Database of binaural room impulse responses of an apartment-like environment

Fiete Winter[1], Hagen Wierstorf[2], Ariel Podlubne[3], Thomas Forgue[3], Jérôme Manhès[3], Matthieu Herrb[3], Sascha Spors[1], Alexander Raake[2], and Patrick Danès[4]

[1]*Institute of Communications Engineering, University of Rostock, Rostock, D-18119, Germany*
[2]*Audiovisual Technology Group, Technische Universität Ilmenau, Ilmenau, D-98693, Germany*
[3]*LAAS-CNRS, Université de Toulouse, CNRS, Toulouse, France*
[4]*LAAS-CNRS, Université de Toulouse, CNRS, UPS, Toulouse, France*

Correspondence should be addressed to Fiete Winter (`fiete.winter@uni-rostock.de`)

ABSTRACT

We present a database of binaural room impulse responses (BRIRs) measured in an apartment-like environment. The BRIRs were captured at four different sound source positions, each combined with four listener positions. A head and torso simulator (HATS) with varying head-orientation in the range of $\pm78°$ with $2°$ resolution was used. Additionally, BRIRs of 20 listener positions along a trajectory connecting two of the four positions were measured, each with a fixed head-orientation. The data is provided in the Spatially Oriented Format for Acoustics (SOFA) and it is freely available under the Creative Commons (CC-BY-4.0) license. It can be used to simulate complex acoustic scenes in order to study the process of auditory scene analysis for humans and machines.

1 Introduction

One of the outstanding capabilities of the human auditory system is to recover information on single auditory objects out of a mixture of sounds [1]. The EU FET-OPEN project TWO!EARS aims at developing a computational model that mimics this behavior. Listeners are regarded as multi-modal agents that develop their concept of the world by active, exploratory sensing. In the course of this process, they interpret percepts, applying existing and collecting new knowledge and concepts accordingly. Consequently, the TWO!EARS model involves bottom-up (signal-driven) as well as top-down (hypothesis-driven) processes.

A first open-source release of the model software is available for download [2]. It will be complemented until the end of the project so as to provide a comprehensive set of functions enabling new perspectives, including the neighbouring field of robot audition. The model is deployed on a binaural robot, in order to assess its suitability in real dynamic auditory scene analysis (DASA) scenarios. In addition, a software simulation stage of the robot is provided to foster the development process.

The synthesis of ear signals is thus an important basis for the development and evaluation of the TWO!EARS model. It allows to generate reproducible conditions in contrast to slightly varying signals from real-world scenarios. Binaural synthesis using Head-Related Impulse Responses (HRIRs) for anechoic environments and Binaural Room Impulse Responses (BRIRs) for reverberant conditions is a versatile tool to generate
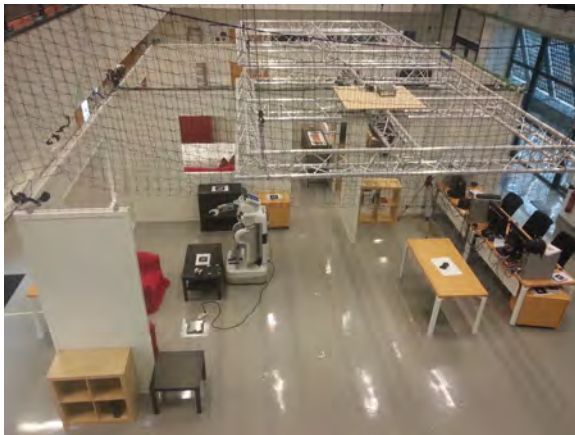
Fig. 1: ADREAM Laboratory

ear signals. For the latter, a-priori measured BRIRs describing the acoustic sound propagation from a sound source to the ears of a human listener are needed.

This engineering brief provides detailed information about the measurement process for the BRIRs of a complex, apartment-like environment. It is organised as follows: The details of the measurement process are described in Sec. 2. A brief data analysis is presented in Sec. 3 followed by information about accessibility of measured data given in Sec. 4.

2 Measurement Details

The impulse response measurements were conducted in the G. Giralt building of the Laboratory for Analysis and Architecture of Systems, *Laboratoire d'Analyse et d'Architecture des Systèmes* (LAAS-CNRS), Toulouse, France, which hosts the Reconfigurable Dynamic Architectures for Embedded Autonomous Mobile Systems, *Architectures dynamiques reconfigurables pour systèmes embarqués autonomes mobiles* (ADREAM) project. As depicted in Fig. 1, this environment constitutes a drywall installation built into in a big hall, representing an apartment with two closed rooms ($L \times W \times H = 4 \times 4 \times 2.5$ m and $3 \times 4 \times 2.5$ m) and one half-open area ($4 \times 4 \times 2.5$ m). The whole apartment has no ceiling.

2.1 Measurement Setup

The involved hardware and the data connections between the devices are depicted in Fig. 2. The BRIRs were measured using logarithmic sweeps of 2^{19} samples length at a sampling rate of 44.1 kHz. The
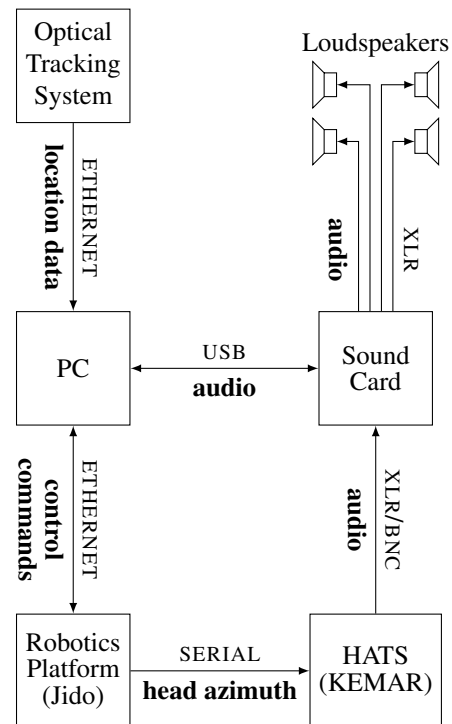
Fig. 2: Involved hardware for the measurement including data flows between the components. The type of the respective physical data connections is written in capital letters, while the data content is marked in bold font.

measurement signals are D/A-converted by a RME FIREFACE UC USB-soundcard and emitted by one out of four Genelec 8020A two-way active loudspeakers. Simultaneously, the emitted sound is measured using a Head and Torso Simulator (HATS), namely the G.R.A.S. Knowles Electronics Manikin for Acoustic Research (KEMAR) 45BB-4 [3] with Large Pinnae (KB0091 + KB0090). The length of the BRIRs was limited to 2 seconds and the delay of the measurement chain was compensated a-priori. The HATS was mounted on a mobile robotics platform, named Jido (see Fig. 3a). The manikin is furthermore endowed with a servo motor to enable azimuthal head rotations. The motor is associated to a gearhead (ratio of 9) and a quadrature incremental encoder with 2048 steps (4 pulses by step), so that the 73728 pulses per revolution ensure an accuracy of 4.88×10^{-3} degrees. It is interfaced via a serial connection through an ELMO controller that manages the sensors (Hall effect sensors, encoders and limit sensors) and controls the
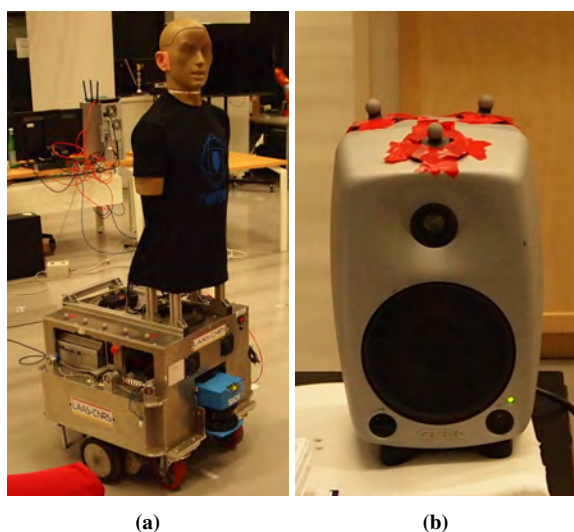
Fig. 3: (a) Jido mobile robot with the KEMAR mounted on its top. (b) Genelec 8020A Loudspeaker (×4). Both devices were endowed with optical tracking markers (small gray balls).

motor in position or velocity mode. The setpoints are sent through a standard CAN serial link.

The location[1] of the robotics platform and loudspeakers are captured by an optical tracking system composed of 28 infrared cameras. It uses small optical markers made with a retroreflective material (cf. Fig. 3) mounted on the respective devices. Its accuracy is in the millimeter range.

2.2 Measurement Positions

In the first part of the measurement, BRIRs were captured for four different loudspeaker locations, each combined with four listener locations with a head-above-torso-orientation varying in the range of $\pm 78°$ with $2°$ resolution (see Fig. 5a). Hereby, a negative angle corresponds to turning the head to the right above the torso. While the loudspeakers remained steady over the whole measurement, the HATS had to be moved to the next location by operating the robot manually. As shown in Fig. 5c, the first loudspeaker was oriented towards a wall in order to create a case, where the direct sound has a lower amplitude than the first reflection, due to the directivity of the loudspeaker.

Within the second part, the listener was gradually moved along a trajectory connecting the measurement

---

[1]location subsumes position and orientation within this treatise

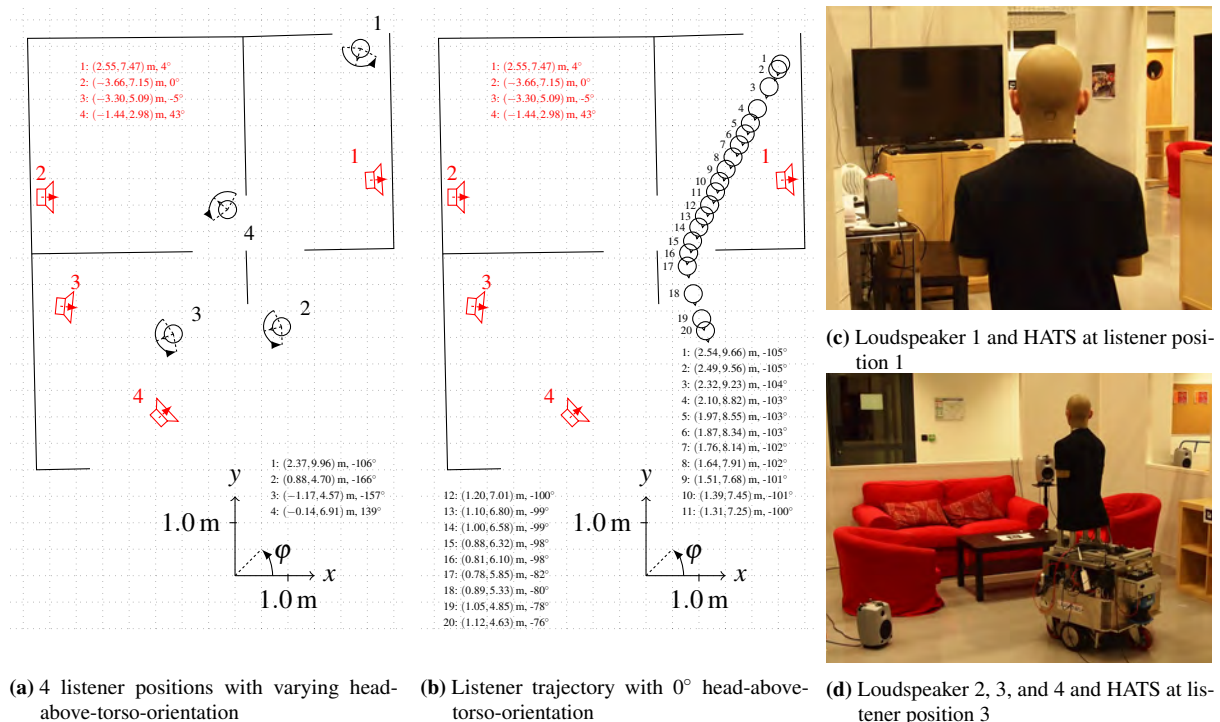**(a)** Loudspeaker 1, Listener Position 1

(b) Loudspeaker 2, Trajectory
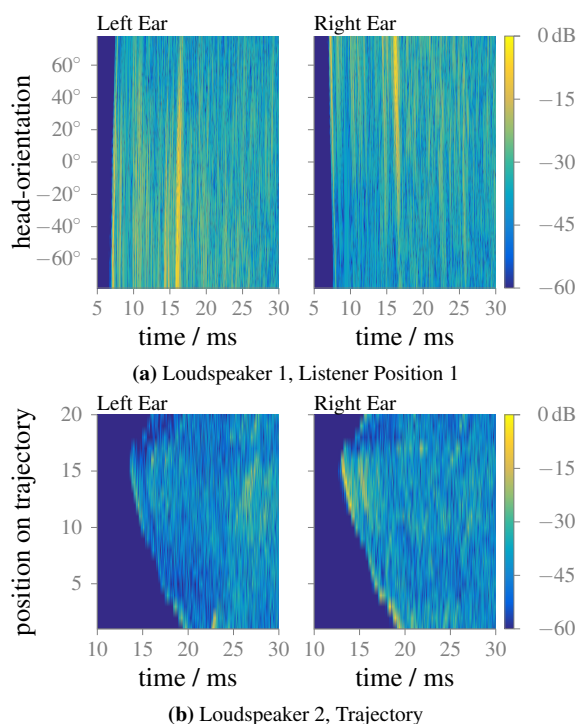
Fig. 4: Magnitude of measured BRIRs

locations 1 and 2 from the first part (see Fig. 5b). BRIRs for each loudspeaker were measured for $0°$ head-orientation every ≈ 25 cm on the trajectory. The loudspeaker positions remained the same compared to part one. During the whole measurement, that is part one and two, the individual sound pressure levels of all four loudspeakers were kept constant.

3 Data Analysis

A small excerpt of the measured impulse responses is shown in Fig. 4 in order to illustrate how the complex environment is affecting the structure of the BRIRs. The directivity of the loudspeakers and the orientation of the first loudspeaker towards the wall leads to the expected strong reflection, which is considerably higher than the direct sound (cf. 4a). The influence of the listener's head is also clearly visible, as the amplitude of the reflection is reduced for the contra-lateral ear. As the listener "moves" along the trajectory and passes the doorframe connecting both closed rooms (cf. location 11 to 16 in Fig. 5b), loudspeaker 2 is no longer occluded by the walls and the direct sound for the right ear (cf. Fig. 4b) significantly increases.

(a) 4 listener positions with varying head-above-torso-orientation

(b) Listener trajectory with 0° head-above-torso-orientation

(c) Loudspeaker 1 and HATS at listener position 1

(d) Loudspeaker 2, 3, and 4 and HATS at listener position 3

Fig. 5: Fig. (a) and (b) show the *xy*-position and horizontal orientation of the listener's torso drawn in black. Respective quantities for the four loudspeakers are drawn in red. The legend shows the respective position followed by the orientation's azimuth angle φ. In addition, an arc around the listener symbol is used in (a) to illustrate the varying head-above-torso-orientation in the range of $\pm 78°$. Fig. (c) and (d) show exemplary combinations of the listener and loudspeaker locations from (a).

4 BRIR Database

All measured BRIRs are stored in the Spatially Oriented Format for Acoustics (SOFA), which is the standard format for spatial acoustic data of the Audio Engineering Society [4]. The data accompanied by photos from the setup is freely available at [5] and is licensed under Creative Commons Attribution 4.0 (CC BY 4.0)[2].

5 Acknowledgements

References

[1] Cherry, E. C., "Some experiments on the recognition of speech, with one and with 2 ears," *J. Acoust. Soc. Am.*, 25, pp. 975–979, 1953.

[2] Two!Ears Project, "Two!Ears Auditory Model 1.2," 2016, doi:10.5281/zenodo.47487.

[3] G.R.A.S. Sound & Vibration, "Instruction Manual - G.R.A.S. 45BB KEMAR Head and Torso / 45BC KEMAR Head and Torso with Mouth Simulator," 2016.

[4] Audio Engineering Society, Inc., "AES69-2015 - AES standard for file exchange - Spatial acoustic data file format," 2015.

[5] Winter, F., Hagen, W., Podlubne, A., Forgue, T., Manhès, J., Herrb, M., Spors, S., Raake, A., and Danès, P., "Binaural room impulse responses of an apartment- like environment," 2016, doi:10.5281/zenodo.49357.

---

[2]http://creativecommons.org/licenses/by/4.0/