ACOUSTICS2008/1573
Monaural segregation of reverberant speech

Zhaozhang Jin and Deliang Wang

Ohio State University, Department of Computer Science & Engineering, 2015 Neil Ave., Columbus, OH 43210, USA

A major source of signal degradation in realistic environments is room reverberation. Monaural speech segregation in reverberant environments is a particularly challenging problem. Although inverse filtering has been proposed to partially restore the harmonicity of reverberant speech before segregation, this approach is sensitive to different room configurations. In this study, we investigate monaural segregation of reverberant speech by employing a supervised learning approach to map from a set of pitch-based auditory features to a grouping cue, which encodes the posterior probability of a time-frequency unit being target dominant given observed features. We devise a novel objective function for the learning process, which relates to the goal of maximizing SNR directly. The models trained using this new objective function yield significantly better results for unit labeling. In our segregation system, a segmentation and grouping framework is utilized in order to capture segments reliably under reverberant conditions and organize them into streams. Systematic evaluations show that our approach produces very promising results.