Alternating least squares-based joint estimation of RETFs and PSDs for multi-channel speech enhancement

Marvin TAMMEN[(1,2)], Ina KODRASI[(3)], Simon DOCLO[(1,2)]

[(1)]Signal Processing Group, Dept. of Medical Physics and Acoustics, University of Oldenburg, Germany
[(2)]Cluster of Excellence Hearing4all, Germany
[(3)]Idiap Research Institute, Switzerland

Abstract

The multi-channel Wiener Filter (MWF) is a well-known speech enhancement technique that can be used to improve speech quality and intelligibility of microphone signals recorded in noisy and reverberant environments [1, 2]. It is commonly assumed that (i) late reverberation and ambient noise can be modeled as a spatially diffuse sound field and (ii) the spatial coherence of the remaining noise is known a-priori. Using these assumptions, the MWF requires estimates of the relative early transfer functions (RETFs) of the target speaker and the power spectral densities (PSDs) of the target speaker, the diffuse and the remaining noise. Recently, we proposed a technique to jointly estimate these quantities by minimizing a model-based error matrix via an alternating least squares (ALS) method [3, 4]. In this contribution, we present extensive simulation results comparing the ALS method with a state-of-the-art reference method based on covariance whitening. We demonstrate the effectiveness of the ALS method in both stationary and dynamic acoustic scenarios by using the estimates in an MWF and evaluating its noise reduction and dereverberation performance with respect to the improvement in speech quality. Results show that the ALS method yields more accurate estimates than the reference method, especially in the presence of strong uncorrelated noise.
Keywords: Multichannel Wiener Filter, Joint Parameter Estimation, Speech Enhancement

REFERENCES

[1] S. Doclo, S. Gannot, M. Moonen, and A. Spriet. Acoustic beamforming for hearing aid applications, in Handbook on Array Processing and Sensor Networks, pp. 269–302. Wiley, 2010.

[2] S. Doclo, W. Kellermann, S. Makino, and S.E. Nordholm, Multichannel Signal Enhancement Algorithms for Assisted Listening Devices: Exploiting spatial diversity using multiple microphones, IEEE Signal Processing Magazine, vol. 32, no. 2, pp. 18–30, Mar. 2015.

[3] M. Tammen, I. Kodrasi, and S. Doclo, Iterative Alternating Least-Squares Approach to Jointly Estimate the RETFs and the Diffuse PSD, in Proc. ITG Conference on Speech Communication, Oldenburg, Germany, Oct. 2018, pp. 221–225.

[4] M. Tammen, S. Doclo, and I. Kodrasi, Joint Estimation of RETF Vector and Power Spectral Densities for Speech Enhancement Based on Alternating Least Squares, in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 2019, pp. 795–799.