



DEPARTMENT OF PHYSICS

Seminar series

You are cordially invited
to the following Seminar

Composite Recognition Scheme for Babble Noise Mitigation.

presented by

Mr. Jamin Atkins

Date: 7th December, 2015

Time: 1:00pm

Location: FST Rm. 413

Supervisor(s): Dr. Davinder Sharma



ABSTRACT

Speech recognition has become an established technology with many mainstream implementations. Although performance in laboratory environments had approached close to 99%, recognition accuracy decreases dramatically in the presence of babble noise. In this presentation, we shall discuss a method of improving the recognition of speech in the presence of babble noise. Exploratory analysis performed in initial experimentation has shown that there is a distinct difference between babble corrupted and clean speech. Both of the currently accepted models have shown distinct trajectories, which were different from clean speech. The distinction between the signals becomes less distinct at lower Signal-to-Noise Ratios (SNR). Using this information, we have developed a scheme to score frames based on the level of corruption. Traditional speech recognition was using the first 13 Mel Frequency Cepstral Coefficients (MFCCs) along with their delta and acceleration coefficients. MFCCs performed very well in the laboratory setting but their versatility was lowered in the presence of babble noise. In order to improve the versatility of the 39 dimensional MFCC feature vector, we calculated the spectral flux difference between two successive frames. The traditional 39 dimensional MFCC feature vector was augmented by adding spectral flux difference as a 40th coefficient. The spectral flux value was used as an indication of the level of noise corruption in the frame. Support Vector Machines, (SVMs) was then used for frame classification via the use of a Hidden Markov Model Intermediate Matching Kernel. For missing feature insertion of corrupted frames, N-gram modelling based rule was used. The resulting scheme showed 7-11% increase in recognition accuracy in the presence of both continuous and fluctuating SNR.