Abstract

In applications such as speech and audio denoising, music transcription, music and audio based forensics, it is desirable to decompose a single-channel recording into its respective sources, commonly referred to as blind source separation (BSS). One of the techniques used in BSS is non-negative matrix factorization (NMF). In NMF both supervised and unsupervised mode of operations is used. Among them supervised mode outperforms well due to the use of pre-learned basis vectors corresponding to each underlying sources. In this paper NMF algorithms such as Lee Seung algorithms (Regularized Expectation Minimization Maximum Likelihood Algorithm (EMML) and Regularized Image Space Reconstruction Algorithm (ISRA)), Bregman Divergence algorithm (Itakura Saito NMF algorithm (IS-NMF)) and an extension to NMF, by incorporating sparsity, Sparse Non-Negative Matrix Factorization (SNMF) algorithm are used to evaluate the performance of BSS in which supervised mode is used. Here signal to distortion ratio (SDR), signal to interference ratio (SIR) and signal to artifact ratio (SAR) are measured for different speech and/or music mixtures and performance is evaluated for each combination.


http://www.telecom.tuc.gr/nikos/BSS. Nikos.html (for obtaining speech/music database)


Index Terms

Computer Science

Applied Mathematics

Keywords

NMF EMML ISRA IS-NMF SNMF Bregman Divergence