A High-Resolution Spatial Room Impulse Response Database

Tim Lübeck[1,2], Johannes M. Arend[1,2], Christoph Pörschmann[1]
[1] Institute of Communications Engineering, TH Köln - University of Applied Sciences, Cologne, Germany
[2] Audio Communication Group, Technical University of Berlin, Berlin, Germany
Email: tim.luebeck@th-koeln.de

Introduction

Re-synthesizing acoustic spaces based on captured spatial room impulse responses (SRIRs) is a popular field of research. In particular, research on Ambisonics sound field en- and -decoding is often based on impulse responses (IRs) measured with a spherical microphone array (SMA). IR Measurements on high spatial-resolution sampling grids for spherical harmonics (SH) transform at high orders are barely available. One example is the IR database by Stade et al. [1], which contains SMA IRs on 29th order Lebedev grids. To our knowledge, there are no available SMA IRs measured with spatial resolutions higher than $N = 29$. However, in theory it may be reasonable to capture sound fields with higher spatial resolution. According to Rafaely [2, p. 80], when transforming spherical data to the SH domain, spatial aliasing and SH order truncation artifacts arise from about $f_a = {}^{Nc}/_{2\pi r}$). Assuming a rigid sphere array radius of $r = 0.0875\,\mathrm{m}$ and a speed of sound of $c = 343\,\mathrm{m\,s^{-1}}$, SH orders higher than $N = 32$ are necessary for artifact-free SH transformation over the entire audible bandwidth (20 Hz-20 kHz). Moreover, it has been shown that head-related transfer functions (HRTFs) have a spatial order greater than N = 40 [3]. For a binaural reproduction of SMA captures, the HRTFs and the sound field are merged with matching orders in the spatial domain. Thus, theoretically when using SMA data captured on grids with spatial orders below $N = 40$, the employed HRTF set must be limited (i.e., truncated) in its spatial order. To avoid SH order truncation or pre-processing of the involved HRTFs for binaural rendering and guarantee an artifact-free spatial sound field representation for the entire audible bandwidth, we performed SMA measurements on a Lebedev grid with a spatial order of $N = 44$. Using the automated single-microphone measurement system VariSpehar [4], we performed measurements in 3 different rooms at TH Köln for two receiver and three source positions each. Additionally, we measured binaural room impulse responses (BRIRs) with a Neumann KU100 dummy head as well as omnidirectional IRs with an Earthwork M30 microphone at the same positions.

This paper introduces the open-source database, available at [5], containing the measurements in SOFA format as well as floor plans and 3D models of the three rooms. The paper presents detailed information about the measurement setup and a basic room acoustical analysis of the measured rooms.

Measurement Setup

In all rooms, we measured RIRs for two receiver positions (Pos.1 and Pos.2) and three source positions (Loudspeaker Left, Center, and Right (LSL, LSC, LSR)). For each of the 6 source-receiver combinations, we measured a set of SMA IRs, a set of dummy head BRIRs, and an omnidirectional IR. For the SMA measurements, we employed the robotic measurement system VariSphear [4] with rigid-body SMA extension. The SMA extension consists of a wooden rigid-body of radius $r = 0.0875\,\mathrm{m}$ with an embedded Earthwork M30 microphone inside. The VariSphear system rotates the rigid-body extension with the microphone so that we could sequentially measure 2702 sampling positions of an N=44 Lebedev [6] grid along a virtual sphere. The dummy head BRIR measurements were performed for 360 directions along a horizontal circle in 1° steps. The rotation was also performed with the VariSphear motor. Finally, with the same Earthworks M30 microphone as used for the SMA measurements, we measured an omnidirectional IR.

As sound sources, we used a Neumann KH120 loudspeaker for the smaller room and a Neumann KH420 loudspeaker for the larger rooms. All measurements were post-equalized with a digital low-shelf filter with a cutoff at 100 Hz that replicates the acoustic controls for the bass range of the speakers as found in their manuals. As DA/AD converter and microphone preamplifier, we employed an RME Babyface. Data acquisition was done with the Matlab VariSphear software using the exponential sine sweep method [7]. Deconvolution was done with a measured feedback loop also compensating for the time delay and frequency response introduced by the measurement chain.

Rooms

The database includes RIR measurements for three different rooms at TH Köln's Deutz Campus, whose geometries and parameters are summarized in Table 1.

Audiolab

The Audiolab has dimensions of 7.15 m × 4 m × 2.95 m (L × W × H). It has a concrete floor and ceiling and sand-lime brick walls. Because of the small room dimensions and rather short source-receiver distances (2.8 m - 3.5 m), we employed the smaller Neumann KH120 loudspeaker for measurements in this room. The acoustical center of the source was at the same height as the receiver. The exact measurement positions are shown in

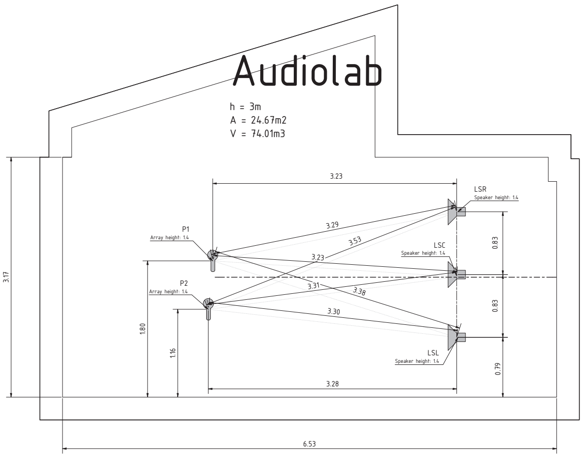

Figure 1: Floor plan of the Audiolab (room number 0-90-a-e at TH Köln - Deutz Campus), including the source and receiver positions. Source and receiver were always on the same height.

Figure 2: Photograph of the Audiolab showing the measurement setup of source position LSC and receiver position 1. It shows the VariSphear with rigid-body extension and the Neumann KH120 loudspeaker.

Figure 1. The Audiolab has an average reverberation time (RT_{60}) of 0.70 s (between 500 and 8000 Hz, averaged across all source-receiver combinations from the omnidirectional measurement).

Classroom

The Classroom is a shoebox-shaped seminar room that seats 74 persons. It has a linoleum floor, walls of plasterboard, and a ceiling with perforated ceiling panels. During the measurements, it was equipped with wooden chairs and tables as can be seen in Figure 4. Because of the larger room dimensions and larger source/receiver distances, we employed the Neumann KH420 loudspeaker. It allows to playback with a higher playback level and thus improve the signal-to-noise ratio. As in the Audiolab, the acoustic center of the source and the receivers were at the same height.

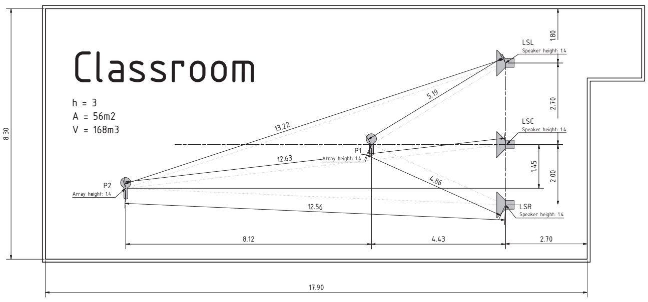

Figure 3: Floor plan of the Classroom (room number ZW-8-3 at TH Köln - Deutz Campus), including the source and receiver positions. Source and receiver were always on the same height.

Figure 4: Photograph of the Classroom showing the VariSphear at position 1 and the Neumann KH420 loudspeaker at position LSR.

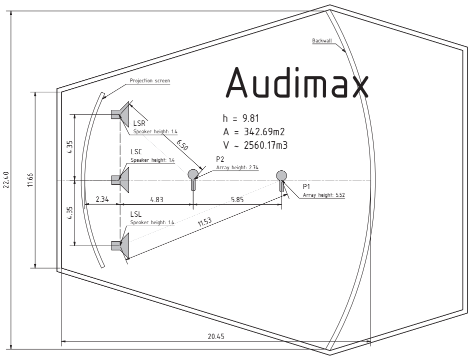

Figure 5: Floor plan of the Audimax (room number Z-2-10 at TH Köln - Deutz Campus), including the source and receiver positions.

Figure 6: Photograph of the Audimax showing the VariSphear at position 1 and the Neumann KH420 loudspeaker at position LSR. The sound source is slightly elevated downwards.

Audimax

The Audimax is a large lecture hall seating 357 persons. It has concrete walls, a linoleum floor, and a ceiling consisting of metal steps. It is equipped with wooden benches. We employed the Neumann KH420 loudspeaker, but compared to the other setups, source and receiver were at different heights, as can be seen in Figure 6.

Room Acoustic Analysis

We used the ITA Matlab toolbox [8] to calculate the room acoustic parameters reverberation time (RT_{30}), energy decay time (EDT), and interaural cross correlation (IACC). The reverberation times (Fig. 7) and the

Room	Volume	Area	RT_{60}
Audiolab	$84{,}37\,m^3$	$28{,}6\,m^2$	0.70 s
Classroom	$459\,m^3$	$153\,m^2$	0.90 s
Audimax	$2560.17\,m^3$	$342.69\,m^2$	1.024 s

Table 1: The RT_{60} values are averaged over all source-receiver positions between 500 and 8000 Hz.

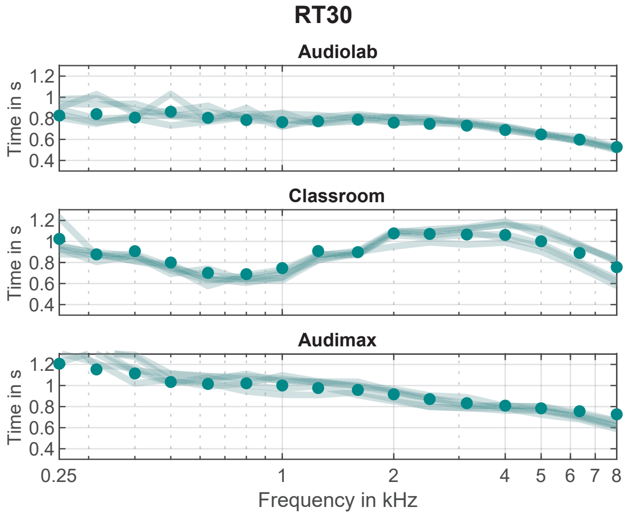

Figure 7: RT_{30} values for all rooms based on the omnidirectional impulse responses.

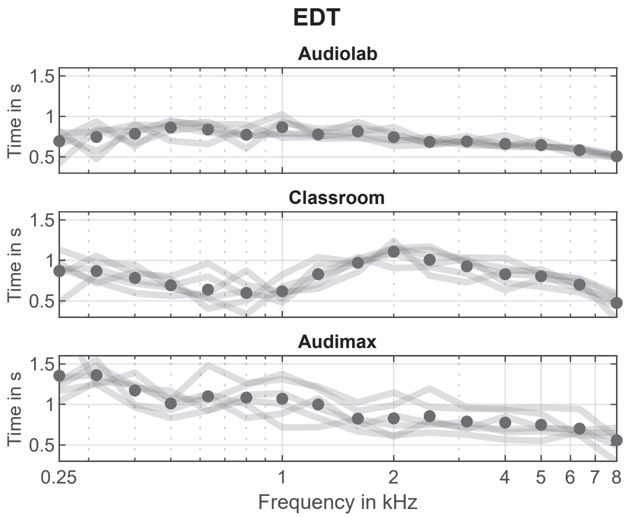

Figure 8: Energy decay times for all rooms based on the omnidirectional impulse responses.

energy decay times (Fig. 8) are based on the omnidirectional IRs. The interaural cross-correlation values (Fig. 9, bottom) are based on the dummy head measurements for frontal head orientation, which were obtained separately for each source/receiver position using onset detection. The slightly transparent lines in the figures show the values for each individual source/receiver position, the dots show the values averaged across all source/receiver positions.

Additionally, we calculated the direct-to-reverberation ratio (DRR) for each source/receiver position based on the dummy head BRIRs for frontal head orientation with a direct sound window of 2.3 ms The values are shown in Tab. 2.

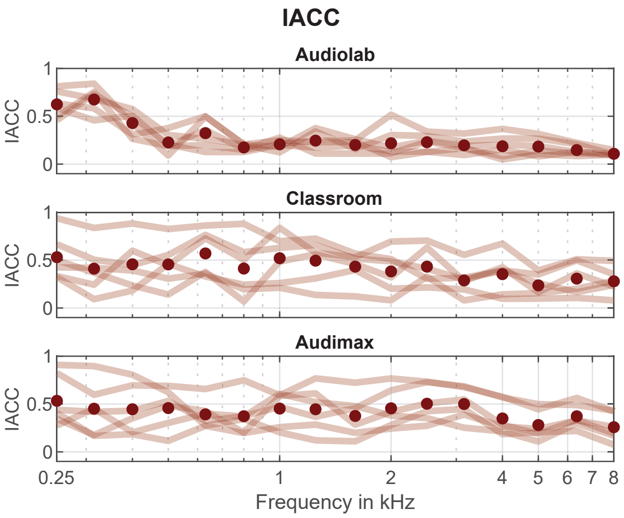

Figure 9: Interaural cross correlation values based on the dummy head BRIRs for frontal sound incidences.

Room	Pos. 1	DRR*	Pos. 2	DRR
Audiolab	LSL	-12.5	LSL	-8.38
	LSC	-11.2	LSC	-12.52
	LSR	-12.42	LSR	-16.08
Classroom	LSL	-2.01	LSL	-8.92
	LSC	0.45	LSC	-8.24
	LSR	-2.85	LSR	-3.26
Audimax	LSL	-4.64	LSL	-2.09
	LSC	-6.03	LSC	-1.03
	LSR	-3.63	LSR	-4.16

* DRR values in dB

Table 2: DRR values calculated from the dummy head measurements for frontal head orientation.

Conclusion and Future Work

We presented a database of spatial room impulse responses including spherical microphone array measurements on a $N = 44$ Lebedev grid, dummy head measurements on a circular grid, and omnidirectional measurements. The data can be used for research on high-resolution sound field analysis or Ambisonics en- and decoding. Furthermore, we provide basic CAD models of all 3 rooms including the source and receiver coordinates. The models can be used for research on room acoustical simulations. We plan to extend the models with proper materials and convert them to a data format readable by Unity or SketchUp. All updates will be published at [5].

Acknowledgements

The authors would like to thank Prof. Dr. Dieter Leckschat and Christian Epe from the Faculty of Media at the Hochschule Düsseldorf - University of Applied Sciences for kindly providing the Neumann KH420 loudspeaker.

References

[1] P. Stade, B. Bernschütz, and M. Rühl, “A Spatial Audio Impulse Response Compilation Captured at the WDR Broadcast Studios,” in *Proceedings of the 27th Tonmeistertagung - VDT International Convention*, 2012, pp. 551—-567.

[2] B. Rafaely, *Fundamentals of Spherical Array Processing*. Springer, 2015.

[3] W. Zhang, T. D. Abhayapala, R. A. Kennedy, and R. Duraiswami, “Insights into head-related transfer function: Spatial dimensionality and continuous representation,” *The Journal of the Acoustical Society of America*, vol. 127, no. 4, pp. 2347–2357, 2010.

[4] B. Bernschütz, C. Pörschmann, S. Spors, and S. Weinzierl, “Entwurf und Aufbau eines variablen sphärischen Mikrofonarrays für Forschungsanwendungen in Raumakustik und Virtual Audio,” in *Proceedings of 36th DAGA*, 2010, pp. 717–718.

[5] Lübeck, Tim and Arend, Johannes M. and Pörschmann, Christoph, “Zenodo repository,” https://zenodo.org/record/5031335.

[6] V. I. Lebedev, “Spherical quadrature formulas exact to orders 25-29,” *Siberian Mathematical Journal*, vol. 18, no. 1, pp. 99–107, 1977.

[7] S. Müller and P. Massarani, “Transfer-function measurement with sweeps,” *Journal of the Audio Engineering Society*, vol. 49, no. 6, pp. 443–471, 2001.

[8] P. Dietrich, M. Guski, J. Klein, M. Müller-Trapet, M. Pollow, R. Scharrer, and M. Vorländer, “Measurements and room acoustic analysis with the ita-toolbox for matlab,” in *40th Italian Annual Conference on Acoustics (AIA) and the 39th German Annual Conference on Acoustics (DAGA)*, 2013.