Measurement of a Binaural Room Impulse Response with High Peak-to-Noise Ratio

Data Processing and Documentation: Wiebke Hahne, Sascha Spors
Measurements: Bassem Zaitoon, Vera Erbes, Sascha Spors

November 23, 2021

1 General Information

Experiment title	Measurement of a Binaural Room Impulse Response with High Peak-to-Noise Ratio
Experiment id	high-PNR-BRIR
Principal investigators	Bassem Zaitoon, Vera Erbes
Affiliation	Institute of Communications Engineering, University of Rostock, Germany
E-Mail	acoustics@comlab.uni-rostock.de
Date	November 23, 2021
Type	data measurement and data preprocessing
Location	Richard-Wagner-Str. 31 (Haus 8), 18119 Rostock
Keywords	BRIR measurement
Language	English
Rights	CC BY 4.0

1.1 Objective

Generation of a binaural room impulse response with high peak-to-noise ratio by averaging measurements.

https://doi.org/10.18453/rosdok_id00002986

1.2 Problem Statement

During their measurement, binaural room impulse responses (BRIRs) are impaired by equipment and background noise. The resulting noise floor might become perceptible when BRIRs are used for auralization purposes, i.e. by convolution with a dry source. To avoid the audibility of noise, high peak-to-noise ratio BRIRs are desired. A high peak-to-noise ratio (PNR) can be computed by averaging of multiple measured BRIRs. With this measurement at least a 90 dB PNR-BRIR is aspired. This data set was measured in context of the master thesis by Bassem Zaitoon [1].

2 Data Description

The published data set contains the sweep signal (*.wav*) as excitation signal, the captured room responses as raw-data (*.wav*) which are measured from the frontal and the lateral side in two sessions each and the resulting binaural room impulse responses (*.sofa*). Additionally, the computed averaged high-PNR-BRIRs (*.sofa*) and the high-PNR-Raw-Data (*.wav*) are included.

Data:

- LogSweepFIFO.wav — excitation signal
- *yyyymmdd*_BRIR.sofa — individual BRIRs of measurement session on date *yyyymmdd*, 600 individual measurements, sofa-format after AES69-2020 [2]: AES standard for file exchange - Spatial acoustic data file format
- *yyyymmdd*_BRIR_highPNR.sofa — high-PNR-BRIR of measurement session on date *yyyymmdd* sofa-format after AES69-2020
- ***yyyymmdd*_RAW**:
  - *yyyymmdd*_RAW_*xxx*.wav — raw measurement data of date *yyyymmdd*, wav-File with 44.1 kHz sampling rate and 24 bit resolution, individual measurement *xxx* (000-599)

- *yyyymmdd*_RAW_inventory.csv — meta data of raw measurement wav-files,
  columns:
  index, Filename, Date [yyyy-mm-dd], Time [HH:MM:SS], Temperature [K]

- ***yyyymmdd*_RAW_highPNR**:
  - *yyyymmdd*_RAW_highPNR.wav — averaged raw measurement data of date *yyyymmdd*,
    wav-File with 44.1 kHz sampling rate and 24 bit resolution
  - *yyyymmdd*_RAW_inventory.csv — meta data of averaged raw measurement wav-files,
    columns:
    index, Filename, Date [yyyy-mm-dd], Time [HH:MM:SS]

Python:

- Generate_HighPNR_BRIR_frontal.ipynb — Jupyter notebook for the generation of the frontal orientated high-PNR-BRIR
- Generate_HighPNR_BRIR_lateral.ipynb — Jupyter Notebook for the generation of the lateral orientated high-PNR-BRIR
- functions.py — python-file which includes functions for the Jupyter notebooks

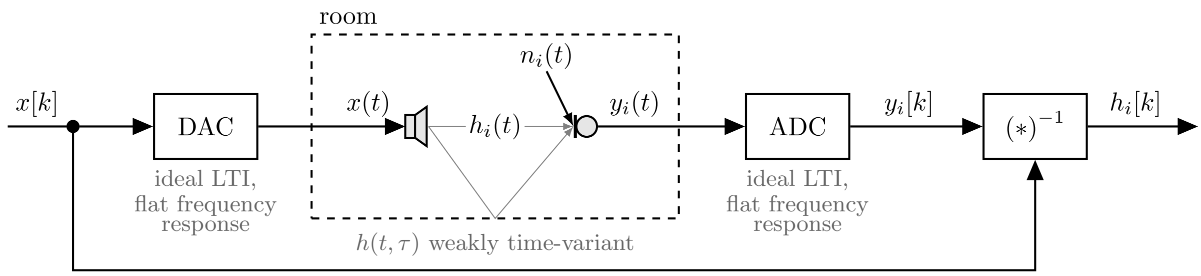


Figure 1: Illustration of the assumed signal model for measurement of BRIRs. For ease of illustration only one of the two channels of the recorded binaural response is visualized.

3 BRIR Measurements

The left/right continuous BRIR $h_{\mathrm{L,R}}(t)$ captures the acoustic room impulse response from a source to a pre-defined position in the ear canal. For ease of illustration, the indexes L and R for the left and right BRIR are omitted in the following. The impairment of BRIRs by noise takes place in the measurement process. Therefore, a common signal model is introduced in a first step. Figure 1 illustrates the assumed signal model for the measurement of BRIRs. After digital analogue conversion (DAC) a specific excitation signal $x(t)$ is played back by a loudspeaker in the recording room. The continuous BRIR is in general assumed to be only weakly time-variant, e.g. due to slight changes in temperature. The resulting temporal variances are reflected by considering a different BRIR $h_i(t)$ for each measurement i. Background and equipment noise are modeled by the additive random signal $n_i(t)$. The captured response $y_i(t) = h_i(t) * x(t) + n_i(t)$ undergoes analogue digital conversion (ADC) and results in its discrete counterpart $y_i[k]$ with the discrete-time index k.

In this work, the finite impulse response (FIR) model serves as the discrete-time model $h_i[k]$ of the continuous BRIR $h_i(t)$. The discrete BRIR $h_i[k]$ is computed by deconvolution with the excitation signal $x[k]$. Applying the Discrete-Time Fourier Transform (DTFT) to both signals, the BRIR is given as

$$h_i[k] = \mathcal{F}_*^{-1} \underbrace{\left\{ \frac{Y_i\left(\mathrm{e}^{\mathrm{j}\Omega}\right)}{X\left(\mathrm{e}^{\mathrm{j}\Omega}\right)} \right\}}_{H_i(\mathrm{e}^{\mathrm{j}\Omega})} \tag{1}$$

where $Y_i\left(\mathrm{e}^{\mathrm{j}\Omega}\right) = \mathcal{F}_*\{y_i[k]\}$ and $X\left(\mathrm{e}^{\mathrm{j}\Omega}\right) = \mathcal{F}_*\{x[k]\}$ denote the DTFTs of the captured and excitation signal, respectively. The binaural room transfer function (BRTF) for the i-th measurement is denoted by $H_i\left(\mathrm{e}^{\mathrm{j}\Omega}\right) = \mathcal{F}_*\{h_i[k]\}$.

3.1 Procedure

While measuring it need to be quiet to have as less noise disturbance as possible. Therefore the measurements are executed over night. The BRIRs are measured in two positions of the HATS - in frontal and lateral orientation to the loudspeaker. Both BRIR measurement sessions of each orientation were executed with 600 measurements per session.

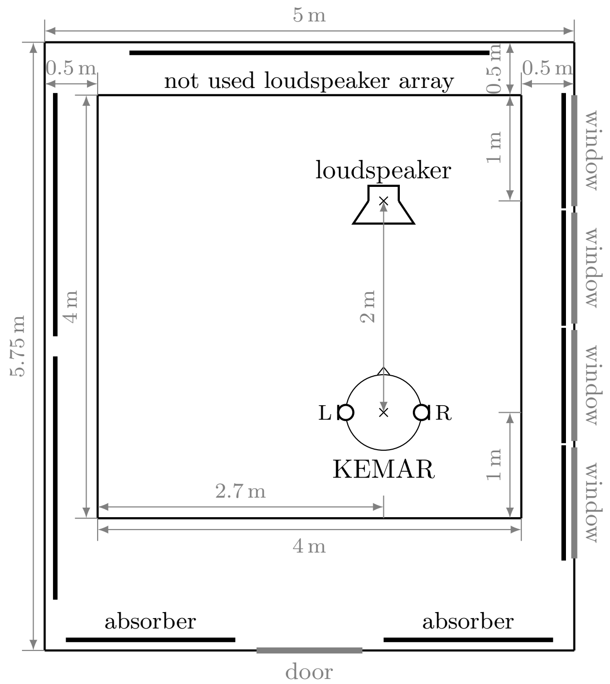


Figure 2: Schematic measurement setup in the audio laboratory of the University of Rostock.

3.2 Measurement Setup

Experimenter:	Bassim Zaitoon
Date:	frontal source, session 1: 18.10.2017 frontal source, session 2: 01.11.2017 lateral source, session 1: 02.11.2017 lateral source, session 2: 07.11.2017
Room:	Audio Laboratory University of Rostock (Germany) Size: 5.75 m × 5.0 m × 3.0 m acoustically treated, reverberation time $T_{60} \approx 0.3\,\mathrm{s}$
Software:	Matlab Version R2015a
Excitation Signal:	Exponential Sine Sweep duration $T \approx 47.6\,\mathrm{s}$ start frequency 20 Hz, stop frequency 20 kHz $f_s = 44.1\,\mathrm{kHz}$
Sound Card:	RME Fireface UC
Power Amplifier:	Brüel & Kjær, Type 2734-A

Loudspeaker:	omni power sound source Brüel & Kjær, Type 4292-L 12 drivers in a dodecahedral configuration
Dummy Head:	KEMAR manikin 45BA G.R.A.S. large pinnae (type KB0065 and KB0066) pressure microphones 40AO G.R.A.S.
Pre-Amplifier:	Lake People C360

4 High-PNR-BRIR Generation

4.1 General Principal

A high-PNR-BRIR is generated by averaging several BRIR measurements. The rationale behind averaging the BRIRs $h_i[k]$ is that if the additive noise $n_i(t)$ is uncorrelated for different measurements, the overall noise will be reduced [3]. This results in the average BRIR

$$\bar{h}[k] = \frac{1}{I} \sum_{i=0}^{I-1} h_i[k] \tag{2}$$

where I denotes the total number of individual measurements. As a measure of the noise level in a BRIR, the peak-to-noise ratio (PNR) is used [4]. It is defined as the squared maximum value of the impulse response to the noise variance σ_N^2

$$\mathrm{PNR} = 10 \log_{10} \left(\frac{\max |h[k]|^2}{\sigma_N^2} \right) \tag{3}$$

in dB. A single impulse response $h_i[k]$ has a PNR of approximately 68 dB.

The repeated measurement of BRIRs over an extended period of time might be subject to outliers. For instance by bursts of ambient noise. For detection of such outliers, the normalized system distance between adjacent BRIR measurements is used [5]. It is defined as

$$D_i = \sqrt{\frac{\sum\limits_{k=0}^{K-1} |h_i[k] - h_{i-1}[k]|^2}{\sum\limits_{k=0}^{K-1} |h_{i-1}[k]|^2}}. \tag{4}$$

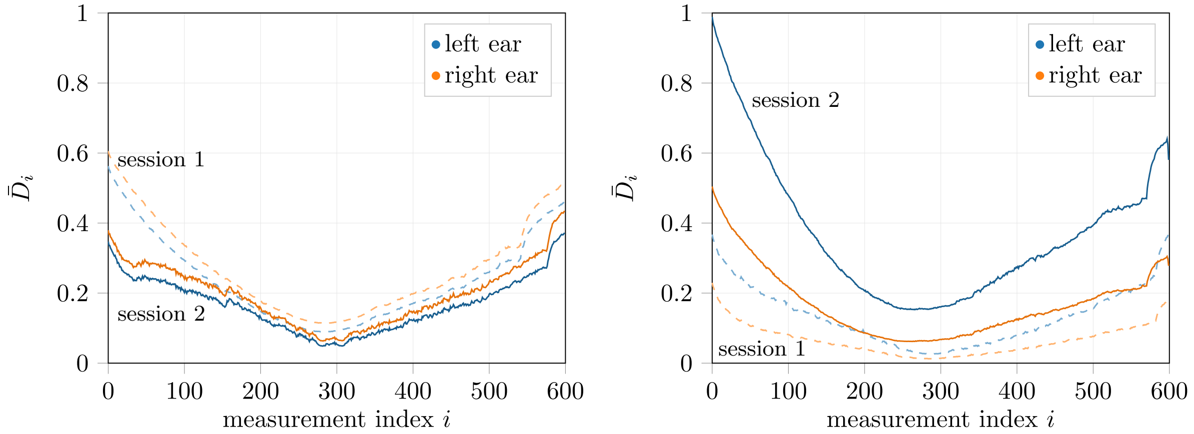
Hereby, $h_i[k]$ denotes the i-th measured BRIR and K the total length of the impulse response. If the system distance D_i between measurement i and $i-1$ exceeds a pre-defined threshold, the measurement $h_i[k]$ is neglected for the computation of the average in (4). The system distance calculation proceeds then with the next measurement $h_i[k] \rightarrow h_{i+1}[k]$.

4.2 Selection Criterion of the Measurement Session

The decision about the less time-variant measurement session is made by a modification of the system distance in equation (4). Instead of depending on the preceding impulse response $h_{i-1}[k]$ the modified system distance

$$\bar{D}_i = \sqrt{\frac{\sum\limits_{k=0}^{K-1} |h_i[k] - \bar{h}[k]|^2}{\sum\limits_{k=0}^{K-1} |\bar{h}[k]|^2}} \tag{5}$$

depends on the average impulse response $\bar{h}[k]$ of this session. In Figure 3 the modified system distance $\bar{D}_i$ of both measurement sessions is visualized. The lower the curve the less disturbance was presented in the measurement session. A greater difference in



(a) frontal orientation (b) lateral orientation

Figure 3: The modified system distance $\bar{D}_i$ of the two BRIR measurement sessions for (a) frontal orientation and (b) lateral orientation. The lower the curves the less disturbances are included in the measurements.

the modified system distance between two measurements in Figure 3 represents a higher value in the system distance according to equation (4). The resulting bathtub curves are explainable due to the changing temperature during the night, the sunrise, starting of the heater or working people in the morning.

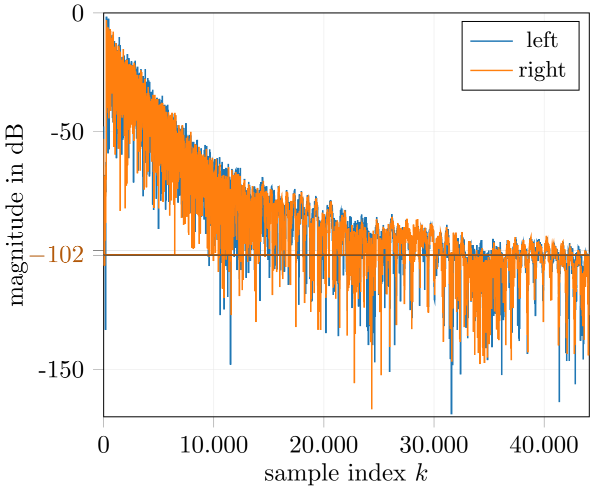
Fontal Orientation The difference between the left and the right ear in each session of the frontal orientation might be explainable due to the off center positioning of the HATS and the window front on the right side of the room. Since in measurement session 2 the bathtub curve is lower this session is used for further processing.

Lateral Orientation In the lateral orientation of the HATS, the difference between the left and the right ear of each session might be explainable due to the orientation of the ears. The left ear is facing the loudspeaker. Supposed the change in the BRIR due to temperature and environment noise in time is the same for the left and the right ear, the higher amplitudes of the right ear result in a lower modified system distance $\bar{D}_i$. Since in measurement session 1 the bathtub curve is lower this session is used for the calculation of a high PNR-BRIR

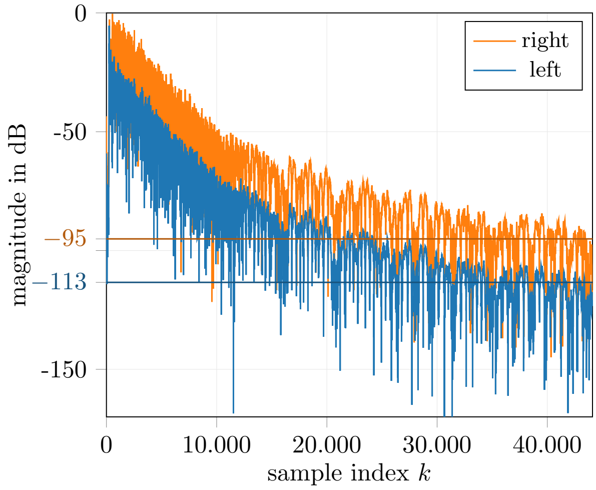
4.3 Realization

The system distance D_i is used to detect and disregard outliers in the measurement session. The first 15 measurements were disregarded for getting a good starting measurement without disturbance due to the experimenter. The realization and the results are summarized in Table 2.

Frontal Orientation Outliers with the self-defined system distance threshold of $D_i = 0.015$ appear in the last measurements when disturbance occurrence might be plausible. From the 600 measurements $I = 544$ are selected to calculate the average impulse response



(a) frontal orientated high-PNR-BRIR



(b) lateral orientated high-PNR-BRIR

Figure 4: Visualisation of the high PNR BRIR measurement with (a) frontal orientation by averaging $I = 544$ measurements and (b) lateral orientation by averaging $I = 563$ measurements.

Table 2: Realization and Result of the high PNR-BRIR

	session	**system distance threshold**	**Number selected measurements**	**resulting PNR**
frontal	2	left: 0.0155 right: 0.0155	544	left: 101.9 dB right: 101.8 dB
lateral	1	left: 0.0079 right: 0.0155	563	left: 113.4 dB right: 95.1 dB

$\bar{h}[k]$ using equation (2). This results in a PNR-level of 101.9 dB for the average impulse response $\bar{h}[k]$ of the left ear and a PNR-level of 101.8 dB for the right ear, as shown in Figure 4a.

Lateral Orientation Since the amplitudes of the frontal and the far lateral orientation are different there is a need for different thresholds of D_i. For the left ear the D_i threshold is defined as 0.0.0079 and for the right ear as 0.0155. This results in $I = 563$ measurements which are used to calculate the average impulse response $\bar{h}[k]$. Hence, the average impulse response $\bar{h}[k]$ has an PNR-level of 113.4 dB for the left ears and 95.1 dB for the right ear.

References

[1] Bassim Zaitoon. Threshold of detection of noise in a binaural room impulse response. Master's thesis, University of Rostock, March 2018.

[2] Aes standard for file exchange - spatial acoustic data file format. Standard AES69-2020, Audio Engineering Society, December 2020.

[3] Swen Müller and Paulo Massarani. Transfer-function measurement with sweeps. *J. Audio Eng. Soc*, 49(6):443–471, 2001.

[4] M Guski and M Vorländer. Comparison of noise compensation methods for room acoustic impulse response evaluations. *Acta Acustica united with Acustica*, 100(2):320–327, 2014.

[5] Nara Hahn and Sascha Spors. Comparison of continuous measurement techniques for spatial room impulse responses. In *Signal Processing Conference (EUSIPCO), 2016 24th European*, pages 1638–1642. IEEE, 2016.