Recent Advances of Exemplar-Based Sparse Representation for Voice Conversion

Haizhou Li[†1]

Abstract: The exemplar-based sparse representation marks a success by performing voice conversion with a very limited amount of training data. In this framework, the activation matrix is shared between the source and target speakers. In a sense, such activation matrix serves as the bridge between the speakers. Ideally the activation matrix should capture the speaker invariant phonetic/prosodic information.

Recently, phonetic and prosodic information has been introduced in different ways to the exemplar-based sparse representation framework to improve the voice conversion quality. In this talk, we will discuss the idea of speaker dependent and speaker independent features in voice conversion. We aim to carry over the speaker independent features from source to target, and transform the speaker dependent features of source to those of target. We will report the latest advances in phonetic sparse representation, and the use of Phonetic PosteriorGrams features and prosodic features, as the case studies on how to improve the activation matrix.

---

†1 National University of Singapore