Supervised Speech Enhancement

Nasser Mohammadiha, Simon Doclo
Dept. of Medical Physics and Acoustics and Cluster of Excellence Hearing4all
University of Oldenburg, Germany

Introduction

In many applications, such as hearing aids and cochlear implants, reducing interference from noisy microphone recordings is very beneficial to improve the speech intelligibility and quality. When information about the acoustic environment is available, supervised speech enhancement methods can be used to obtain a higher-quality signal compared to the unsupervised speech enhancement methods. In this contribution, we present an overview of recently developed supervised methods for speech enhancement.

Methods

In supervised speech denoising methods, e.g., based on hidden Markov models (HMM) and nonnegative matrix factorization (NMF), a model is considered for both the speech and the noise signals and the model parameters are estimated using training samples from those signals. The speech and the noise models are then combined to construct a model for the noisy observations, using which the noise reduction task is carried out. An advantage of these supervised methods is that there is no need to estimate the noise power spectral density using a separate algorithm.

Results

Under matched training and testing conditions, supervised denoising methods can significantly outperform unsupervised methods. For example, a recently developed NMF-based noise reduction method outperformed a state-of-the-art unsupervised method by 0.2 MOS in perceptual evaluation of speech quality (PESQ), averaged over different noise types. This performance gap becomes smaller if there is a mismatch between training and testing data, where online learning methods can be used to improve the performance.

Conclusion

Supervised noise reduction methods can be used to reduce highly non-stationary noise, such as babble noise, given that a noise-dependent model can be learned a-priori. If the speech signal is degraded with a different noise type than the training noise type, the enhancement performance might degrade substantially. Online model learning paradigms may then provide a solution to reduce the effect of the mismatch.