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 the simulation-based room auralization of a small room, objective measures were evaluated in particular: early-decay time (EDT), 1st sound decay time (TS1, TS2), clarity (CT), C50, C80, interaural cross-correlation (IACC), speech transmission index (STI), direct-to-reverberant ratio (DRR), impulse responses (IR), and room impulse responses (RIR).

Room Acoustic Measures

Figure 1: Reverberation time (T20) measured in octave bands, measured at 7 source and 4 receiver positions. The blue and red curves indicate the ODEON model and the reference room, respectively. The remaining curves are auralized versions of the room.

Figure 2: Clarity (C50) measured in octave bands, measured at 7 source and 4 receiver positions. The blue and red curves include the ODEON model and the reference room, respectively. The remaining curves are auralized versions of the room.

Binaural Measures

Figure 3: Root mean square error (RMSE) of the clarity measures relative to the ODEON model. The dashed line indicates the perceptual just noticeable difference for clarity (0.5 dB, Astner et al. 2011).

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 CT
• Similar performances were obtained across leading to higher errors in e.g. EDT and CT

Acknowledgements

The authors would like to thank the dependants of the ITA loudspeakers from RWTH Aachen as well as the descriptions of the Fraunhofer’s Institute. Special thanks to the ODEON-Multimedia Research Centre in Exeter, for providing the speech material of the Centre 9.

This project was partly funded by the Oticon Centre of Excellence for Hearing and Speech Sciences (OCEHSS).

Literature


