

# Research on speaker information transformation in speech using time-domain decomposition techniques

Than The Huyen

**Abstract:** This study proposes a novel method for speaker information transformation in speech using time-domain decomposition techniques. Unlike conventional approaches that rely on spectral or parametric representations, the proposed method operates directly on raw waveforms by decomposing speech signals into temporal components that separately encode speaker identity and linguistic content. A neural network architecture is designed to learn and manipulate these decomposed representations, enabling the conversion of a source speaker's voice to a target speaker's voice while preserving the original linguistic information. Experimental evaluations on standard Vietnamese speech datasets demonstrate that the method achieves high naturalness and speaker similarity in transformed speech, outperforming baseline systems in objective metrics such as Mel-cepstral distortion and speaker verification accuracy. The results indicate that time-domain decomposition offers an effective and computationally efficient alternative for voice conversion, with potential applications in personalized speech synthesis, privacy protection, and assistive communication technologies.

**Keywords:** voice conversion, speaker information, time decomposition, speech processing