Investigations on Loudspeaker-based Auralization of Immersively Connected Rooms

Henning Piper, Marcel Nophut, Robert Hupke, Stephan Preihs, Jürgen Peissig
Leibniz Universität Hannover, Institut für Kommunikationstechnik, 30167 Hannover, Deutschland,
Email: nophut@ikt.uni-hannover.de

Abstract

The research project LIPS (Live Interactive PMSE Services[1]) aims at enabling tele-conferencing and distributed music performances in a most natural and immersive way. To provide a most plausible acoustic environment for the distributed interacting participants the Institut für Kommunikationstechnik (IKT) is working on Immersively Connected Rooms, where two remote rooms interact as they were coupled through a virtual window building a virtually combined larger room. In this contribution the geometric acoustic simulation tool RAVEN was used to compute the sound field in the virtually combined room. By using the spatial audio reproduction techniques Higher-Order Ambisonics (HOA) and Vector-base Amplitude Panning (VBAP) auralizations of these sound fields were generated for a unidirectional case, where the listening and source position were on different sides of the virtual window. These auralizations were then played back through a 36.2 loudspeaker setup. To evaluate how well the auralized room acoustic corresponds to the simulated acoustic environment and how the reproduction room influences the auralization, a set of room acoustic parameters was obtained from simulated and measured room impulse responses, which were then compared to each other.

Introduction

Immersively Connected Rooms means an audio-visual connection of two real rooms at geographically distant locations. A *virtual window* connects the two rooms, as depicted in Figure 1, which allows a visually plausible view into the other room. Furthermore, the room acoustics in each room change accordingly. Therefore, it sounds like being present in this virtually coupled room. This virtual room acoustic shall be auralized in each of the rooms through a multichannel loudspeaker setup.

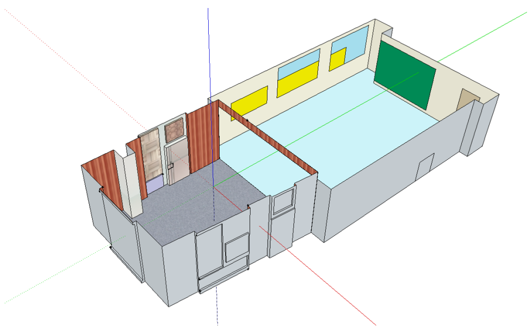

Figure 1: Model of rooms coupled through a virtual window. Here, coupling of Immersive Media Lab at IKT (left) and a seminar room from GRAP database (right).

[1] http://www.lips-project.de

To achieve this, the room acoustic properties of the two rooms A and B which are coupled through the virtual window have to be modeled. By picking up the sound sources with close-up microphones and tracking the position of all listeners and sources a corresponding sound field in the coupled room can be modeled by means of geometric acoustical simulations. This simulated sound field can then be rendered to an acoustic scene using the VBAP or Ambisonics technique for a playback through a multichannel loudspeaker setup. This acoustic scene can be supplemented by a corresponding video playback through VR goggles or a projection screen. The general signal flow is depicted in Figure 2.

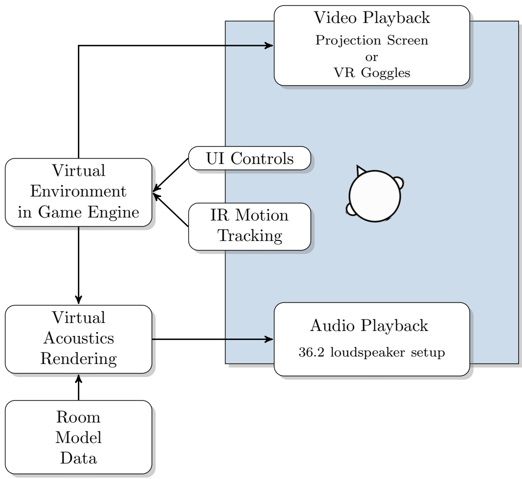

Figure 2: General signal flow and dependencies of processing instances in one of the rooms.

In future setups the room acoustical relevant parameters that are necessary to model the rooms might be obtained by automated measurement techniques. This is also a topic of current research [1].

Methods

In this work a uni-directional scenario with a listener in a real room A and a source in a virtual room B was considered using the following tool-chain:

1. **SketchUp**: Creation of coupled room models.
2. **RAVEN**: Simulation of sound fields for several source and receiver positions, computation of respective acoustic impulse responses and rendering of VBAP or HOA auralizations.
3. **Max/MSP**: Lookup of auralizations for given positions, real-time convolution/decoding and audio playback through a setup of 36 loudspeakers.

The models of the two Rooms A and B were created using the design tool SketchUp. To add the acoustic properties SketchUp was extended by the RAVEN Plug-in [2]. Subsequently, the sound fields for a source and receiver, both with omnidirectional characteristics, at several positions were computed using RAVEN. By means of the same tool these sound fields were rendered for the spatial reproduction techniques, Vector Base Amplitude Panning (VBAP) and Higher Order Ambisonics (HOA). The simulated sound fields were then auralized in the IKT's Immersive Media Lab [3] through a setup of 36 loudspeakers.

Experimental Evaluation

For the auralized *Immersively Connected Rooms* depicted in Figure 1 the room acoustic metrics Reverberation Time (T_{30}), Early Decay Time (EDT), Clarity (C_{80}) and Definition (D_{50}) were evaluated. The Immersive Media Lab at IKT, Room A, was modeled using data from previous measurements, whereas the seminar room, Room B, was taken from the room database GRAP [4]. Room impulse responses were measured by playing back auralized log-sweeps from room B to room A for several source and sink positions. The positioning of the source and receiver was carried out in accordance with the measurement setup suggestions described in the standard DIN 3382-2 [5]. The metrics were then derived from the obtained RIRs. This was evaluated for both HOA and VBAP auralization renderings. The measurements of the auralizations were then compared to the pure simulation considering the Just Noticable Difference (JND) of the metrics [6]. The upper and lower JND thresholds from the pure simulation results were then used as tolerance intervals. Results are depicted in Figures 3-6.

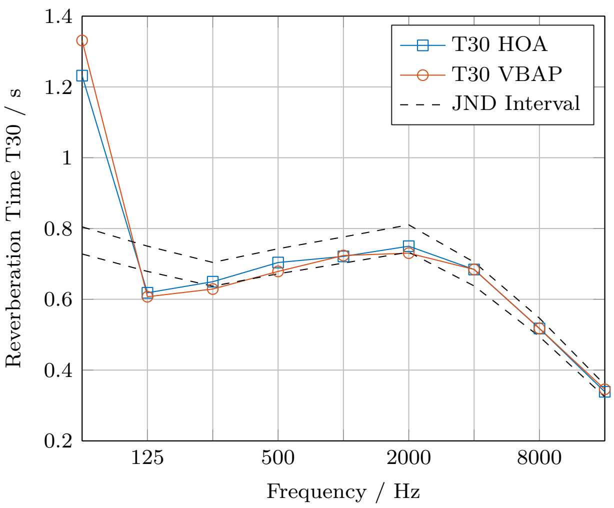

Figure 3: Reverberation Time (T_{30}) vs. Frequency.

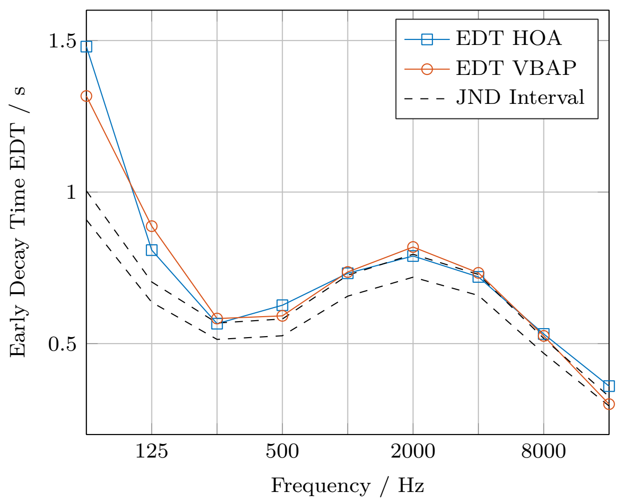

Figure 4: Early Decay Time (EDT) vs. Frequency.

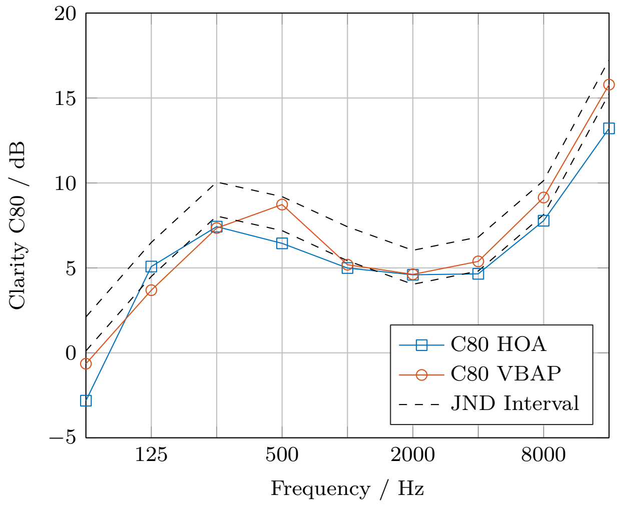

Figure 5: Clarity (C_{80}) vs. Frequency.

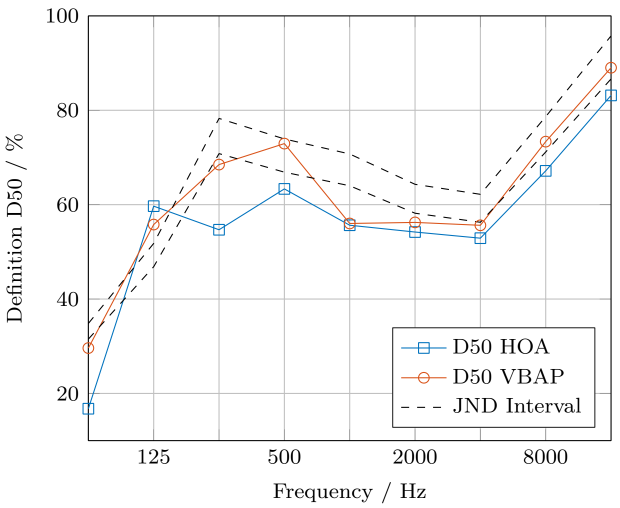

Figure 6: Definition (D_{50}) vs. Frequency.

For most octave bands the JND interval of the four metrics is met or is met with minor deviations, whereby the lowest frequency bands show larger deviations than the higher frequency bands. Only the Definition D_{50} shows some larger deviations in the mid frequency bands. Both reproduction techniques show a similar performance and none of them outperforms the other significantly.

Conclusion

The coarse characteristics of the considered room acoustical metrics could be reproduced in this study. However, there are deviations from the JND corridor of the simulation results. The main reason for this might be the influence of the reproduction room on the metrics (cf. [7]). Further research could include the compensation of the reproduction room's influences or the extension of the framework for dynamic real-time auralizations.

Acknowledgement

The project LIPS-CISS is co-funded by the Federal Ministry for Economic Affairs and Energy under the reference number 01MD18010G.

References

[1] Sebastian Lauster, Robert Hupke, Marcel Nophut, Stephan Preihs, and Jürgen Peissig, "Bestimmung des reflexionsfaktors mit einem sphärischen mikrofonarray," in *Fortschritte der Akustik: DAGA 2020*. 2020.

[2] Sönke Pelzer, Lukas Aspöck, Dirk Schröder, and Michael Vorlaender, "Integrating Real-Time Room Acoustics Simulation into a CAD Modeling Software to Enhance the Architectural Design Process," *Buildings*, vol. 4, pp. 113–138, 04 2014.

[3] Robert Hupke, Marcel Nophut, Song Li, Roman Schlieper, Stephan Preihs, and Jürgen Peissig, "The Immersive Media Laboratory: Installation of a Novel Multichannel Audio Laboratory for Immersive media applications," in *Audio Engineering Society Convention 144*, May 2018.

[4] David Ackermann und Maximilian Ilse, "The Simulation of Monaural and Binaural Transfer Functions for a Ground Truth for Room Acoustical Analysis and Perception (GRAP)," Masterarbeit, Technische Universität Berlin, 2015.

[5] DIN Deutsches Institut für Normung e.V, "Akustik – Messung von Parametern der Raumakustik: Teil 2: Nachhallzeit in gewöhnlichen räumen (DIN EN ISO 3382-2)," 2009.

[6] DIN Deutsches Institut für Normung e.V, "Akustik – Messung von Parametern der Raumakustik: Teil 1: Aufführungsräume (DIN EN ISO 3382-1)," 2009.

[7] Michael Kohnen and Michael Vorländer, "Virtual scene adaption for compensation of the reproduction room," in *Proceedings of INTERNOISE 2019*, Madrid, 2019.

[8] Dirk Schröder, *Physically based Real-Time Auralization of interactive virtual Environments*, Ph.D. thesis, Rheinisch-Westfälischen Technischen Hochschule Aachen, 2011.

[9] Michael Vorländer, *Auralization*, Springer-Verlag, Berlin, Heidelberg, 2008.