



Robust Speech Recognition Algorithm from Acoustic Modelling using Wavlets and EEG Data

KoyaJeevan Reddy¹, Dr S P Venu Madhava Rao²

¹*Dept of ECE, SNIST*

²*Principal, St Martins Engg College, Hyderabad*

ABSTRACT

Automatic speech recognition systems commonly separate speech enhancement, including localization, beamforming and post filtering, from acoustic modeling. In this paper, we perform multichannel enhancement jointly with acoustic modeling in a deep neural network framework. Inspired by beamforming, which leverages differences in the fine time structure of the signal at different microphones to filter energy arriving from different directions, we explore modeling the raw time-domain waveform directly. We introduce a neural network architecture which performs multichannel filtering in the first layer of the network and show that this network learns to be robust to varying target speaker direction of arrival, performing as well as a model that is given oracle knowledge of the true target speaker direction. Next, we show how performance can be improved by factoring the first layer to separate the multichannel spatial filtering operation from a single channel filterbank which computes a frequency decomposition. We also introduce an adaptive variant, which updates the spatial filter coefficients at each time frame based on the previous inputs. Finally we demonstrate that these approaches can be implemented more efficiently in the frequency domain. Overall, we find that such multichannel neural networks give a relative word error rate improvement of more than 5% compared to a traditional beamforming-based multichannel ASR system and more than 10% compared to a single channel waveform model.