Evaluating the auralization of a small room in a virtual sound environment using objective room acoustic measures

Ahrens, Axel; Marschall, Marton; Dau, Torsten

Publication date:
2016

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

Link back to DTU Orbit

Citation (APA):
Evaluating the auralization of a small room in a virtual sound environment using objective room acoustic measures

Axel Ahrens, Marton Marschall, Torsten Dau
Hearing Systems group, Technical University of Denmark, Lyngby, Denmark
aahr@elektro.dtu.dk

Introduction

To study human auditory perception in realistic environments, loudspeaker-based reproduction techniques have recently become state-of-the-art. To evaluate the accuracy of a simulation-based room auralization of a small room, objective measures were evaluated. In particular: early-decay time (EDT), ITD, JND, clarity, C50, C70, $\epsilon_{\text{auralization}}$, $\epsilon_{\text{reproduction}}$, interaural cross-correlation (IACC), speech transmission index (STI), direct-to-reverberant ratio (DRR). The method was implemented in a real-world playback room. The auralizations were realized using higher order ambisonics (HOA), mixed-order ambisonics (MOA), and a nearest loudspeaker method (NL) and reproduced in a virtual sound environment.

Method

Reference Room

Room Acoustic Model

Virtual Sound Environment

Reproduction techniques

- Nearest loudspeaker (NL), (Ferrante & Mushin, 2010)
- Higher-order ambisonics (HOA, 5th order)
- Mixed-order ambisonics (MOA, 7th/5th order; Denal et al. 2003)

Modeling

OEDEON v3.04 (Prehoelter, 1995) model of IEC listening room (7.5% 75% 2.8%)
- Material properties optimized using OEDEON’s genetic material optimizer (Christensen et al., 2014)

IR recording

- 7 source positions (Dynaudio BM5)
- 4 receiver positions (B&K 4192 and B&K HATS Type 4100)

- Processing and analysis using ITA-toolbox and TwoEar framework

Room Acoustic Measures

Binaural Measures

- Early-decay time (EDT)
- ITD (top) and JND (bottom)
- Clarity (C50, C70)
- Speech transmission index (STI)

Binaural Direct-to-Reverberant Ratio

- DRR (left ear: 360°, right ear: 0°)

Speech Intelligibility and STI

- SRT and SPEAK ratings
- Log-SNR and log-SRT

Conclusions

- Long-term, averaged measures are reproduced in the range of ~1 JND (T20/30, C50/80, STI, IACC)
- Short-term features of the impulse response are more difficult to capture leading to higher errors in e.g. EDT and STI
- Similar performances were obtained across reproduction techniques
- Auralization errors (auralization vs. model) are in the range of modelling errors (model vs reference)
- Dynamic binaural cues appear to be well captured
- Perceptual differences (e.g. speech intelligibility) occur, but not reflected in standard objective measures

Further investigations needed to link perceptual differences to objective measures

Literature

- Morales et al. (2016). Speech transmission indexes (STI) and speech reception thresholds (SRT) as target measures.
- JND
- Further investigations needed to link perceptual differences to objective measures
- The authors would like to thank the developers of the ITA-toolbox from RWTH Aachen as well as for the description of the ITA-toolbox’s parameters.
- The project was partly funded by the Oticon Centre of Excellence for Hearing and Speech Sciences (OCHSS)