

Research on speech coding method using non-negative matrix factorization technique

Pham Manh Cuong

Abstract: This study presents a novel speech coding method based on non-negative matrix factorization (NMF). Unlike conventional codecs that rely on linear predictive coding or transform coding, the proposed approach leverages NMF to decompose the speech spectrogram into a dictionary of spectral basis vectors and a sparse matrix of activation weights. This decomposition efficiently captures the non-negative and additive nature of speech spectral components, enabling a compact representation. The coding process involves encoding the activation parameters and a pre-trained or adaptively updated dictionary. Experimental evaluations on clean and noisy speech datasets demonstrate that the proposed NMF-based coder achieves competitive perceptual quality and intelligibility at low bit rates compared to standard parametric codecs. Furthermore, the method exhibits inherent robustness to background noise due to its parts-based representation. The findings suggest that NMF offers a promising alternative for efficient and robust speech coding, particularly in applications requiring low computational complexity and noise resilience.

Keywords: speech coding, non-negative matrix factorization, audio signal processing, audio compression