Researchers propose novel blind source separation framework for sound mixing

March 22 2022, by Li Yuan

Credit: Pixabay/CC0 Public Domain

Blind source separation (BSS) aims to estimate source signals from observed mixtures without prior information about the source or mixing system.

In the case of long reverberation times, the full-rank spatial covariance matrix (SCM) has been introduced, which shows improved separation performances. However, the full-rank SCM is still short of physical meaning.

Recently, researchers from the Institute of Acoustics of the Chinese Academy of Sciences (IACAS) proposed a BSS framework based on the frequency-domain convolution transfer function, which provides a new idea for solving the BSS problem in highly reverberant environments.

The study was published online in *IEEE/ACM Transactions on Audio, Speech, and Language Processing* on Jan. 25.

Without employing the narrowband assumption, they approximated the time-domain convolutive mixture using a frequency-wise convolutive mixture, and proposed a convolution transfer function (CTF)-based multichannel nonnegative matrix factorization (MNMF) framework for BSS in highly reverberant environments.

The full-rank SCM can be derived based on the proposed CTF framework and slowly time-variant source variances, which clearly explains why the full-rank spatial model works well in practice.

Based on the CTF framework, the researchers proposed a CTF-based MNMF algorithm for overdetermined BSS. Experiments showed that the proposed algorithm achieved a higher separation performance in reverberant environments.

  More information: Taihui Wang et al, Convolutive Transfer Function-Based Multichannel Nonnegative Matrix Factorization for Overdetermined Blind Source Separation, *IEEE/ACM Transactions on Audio, Speech, and Language Processing* (2022). DOI: 10.1109/TASLP.2022.3145304