

This poster presents a new method for reverberant speech separation, based on the combination of binaural cues and blind source separation (BSS). The main idea is to model interaural phase difference, interaural level difference and frequency bin-wise mixing vectors by Gaussian mixture models for each source and then evaluate that model at each T-F point and assign the units with high probability to that source. The model parameters and the assigned regions are refined iteratively using the Expectation-Maximization (EM) algorithm. We also address the permutation problem of the frequency domain BSS by initializing the mixing vectors appropriately. The EM algorithm starts with binaural cues and after a few iterations the estimated mask is used to initialize the mixing vector model parameters. We performed experiments on speech mixtures, and showed an average of about 0.8 dB improvement in signal-to-distortion (SDR) over the binaural-only baseline.