Exploiting Deep Neural Networks and Head Movements for Robust Binaural Localization of Multiple Sources in Reverberant Environments - DTU Orbit (10/11/2019)

This paper presents a novel machine-hearing system that exploits deep neural networks (DNNs) and head movements for robust binaural localization of multiple sources in reverberant environments. DNNs are used to learn the relationship between the source azimuth and binaural cues, consisting of the complete cross-correlation function (CCF) and interaural level differences (ILDs). In contrast to many previous binaural hearing systems, the proposed approach is not restricted to localization of sound sources in the frontal hemifield. Due to the similarity of binaural cues in the frontal and rear hemifields, front-back confusions often occur. To address this, a head movement strategy is incorporated in the localization model to help reduce the front-back errors. The proposed DNN system is compared to a Gaussian-mixture-model-based system that employs interaural time differences (ITDs) and ILDs as localization features. Our experiments show that the DNN is able to exploit information in the CCF that is not available in the ITD cue, which together with head movements substantially improves localization accuracies under challenging acoustic scenarios, in which multiple talkers and room reverberation are present.

General information
Publication status: Published
Organisations: Department of Electrical Engineering, Hearing Systems, University of Sheffield
Contributors: Ma, N., May, T., Brown, G. J.
Number of pages: 10
Pages: 2444-2453
Publication date: 2017
Peer-reviewed: Yes

Publication information
Journal: IEEE/ACM Transactions on Audio, Speech, and Language Processing
Volume: 25
Issue number: 12
ISSN (Print): 2329-9290
Ratings:
BFI (2017): BFI-level 2
Scopus rating (2017): CiteScore 4.4 SJR 0.867 SNIP 4.093
Web of Science (2017): Impact factor 2.95
Web of Science (2017): Indexed yes
Original language: English
Keywords: Binaural sound source localisation, Deep neural networks, Head movements, Machine hearing, Multi-conditional training, Reverberation
Electronic versions:
taslp_ma_2750760_proof_color.pdf
DOIs:
10.1109/TASLP.2017.2750760
Source: PublicationPreSubmission
Source ID: 139430671
Research output: Contribution to journal › Journal article – Annual report year: 2017 › Research › peer-review