MODIFIED LATE REVERBERATION IN AN AUDIO AUGMENTED REALITY SCENARIO

Christian Schneiderwind *

Electronic Media Technology Group
Technische Universität Ilmenau
Ilmenau, Germany
christian.schneiderwind@tu-ilmenau.de

Annika Neidhardt †

Institute of Sound Recording
University of Surrey
Guildford, UK
a.neidhardt@surrey.ac.uk

ABSTRACT

This paper presents a headphone-based audio augmented reality demonstrator showcasing the effects of manipulated late reverberation in rendering virtual sound sources. The setup is based on a dataset of binaural room impulse responses measured along a 2 m long line, which is used to imitate the reproduction of a pair of loudspeakers. Therefore, listeners can explore the virtual sources by moving back and forth and rotating arbitrarily on this line. The demo allows the user to adjust the late reverberation tail of the auralizations interactively from shorter to longer decay times regarding the baseline decay behavior. Modification of the decay times is based on resynthesizing the late reverberation using frequency-dependent shaping of binaural white noise and modal reconstruction. The paper includes descriptions of the frameworks used for this demo and an overview of the required data and processing steps.

1. INTRODUCTION

Successful implementations of audio augmented reality must consider room acoustics. For embedding virtual sound sources into the real environment, the virtual room acoustic behaviour has to match the room acoustic behaviour of the reproduction environment perceptually [1, 2]. This includes considering early reflections caused by physical structures in the environment and the statistical properties of the emerging diffuse late reverberant field. This can be achieved by measuring the transfer function, for example, using a dummy head, to get binaural room impulse responses or simulation techniques based on wave-based or geometrical acoustic methods. Previous research has already established that a mismatch between the real and virtual room acoustic information can lead to severe degradation of the quality of the auralization, from timbre and localization errors to even the absence of externalization.

However, experiments as in [3, 4, 5] indicate that there is potential to actively deviate from the actual room acoustic properties and still provide scenarios where virtual sources are successfully integrated into the environment. These results have raised the question of to what degree the virtual room acoustic information can diverge from the real room acoustics while still providing perceptually plausible auralizations. The presented demo is based on the experiment conducted in [5] to find how much the decay length of the late reverberation tail can be shortened or extended before listeners notice a degradation in the quality of the illusion. Manipulation of the late reverberation was implemented by extracting and resampling frequency-dependent decay curves of the late reverberation and applying them to binaural white noise to create a synthetic binaural late reverberation tail. Reported effects of a degradation in audio-visual plausibility included slightly distorted distance perception when the decay length was shortened by over 60 %. On the other hand, extending the diffuse decay time by 160 % and higher was reported to make it challenging to locate the sound source position precisely. However, a complete breakdown of the illusion was only noticed in cases where late and early reflections were removed entirely. Although only one room was investigated, the results suggest that modifying the late reverberation does not necessarily impact the audio-visual plausibility of the presented scene, depending on the amount of deviation.

The presented demo offers a position-dynamic binaural audio augmented reality (AAR) experience similar to the one described above. This demo focuses on the listener's freedom to actively adjust the reverberation length in real-time and explore the perceptual impact of a virtual sound source reproduction when the characteristics of the real and virtual acoustic environment deviate.

2. AUDIO AUGMENTED REALITY DEMO

The presented demonstrator creates an AAR scenario where a listener can experience the binaural simulation of two loudspeakers playing. The listener can switch between the real loudspeakers reproducing the sound and the simulation of the loudspeakers using a headphoned-based binaural reproduction. In this demo, a 2 m long path is indicated by a line on the floor, which the listeners can move on freely and explore the scene. This demo focuses on the match and deviations of the virtual acoustics from the real acoustics of the reproduction room. The binaural simulation of the loudspeakers allows the listener to adjust the room reverberation. To be precise, the late reverberation characteristics of the virtually reproduced room acoustics can be adjusted to yield a shorter or longer decay of the diffuse field.

Fig. 2 shows an example of the demo setup used for an experiment in an empty seminar room at Technische Universität Ilmenau. The line in the middle indicates the path that listeners can move along. The loudspeakers are placed at a distance of 1.25 m to this line, with one placed in front of the line and the other at the side. Both loudspeakers are directed towards the 2 m-line.

* This work was funded by the Carl-Zeiss-Stiftung under the project: “Co-Presence of Humans and Interactive Companions for Seniors (CO-HUMANICS)”.

† This work was funded by the Engineering and Physical Sciences Research Council (EPSRC) under the ‘SCalable Room Acoustics Modelling’ grant EP/V002554/1.

Figure 1: A participant exploring the real and virtual loudspeaker arrangement in the seminar room, illustrated in Fig.2. The participant wears a pair of AKG K1000 with the attached Generic Vive Tracker. The red line marks the explorable path. In this specific scenario, an arrangement where the frontal loudspeaker was turned around by 180 ° was chosen.

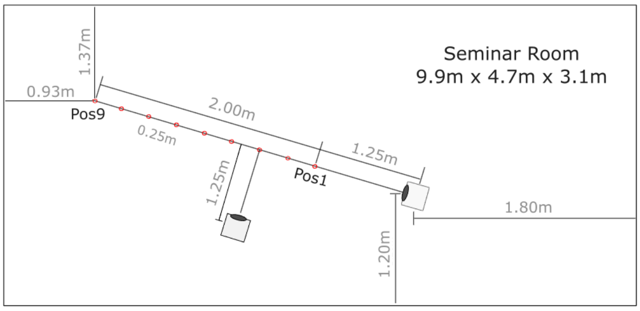

Figure 2: *An illustration of the measurement and auralization arrangement used for an experiment in a seminar room at Technische Universität Ilmenau. The red dots on the 2 m-line indicate the measurement positions of the dummy head. The squares show the positions of the loudspeakers facing the line.*

2.1. BRIR Data Set: Measurement Setup

The auralization of the virtual loudspeakers in this demo is based on a dataset of measured binaural room impulse responses. For that purpose, a dummy head mounted on an electric turntable was placed on the 2 m line at nine different positions, sampling the line with a resolution of 0.25 m. Using the electric turntable, head orientations in the azimuthal plane were captured with a 4 ° resolution, resulting in 90 angles. BRIRs were captured for only one fixed height. Moreover, the measurement setup did not allow for measuring BRIRs of orientation angles other than yaw. Although this limits the degrees of freedom the listeners can move in, it has little impact on this specific scenario, as previous research has shown. However, other measurement strategies might also be suitable, e.g., using a microphone array from which BRIRs can be rendered using the Spatial Decomposition Method (SDM, [6]) and HRTFs.

Another consideration when using measured BRIRs in such an AAR scenario is incorporating the headphones' influence on the real sound field. The headphones on the listener's head introduce distortions to the sound field created by the real sound sources in the environment. Even extra-aural models such as the AKG K1000 lead to spectral deviations in frequencies above 2 kHz [7, 8], depending on the incident angle of sound and the adjusted angle of the headphones' speakers. Therefore, the headphones used in the demo were placed on the dummy head for the duration of the measurements to account for the distortions. The dataset collected in the seminar room (Fig.2) can be downloaded at [9]:

zenodo.org/record/3457782

An in-depth documentation of the measurement, as well as a physical analysis of the data, can be found in [10]. A similar set of measurements will be conducted in the DAFx demonstration room. The documentation of the measurement setup, as well as a physical analysis of the captured impulse responses will be presented as part of the demonstration.

2.2. Auralization Framework

The core of the demo setup is based on the Python-based framework *pyBinSim*. This framework was introduced by Neidhardt et al. [11] for position-dynamic real-time auralization of BRIR datasets. It uses a uniform partitioned convolution scheme to process arbitrary audio signals with measured or simulated BRIRs. To allow for (position-)dynamic binaural reproduction in real-time, the partitioned BRIR filters are switched in the processing by receiving external filter update commands depending on the changes in positions and orientation of the listener. In this specific setup, the HTC Vive Tracking system tracks the listener's movement, including an HTC Vive Generic Tracker placed on top of the headphones and two Lighthouse base stations.

For this specific demo, the pyBinSim framework was modified to allow for the playback of the real loudspeakers. The listeners can bypass the convolution, and the signal is fed directly to the real loudspeakers.

Furthermore, individual sections of the partitioned filters can be switched independently to manipulate the late reverberation segment and influence the decay time.

The demonstration will provide the opportunity to explore the reproduction with different types of signals. In addition, real and virtual sound sources can be compared with equivalent as well as different signals.

2.3. Modified Late Reverberation

With this demonstrator, listeners can explore the perception of diverging room acoustic conditions in AAR. In particular, the user can experience different scaled versions of the decay length of the original late reverberation in the dynamic binaural reproduction. A parametric approach to (re-)synthesize and modify the late reverberation tail of BRIRs was used to render reverberation tails with different decay times. This approach includes binaural white noise shaping for the high frequencies in the reverberation tail, inspired by [12], and a modal estimation and synthesis technique to reconstruct and modify the low-frequency response of the tail. Here is a short summary of the modification process of the late reverberation tail:

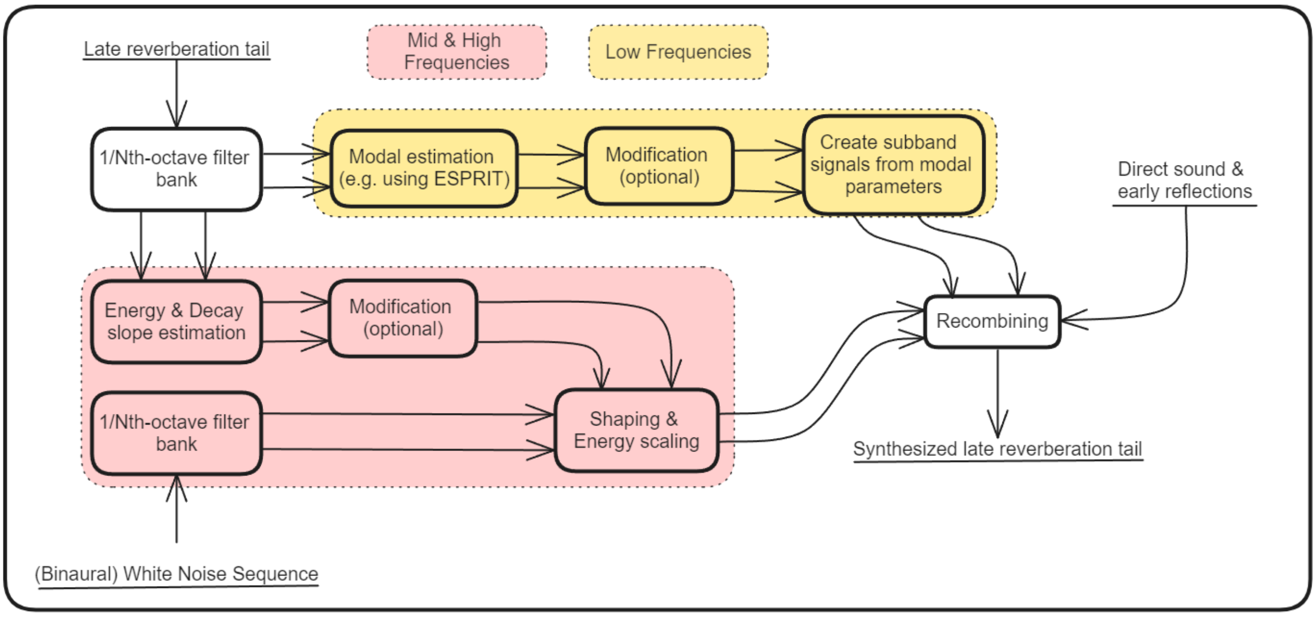

Figure 3: A block diagram of the late reverberation modification process [5]. The yellow and red marked blocks depict the processing paths for the low and high-frequency bands, respectively.

1. The late reverberation tail is extracted using a mixing time estimate to determine the crossover from early reverberation to the diffuse tail [13].
2. The extracted tail is fed through a filter bank to yield subbands with $1/6$-octave resolution [14]. The modification of the decay time is done separately for subbands below (LB) and above the Schroeder frequency (HB).
3. Extract decay parameters:
   - LB: Modal parameter estimation (ESPRIT [15, 16])
   - HB: Decay slope estimation
4. Modify decay parameters:
   - LB: Modify modal parameters to yield a targeted decay time for the individual modes
   - HB: Resampling of decay slopes
5. Reconstruct reverberation tail:
   - LB: Recreating subband signals from modified modes using exponential decaying sinusoids
   - HB: Resampling of decay slopes and applying them to the subbands of a binaural white noise sequence.
   - Combine the reconstructed tails (LB and HB)
6. Concatenate the resulting tail to the early BRIR segment

More information about this modification process can be found in [17]. Additionally, Fig.3 shows a block diagram of the modification algorithm.

Dealing with the low-frequency bands (< Schroeder frequency [18]) separately as a sum of individual decaying sinusoids, the strong modal behavior, often found in small rooms, is considered and can be reconstructed.

To save memory, the late reverberation tail of one BRIR from the dataset was used for modification and applied to the early segments of all other BRIRs in the set. This decision was based on the assumption that the late reverberation tail does not change significantly between the different positions and orientations.

After extracting the decay parameters for the low- and high-frequency subbands, arbitrary modifications can be introduced to these parameters.

By the time of submission of the demo, the following conditions are included:

- Original BRIR
- BRIR with synthesized tail (no modification)
- BRIR with synthesized tail (modified LB only)
- BRIR with synthesized tail (modified HB only)
- BRIR with synthesized tail (modified LB & HB)

However, arbitrary modifications can be introduced due to the parametric approach of synthesizing the late reverberation tails. This includes, for example, transferring the decay behavior of a different room into the synthesis and, hence, creating a mixture of rooms. An additional feature will be to adjust the modal decay behavior depending on the nature of the input signal by emphasizing or dampening characteristic modes or frequency bands.

3. SUMMARY AND OUTLOOK

The presented demo creates an AAR experience that allows listeners to explore two virtual sound sources in a room while being able to modify the late reverberation with a controller. It allows the listener to test how far they can alter the room's true acoustic characteristics until the quality of the experience is noticeably affected. The adjustable modifications are limited to lengthening and shortening the decay of diffuse late reverberation.

Within the range of modifications that do not negatively impact the audio-visual plausibility, future experiments can explore if purposely deviating from the true acoustics can yield desirable effects, such as speech enhancement in communication scenarios or more artistic freedom for designers of AR experiences.

4. REFERENCES

[1] A. Neidhardt, C. Schneiderwind, and F. Klein, "Perceptual matching of room acoustics for auditory augmented reality in small rooms - literature review and theoretical framework," *Trend in Hearing*, Jan. 2022.

[2] S. Werner, F. Klein, T. Mayenfels, and K. Brandenburg, "A summary on acoustic room divergence and its effect on externalization of auditory events," in *2016 8th Int. Conf. on Quality of Multimedia Experience (QoMEX)*, 2016, pp. 1–6.

[3] A. Neidhardt and S. Kamandi, "Plausibility of an approaching motion towards a virtual sound source II: In a reverberant seminar room," in *152nd AES Convention, The Hague, The Netherlands*, 2022.

[4] A. Neidhardt and C. Schneiderwind, "The influence of the DRR on audiovisual coherence of a real loudspeaker playing virtually over headphones," in *47th Annual Conference on Acoustics (DAGA), Vienna, Austria.*, 2021.

[5] C. Schneiderwind, M. Richter, N. Merten, and A. Neidhardt., "Effects of modified late reverberation on audio-visual plausibility and externalization in AR," in *2023 Immersive and 3D Audio: from Architecture to Automotive (I3DA)*, 2023, pp. 1–9.

[6] S. Tervo, J. Pätynen, A. Kuusinen, and T. Lokki, "Spatial decomposition method for room impulse responses," *Journal of the Audio Engineering Society*, vol. 61, no. 1/2, pp. 16–27, 2013.

[7] C. Schneiderwind, A. Neidhardt, and D. Meyer, "Comparing the effect of different open headphone models on the perception of a real sound source," in *150th AES Convention, Online*, May 2021.

[8] D. Satongar, C. Pike, Y. W. Lam, and A. I. Tew, "The influence of headphones on the localization of external loudspeaker sources," *J. Audio Eng. Soc.*, vol. 63, no. 10, pp. 799–810, Oct. 2015.

[9] A. Neidhardt, A. M. Zerlik, and S. Kamandi, *BRIR data set for interactive listener translation in two rooms*, Zenodo, July 2020, More details about the measurement are provided in the following publication: Neidhardt, A., "Data set: BRIRs for position-dynamic binaural synthesis measured in two rooms", In: Proceedings of 5th International Conference on Spatial Audio (ICSA), Ilmenau, Germany, Sept. 2019, doi: 10.5281/zenodo.3457782.

[10] A. Neidhardt, "Data set: BRIRs for position-dynamic binaural synthesis measured in two rooms," in *5th Int. Conference on Spatial Audio, Ilmenau, Germany*, 2019.

[11] A. Neidhardt, F. Klein, N. Knoop, and T. Köllmer, "Flexible python tool for dynamic binaural synthesis applications," in *eBrief 24, 142nd AES Convention, Berlin, Germany*, 2017.

[12] P. Stade and J. M. Arend, "Perceptual evaluation of synthetic late binaural reverberation based on a parametric model," in *Audio Engineering Society Conference: 2016 AES International Conference on Headphone Technology*. Audio Engineering Society, 2016.

[13] A. Lindau, L. Kosanke, and S. Weinzierl, "Perceptual evaluation of model-and signal-based predictors of the mixing time in binaural room impulse responses," *Journal of the Audio Engineering Society*, vol. 60, no. 11, pp. 887–898, 2012.

[14] J. Antoni, "Orthogonal-like fractional-octave-band filters," *The Journal of the Acoustical Society of America*, vol. 127, no. 2, pp. 884–895, 2010.

[15] C. Kereliuk, W. Herman, R. Wedelich, and D. J. Gillespie, "Modal analysis of room impulse responses using subband esprit," in *Proceedings of the International Conference on Digital Audio Effects*, 2018.

[16] O. Das and J. S. Abel, "Modal estimation on a warped frequency axis for linear system modeling," *arXiv preprint arXiv:2202.11192*, 2022.

[17] C. Schneiderwind and L. Treybig, "Late reverberation synthesis in small rooms considering modal properties," in *50th Annual Conference on Acoustics (DAGA), Hannover, Germany.*, 2024.

[18] M. R. Schroeder, "The "Schroeder frequency" revisited," *The Journal of the Acoustical Society of America*, vol. 99, no. 5, pp. 3240–3241, 05 1996.