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):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
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), room transfer function (RTF), clarity (CT), C50, C80, interaural cross-correlation (IACC), speech transmission index (STI), direct-to-reverberant ratio (DRR), impulse responses (IRs), were evaluated for auralized versions of the reference room. The auralizations were realized using higher-order ambisonics (HOA), mixed-order ambisonics (MOA), and a near-field loudspeaker method (NL) and reproduced in a virtual sound environment.

Room Acoustic Measures

Binaural Measures

Binaural-Direct-to-Reverberant Ratio

Speech Intelligibility and STI

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 reproduction techniques.
- Auralization errors (auralization vs. model) are in the range of modelling errors (model vs reference).

Further investigations needed to link perceptual differences to objective measures.

Acknowledgements

The authors would like to thank the developers of the ITA toolbox from RWTH Aachen as well as the descriptions of the ITA toolbox from RWTH Aachen. Additionally, thanks to Oticon Research Centre Eriksholm for providing the speech material of the Dantale II.

This project was partly funded by the Oticon Centre of Excellence for Hearing and Speech Sciences (DAN25).

References

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

Method

Reference Room

Room Acoustic Model

Virtual Sound Environment

Reproduction techniques

- Nearest loudspeaker (NL; Forschungsthemen, 2010)
- Higher-order ambisonics (HOA, 5th order)
- Mixed-order ambisonics (MOA, 7th/5th order; Daniel, 2003)

Modeling

- ODEON v3.04 (Phonelab.de; 1995) model of IEC listening room (7.5*5.75*2.8m)
- Material properties optimized using ODEON’s genetic material optimizer (Christensen et al., 2014)
- IR recording
- 7 source positions (Dynaudio BMM)
- 4 receiver positions (B&K 4192 and B&K HATS Type 4100)
- Processing and analysis using ITA-toolbox and Two!Ears framework

Binaural Measures

- Interaural correlation (IACC) 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 measured versions of the room.

Figure 1: Reversal time (TD) (measured standard deviation) over octave bands, measured at 7 source and 4 receiver positions. The blue and red curves include the ODEON model and the virtual room, respectively. The remaining curves are measured versions of the room.

Figure 2: Clarity (C80) measured-standard deviation over 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 measured versions of the room.

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 measurement noise and EDT (5%; Axelsen et al., 2014).

Figure 4: Root mean-square error (RMSE) of the IACC measures relative to the ODEON model. The early IACC is calculated over the first 80ms of the impulse response. The late IACC from 80ms onwards. The dashed line indicates the perceptual just noticeable difference for IACC (0.03; Álvarez et al., 2014).

Figure 5: Distribution of early-time ITDs (top) and ITDs (bottom) calculated for a source 30’ right of the NATS (new pictogram). Interaural differences were analyzed in 20ms windows with 10% overlap over a 150 ms period. Black and red curves indicate the early-time and late-time distributions calculated for the auralized and reference room.

Figure 6: Root mean-square error (RMSE) of the IACC measures relative to the reference room. The early IACC is calculated over the first 80ms of the impulse response. The late IACC is calculated over the reference room, respectively. The remaining curves are measured versions of the room.

Figure 7: Distribution of RT60 (top) and EDT (bottom). calculated for a source 30’ right of the NATS (new pictogram). Interaural differences were analyzed in 20ms windows with 10% overlap over a 150 ms period. Black and red curves indicate the early-time and late-time distributions calculated for the auralized and reference room.

Figure 8: 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 (5%; Axelsen et al., 2014).

Figure 9: Distribution of RT60 (top) and EDT (bottom) calculated for a source 30’ right of the NATS (new pictogram). Interaural differences were analyzed in 20ms windows with 10% overlap over a 150 ms period. Black and red curves indicate the early-time and late-time distributions calculated for the auralized and reference room.

Figure 10: Speech transmission index (STI) for a target speaker (STI) and reference room. RTDs and EDTs from distinctions shown in Fig. 7 as a function of loudspeaker position and auralization technique. The same frequency range was applied. The pictogram depicts the source/receiver setup.

Figure 11: Speech transmission index (STI) for a target speaker location (STI) and AO-based room auralization system, Acta Acustica united with Acustica, 96 (2).

Further investigations needed to link perceptual differences to objective measures.

Literature

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

The authors would like to thank the developers of the ITA toolbox from RWTH Aachen as well as the descriptions of the ITA toolbox from RWTH Aachen. Additionally, thanks to Oticon Research Centre Eriksholm for providing the speech material of the Dantale II.

The project was partly funded by the Oticon Centre of Excellence for Hearing and Speech Sciences (DAN25).