Speech Recognition in Noisy Environment: Transformation of Acoustic Parameters

Qin Zou, Om Deshmukh, Prof. Carol Y. Espy-Wilson

Background Information:

Currently, most of the speech recognition systems are designed to work in clean environment. If these systems are exposed to normal environment where noise is present, their performance degrades drastically. Therefore, in order to keep the accuracy of these recognition systems, a transformation must be found to reduce the differences between the training and testing environment.

Our Approach:

We have developed a set of Acoustic Parameters (APs) that capture the salient features of various broad classes of speech. However, these APs are not robust to noises. The main aim of this project is to develop algorithms that compensate the transformations undergone by the APs at different Signal-to-Noise ratios (SNRs) relative to their values in clean environment (i.e. at ∞ SNR). These transformed APs will be used in conjunction to the recognition system developed by our lab for robust speech recognition.

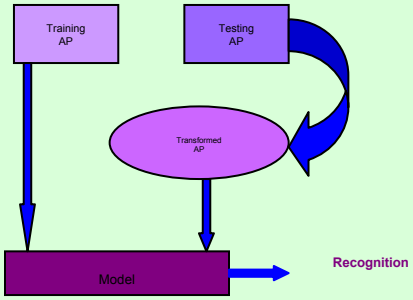

Figure1: illustration of our approach

Transformation Experiment:

In this experiment, the APs we utilized to test our transformation are *energy ratio*, *envelope variance*, *and spectral peak frequency*, only the first AP is discussed here. The Experiment evaluated how effectively the transformation can reduce the error between the original and distorted energy ratio values at different SNRs.

In each SNR case, we
(1) calculate the error by taking the norm of the difference between the distorted energy ratio values and those in clean environment
(2) obtain new values by applying a transformation to the distorted energy ratio values.
(3) compute error after the transformation using the same method as in part (1)

• Transformation at 20dB SNR:

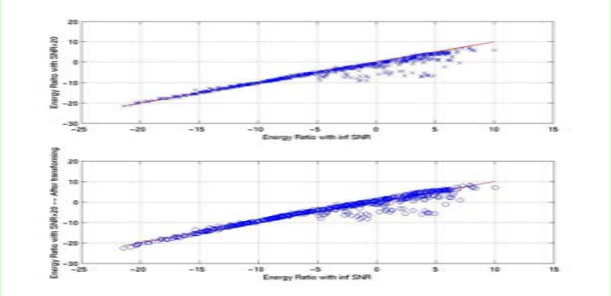

Figure2: (a) Energy Ratio before transform
(b) Energy Ratio after transform

Note: in figure 2, 3, 4, the red line represents the clean environment energy ratio values (i.e. ∞ SNR values)

Transformation at 10dB SNR:

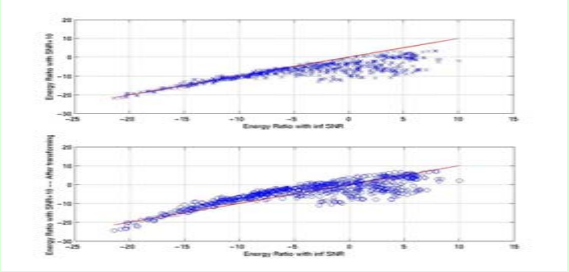

Figure3 (a) Energy Ratio before transform
(b) Energy Ratio after transform

Transformation at 0dB SNR:

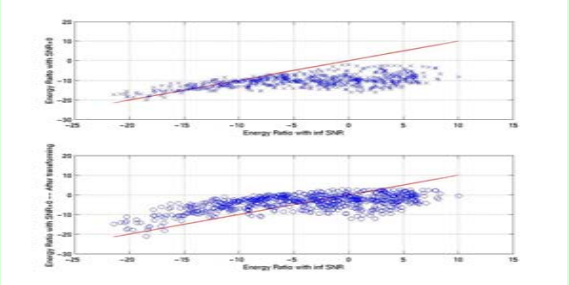

Figure4: (a) Energy Ratio before transform
(b) Energy Ratio after transform

Results:

SNR	Error Before Transform	Error After Transform
20dB	52.233413	44.805808
10dB	128.129526	91.122203
0dB	224.947649	143.892493

From the result table, we notice that the error gets reduced after transformation for each SNR case. Similar results are obtained from other tested APs. In future we will apply the transformed AP values to the recognition system developed by our lab. We believe that this can give improvement in the recognition accuracy in noisy environment.