Objective and Perceptual Evaluation of a Virtual Sound Environment System.

Cubick, Jens; Dau, Torsten

Publication date: 2016

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Objective and Perceptual Evaluation of a Virtual Sound Environment System

Jens Cubick, Torsten Dau
Hearing Systems group, Department of Electrical Engineering, Technical University of Denmark

Introduction

Hearing aid (HA) users often have difficulties following a conversation in challenging listening situations involving multiple talkers, background noise, and reverberation. To improve their listening performance, the algorithms in modern HAs have become increasingly complex. Yet, most HA testing is still performed in unrealistically simple setups.

In this study [1], speech reception thresholds (SRTs) were measured in a real classroom and its virtual counterpart auralized in the ‘Spacelab’ (see Fig. 1) to investigate, how well results obtained in a complex Virtual Sound Environment (VSE) translate to the corresponding real situation.

Method

Auralization technique

- Simulation of the classroom (Room 019, V=180m³, T₂₀=0.5s) in ODEON, see Fig. 2.
- Calibration of the room model to the measured reverberation time T₂₀ and clarity C₅₀, ISO 3382, see Fig. 4.
- Auralization of the results with a spherical 29 loudspeaker array in the ‘Spacelab’ using the LoRA toolbox [2], see Fig. 2.
- Rendering method: Higher Order Ambisonics (HOA) or Nearest Loudspeaker (NLS), where each reflection is mapped to a single loudspeaker.

Listening experiments

- Danzate II test [3] with 8 normal hearing, native Danish speaking listeners.
- Receiver-in-the-ear HAs with power domes providing flat, linear gain of 15 dB.
- Omnidirectional (Omn) microphone and static beamformer (BF).

Results – Physical evaluation

Room acoustic parameters (ISO 3382)

Fig. 4: Average Reverberation time T₂₀ and Clarity C₅₀ measured for 30 source positions in the real Room 019 (square), the VSE with NLS (crosses) and HOA (circles) rendering.

Results – Listening experiments

Fig. 5: Intenational cross-correlation coefficients measured in the real Room 019 (squares), the VSE with NLS (crosses) and HOA (circles) rendering.

Fig. 6: Average Speech reception thresholds measured in Rooms 019, the VSE with NLS coding and HOA coding (upper panel) and benefit from the BF program compared to the omnidirectional microphone (lower panel). The error bars indicate ±1 standard deviation.

Discussion – Physical Evaluation

- LoRA processing transparent with respect to T₂₀ and C₅₀.
- IACC lower in VSE than in Room 019, i.e., more diffuse sound field in the VSE.
- Reduced dynamic range of the directivity pattern over azimuth angle in VSE compared to Room 019, indicating a more diffuse sound field in the VSE.
- Effectiveness of the beamformer reduced in the VSE due to the higher diffuseness.

Discussion – Listening experiments

- SRTs higher in the VSE than in Room 019, probably due to higher diffuseness compared to Room 019.
- However, the SRT benefit from BF over omnidirectional microphone is similar in the VSE and Room 019.
- Higher SRTs with HOA compared to NLS coding, consistent with [4].
- Standard deviations are smaller in the VSE than in Room 019, consistent with the subjective description of the listeners.
- With NLS, SRT differences between 2m and 5m are almost identical to Room 019. Generally: NLS results are closer to reality.
- NLS preserves the dependence of the SRT on the target source distance.

Conclusion

The tested VSE seems to capture many acoustical features of a real environment (an existing room) that might be crucial for speech intelligibility, even though the resulting sound field in the VSE seems to be slightly more diffuse. The SRTs measured in the VSE were higher than those in Room 019. However, the differential outcome measure, BF benefit, was the same in the simulated and real environment.

VSEs could provide a powerful tool for testing HA algorithms, since very different acoustic conditions can be tested flexibly in a controlled and repeatable way. Array microphone recordings and screen projection could provide additional flexibility, for example, in the case of dynamic acoustic scenes.

References