Dataset of Binaural Room Impulse Responses at Multiple Recording Positions, Source Positions and Orientations in a Real Room

Christina Mittag[1,2], Stephan Werner[1], Martina Böhme[1], Florian Klein[1]
[1] *Technische Universität Ilmenau, Ilmenau, Germany; E-Mail:stephan.werner@tu-ilmenau.de*
[2] *now with Fraunhofer Institute for Integrated Circuits, Erlangen, Germany.*

Introduction

Spatial audio reproduction systems enables the listener to perceive a real or a virtual environment. One aim of such systems is a plausible and immersive synthesis of an auditory scene. A binaural synthesis system is one technology to create such impressions. The system uses binaural room impulse responses (BRIRs) to recreate the correct sound signal at the ears of a listener. Figure 1 gives a schematic overview of such a system.

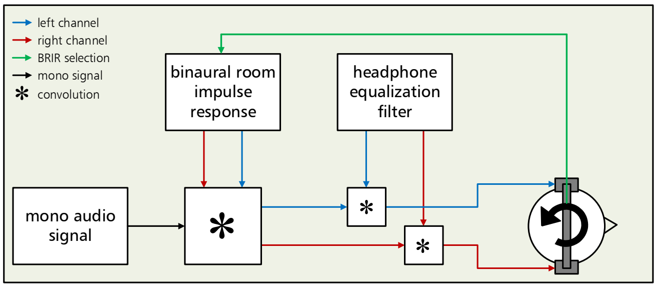

Figure 1: Scheme of a dynamic binaural synthesis system using headphones.

The BRIRs represent the transfer characteristics from a sound source in a room to the left and right ear of a listener at the listening position in the room. The BRIRs can be simulated using an acoustic room simulation or the BRIRs can be recorded in a real room using individual ear recordings or a dummy head. Several head positions have to be considered if head movement is allowed. Furthermore, several source to listener positions have to be considered if a movement of source and/or listener is intended.

Extensive research is conducted to identify the technical requirements for the recording and playback of binaural signals [1, 2, 3]. Previous research has also shown, that the perceived quality of binaural headphone systems is highly dependent on the context of usage. In the case of virtual acoustic environments, the listener is regularly forced into an acoustically different environment compared to the actual listening room. The conflicting room acoustics of the synthesized and the listening room result in the so-called room divergence effect which decreases externalization [4, 5, 6]. Other context related parameters are for example the presence of visual cues, the complexity of the acoustical scene and auditory training effects [7]. The effects are based on the fact, that the quality formation process relies on expectations formed by prior listening experience. These expectations can be changed for example by prior sound exposure or by altering visual cues.

Motivation

The motivation for the presented BRIR recordings is to provide a dataset of recorded BRIRs of a real room to exclude possible room simulation artefacts. The recordings are done for several receiver positions, for two sound source positions and for 360° head rotation at each recording position. The dataset can be used as a basis for the development and evaluation of BRIR interpolation methods [8] or room acoustic simulations. Furthermore, the dataset provides BRIRs for a dynamic binaural synthesis of a real room at different positions.

BRIR Recordings

The recordings are conducted in an office-like room at Technische Universität Ilmenau. A KEMAR head and torso simulator is used as dummy head. Two Geithain MO-2 loudspeakers are used as sound sources at two positions in the room. A logarithmic sine sweep is used as measurement signal to calculate the room impulse response between 50 Hz and 18 kHz. The recordings cover an equidistant grid in the room. Twenty-five receiver positions are recorded. A measurement of 360° is conducted at each position with an angle resolution of 5° using a turn-table. Figure 2 shows the scheme of the recording room and positions.

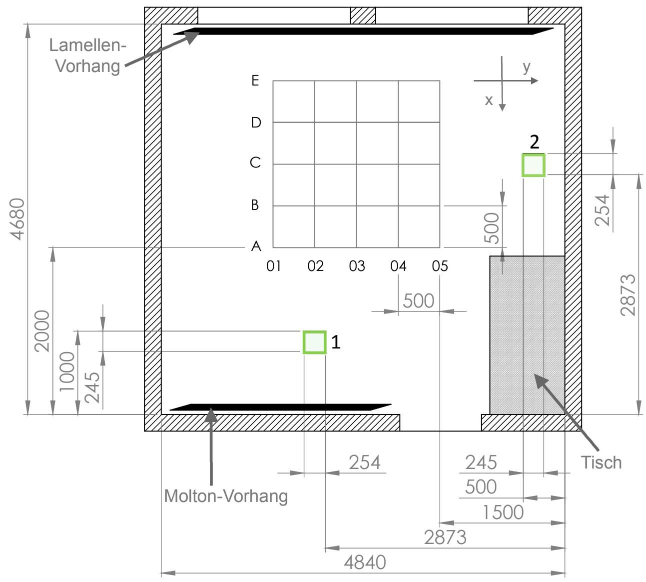

Figure 2: Scheme of the used room, measurement grid, and source positions; green rectangles indicate source positions; figure from [8].

Dataset in an Overview

The Direct-to-Reverberant-energy-Ratio (DRR) is used to illustrate the different recording positions and head orientations. The DRR describes the energy ratio between the direct sound and the reflected, absorbed, and diffracted sound in a room. In the presented case the direct sound is defined as sound which reaches the ears up to 1.5 ms after the undisturbed sound. This time span includes monaural and binaural cues resulting from reflexions and diffractions of the outer ear, head, and torso but not from the room. The sound waves reflected on the room surfaces are called reverberant sound. This sound is delayed and reaches the listening point via several paths. The reverberant sound is defined as sound reaching the listening point after 1.5 ms after the first sound in the presented case. The energy ratio between direct and reverberant sound is described by:

$$DRR = 10\log\left(\frac{\int_o^T h^2(t)dt}{\int_T^\infty h^2(t)dt}\right),$$

with *h(t)* as impulse response between two points in an enclosure and T = 1.5 ms to separate direct sound and room reflections.

Figure 3 shows the DRRs for each recording position and 360° head rotation for the left and right ear for loudspeaker position one. The solid lines indicate positive DRR up to +20 dB. The dashed lines indicate negative up to -20 dB.

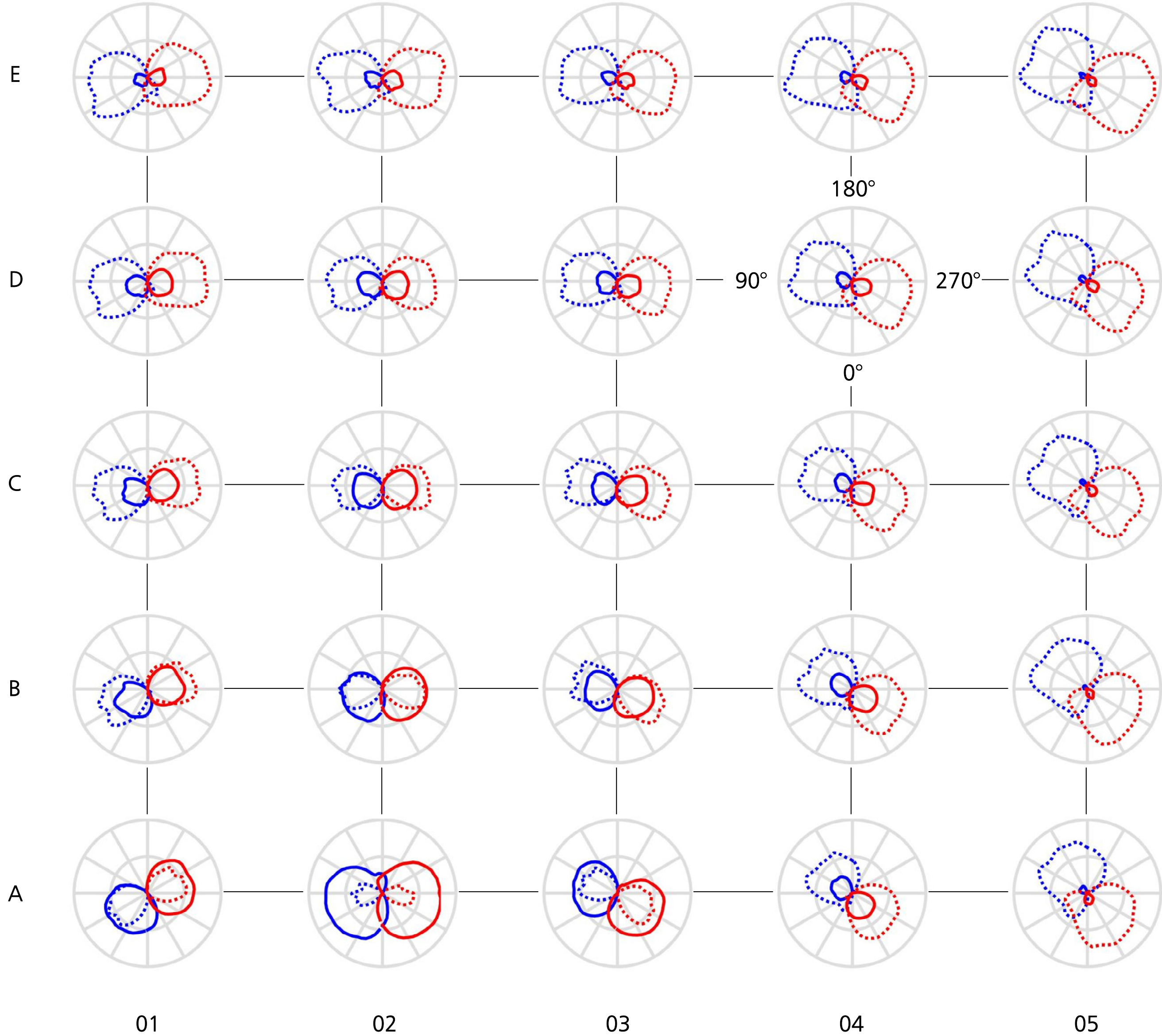

Figure 3: Direct-to-Reverberant-energy-Ratio of the left (red curves) and the right (blue curves) ear of the dummy head for all recording positions, head orientations, and for sound source 1; solid lines indicate positive DRRs, dashed lines indicate negative DRRs; outer radius indicates +/- 20 dB, midpoint 0 dB.

Download

The BRIR dataset is provided for download and underlies the license Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0). The download of the BRIR dataset can be found here:

https://doi.org/10.5281/zenodo.206860

Acknowledgment

This research is supported by a grant of the Deutsche Forschungsgemeinschaft (Grant BR 1333/14-1) and by Thüringer Aufbaubank (2015FGR0090) and the European Social Fund.

References

[1] Silzle, A., "Generation of Quality Taxonomies for Auditory Virtual Environments by Means of Systematic Expert Survey", Ph.D. thesis, Ruhr-Universität Bochum, Shaker Verlag GmbH, Germany, ISBN-13: 978-3832270407, 2007.

[2] Völk, F., "Interrelations of Virtual Acoustics and Hearing Research by the Example of Binaural Synthesis", Ph.D. thesis, Technische Universität München, Verlag Dr. Hut, München, ISBN 978-3843911139, 2013.

[3] Lindau, A., "Binaural resynthesis of acoustical environments - Technology and perceptual evaluation", Ph.D. thesis, Technische Universität Berlin, Fakultät I - Geisteswissenschaften, http://dx.doi.org/10.14279/depositonce-4085, 2014.

[4] Plenge, G., "[The problem of in-head lokalization] Über das Problem der Im-Kopf-Lokalisation," Acustica, 26(5), pp. 241–252, 1972.

[5] Werner, S., Klein, F., Mayenfels, T., and Brandenburg, K., "A Summary on Acoustic Room Divergence and its Effect on Externalization of Auditory Events," in 8th International Conference on Quality of Multimedia Experience (QoMEX), DOI 10.1109/QoMEX.2016.7498973, Portugal, 2016.

[6] Gil-Carvajal, J., Cubick, J., Santurette, S., and Dau, T., "Spatial Hearing with Incongruent Visual or Auditory Room Cues," Nature Scientific Reports DOI:10.1038/srep37342, 2016.

[7] Werner, S. and Klein, F., "Context and Plausibility in Binaural Synthesis," in 2nd International Conference on Spatial Audio (ICSA), Erlangen, Germany, 2014.

[8] Mittag, C., "Entwicklung und Evaluierung eines Verfahrens zur Synthese von binauralen Raumimpulsantworten basierend auf räumlich dünn besetzten Messungen in realen Räumen", Master Thesis, Fachgebiet elektronische Medientechnik, Technische Universität Ilmenau, 2016.