Abstract:

In real-life situations, sound signals often reach our ears as a mixture of target signal, background noise and/or competing speech. A typical example is talking with someone at a crowded group with other interfering noise, music and competing voices. While human can attend to individual sources without much difficulty even with a single ear, the performance of most speech processing systems is easily degraded in this adverse condition. Therefore, it is necessary to extract target sound source(s) from the mixture before any vulnerable processes. This seminar focuses on speech source separation problem in a single-microphone scenario based on how human auditory system organizes and separates a signal mixture.

A collection of psychoacoustic experiments has emerged since the last decade, showing that several essential perceptual cues are used to differentiate individual speech from each other. During the internship in MSRA, a separation system has been proposed by utilizing some of these perceptual cues, namely continuity and context information. An iterative trajectory regeneration algorithm is proposed to estimate the magnitude spectrum of each source. Dynamic spectral information is used as a constraint. Each estimated spectral trajectory, as a result, exhibits similar temporal variations as the original source. From experiments, it is shown that significant amount of energy from interfering source is removed with remarkable reduction in mean-square-error (MSE) of spectral estimate. Asymptotically, we also prove that the regenerated trajectory yields the same time variations as the given dynamic information.

About the speaker
Ms. Yvonne, Lee Siu Wa received her BEng and MPhil, both from the Chinese University of Hong Kong (CUHK) in 2002 and 2004 respectively, where she mainly worked in speech enhancement and robust speech recognition projects. She continues her research as a PhD candidate of the Department of Electronic Engineering there. Her current research interests are digital signal processing and speech processing.

Inquiries: Lo Wai Kit (wklo@se.cuhk.edu.hk), +852-26098338

*** ALL ARE WELCOME ***