chosen listening and source positions.