

Title: Robust Automatic Speech Recognition in the Presence of Babble Noise
Presenter: Mr. Jamin Atkins
Occasion: Upgrade Seminar
Date: Monday 11th June 2012
Time: 11:00 am
Venue: FSA, Rm. 413

Abstract:

Speech recognition is one of the next generation technologies for man-machine interaction. Recent applications of automatic speech recognition systems include mobile telephony, automobile feature controls, caption generation, transcription systems and interactive kiosk terminals. While automatic speech recognition systems have achieved word-accuracy-rates of nearly 100% in laboratory environments, their real world recognition accuracy rates are currently less than 40%. This difference in performance can be attributed to the unavoidable mismatch between training and implementation environments. Recently, a shift in the method for developing robust systems has seen individual environmental noise types targeted for rectification. Babble is one of the major noise types which are responsible for low word-accuracy-rates. Babble is the result of many overlapping conversations in the background and requires special attention since conventional methods for negation fail due to the fact that babble exists in the same frequency range as of clean speech. We present a proposal towards minimizing the effect of babble on the Mel –Frequency Cepstral Coefficients (MFCCs), the preferred features for automatic speech recognition. The proposed method of negation uses unsupervised learning to train Gaussian Mixture Models, with the goal of determining the means and covariance of both clean and babble polluted phonemes. The generated statistics would further not only contribute to the currently limited knowledge on the effect of babble on speech but also provide additional information, which can be used by feature space speech enhancement algorithms to reduce the effect of babble on speech.