

UCLA Electrical Engineering ARR 2011

Session 9. Advanced Algorithms for Speech and Signal Processing

South Bay Room

1:15-1:40

"Noise Robust Fundamental Frequency (F0) Estimation for Speech using SNR-weighted Summary Correlograms from Multi-band Comb Filters"

Lee Ngee Tan and Abeer Alwan

Abstract:

A noise-robust, signal-to-noise ratio (SNR)-weighted correlogram-based F0 estimation algorithm is proposed. A comb-filterbank operates in each frequency band (low, mid and high). Correlograms are obtained by applying autocorrelations on all comb-filtered channel outputs. An SNR-weighting scheme is used for channel selection to yield a summary correlogram from each correlogram. These summary correlograms are averaged and time-smoothed before peak extraction is performed. The proposed algorithm is evaluated on the Keele corpus with additive white or babble noises. In comparison with other widely-used F0 estimation algorithms, the proposed algorithm has the lowest overall gross pitch error, especially in low SNR cases.

1:40-2:05

"Real-time optogenetic functional magnetic resonance imaging (rt-ofMRI) using graphic processing unit (GPU) based parallel computation"

Zhongnan Fang, Jin Hyung Lee

(Abstract not available)

2:05-2:30

"Cooperative Prey Herding Based on Diffusion Adaptation"

Sheng-Yuan Tu and Ali H. Sayed

Abstract:

Mobile adaptive networks consist of a collection of nodes with learning and motion abilities that interact with each other locally in order to solve

distributed processing and distributed inference problems in real-time. In this talk, we develop adaptation algorithms that exhibit self-organization properties and apply them to the model of cooperative hunting among predators. The results help provide an explanation for the agile adjustment of network patterns in the interaction between fish schools and predators.

2:30-2:55

"Noise Robust Automatic Speech Recognition Using a Novel Approach to Log-Spectral Enhancement"

Julien van hout and Abeer Alwan

Abstract:

This talk presents a technique for enhancing the Mel-filtered log-spectrum of noisy speech, with application to noise robust speech recognition. We first compute a simple SNR-based soft-decision mask in the Mel-spectral domain as an indicator of speech presence. We exploit the known time-frequency correlation of speech by treating this mask as an image, and performing median filtering and blurring to remove the outliers and to smooth the decision regions. This mask constitutes a set of multiplicative coefficients (ranging in $[0,1]$) that are used to discard the unreliable parts of the Mel-filtered log-spectrum of noisy speech. As a last step, we apply Log-Spectral Flooring on the filtered spectra of both clean and noisy speech so as to match their respective dynamic ranges and to emphasize the information in the spectral peaks. The noisy MFCCs computed on these modified log-spectra show an increased similarity with their corresponding clean MFCCs. Evaluation on the Aurora-2 corpus shows that the proposed algorithm competes with state-of-the-art front-end procedures, like ETSI-AFE, with less computational complexity.