# Speech enhancement based on Gaussian mixture modeling in the sub-band log-power domain

Tran Huy Dat, Kazuya Takeda, and Fumitada Itakura,

## Abstract

We present a speech enhancement system based on Gaussian mixture modeling in the sub-band log-power domain of the observed noisy speech. The basic idea of this method is fitting the actual behaviors of noise and noisy speech powers in terms of their distributions in each sub-band and employing the statistical methods for the noise power estimation and speech activity discrimination. The conventional two components GMM with standard EM algorithm is applied in each sub-band for each segment of half second of the observed noisy speech power. Two statistical methods of maximum a posterior probability (MAP) and cumulative distribution function equalization (CDFE) are developed in this works for the noise estimation. For the voice activity detection, an adaptable decision rule is proposed for the speech recognition application. The noise power and VAD are used in a Wiener filtering system. In an experimental evaluation on AURORA2 database, we compare the proposed to the conventional VAD and noise estimation method. From the experimental results, the proposed VAD method is superior in the non-speech detection rate and the Wiener filtering system based on proposed noise estimation performed better in speech recognition rate, especially in the case when CDFE estimator is employed.