Speech Processing

Eötvös Loránd University, Budapest, Hungary, 2005 – 2007 Fall, Spring
Instructor: Barnabás Póczos

This course provides a brief introduction to the principles of speech processing. We review the following basic elements: Information theory, ARMA systems, Hidden Markov models, Fourier analysis, energy, autocorrelation, linear predictive methods, source separation, and speech recognition. The course project assignments require programming in Matlab, C++, or Java.

Information Theory

Entropy. Mutual information. Rényi entropy and information. Channel capacity.
Nonparametric methods for entropy and mutual information estimation.
Kernel Mutual Information, Kernel Generalized Variance, Kernel Canonical Correlation Analysis.

Stochastic Process

AR, MA, ARMA systems.
Linear dynamical systems.
Expectation maximization. Maximum likelihood parameter estimation.
System identification with subspace methods.

Frequency domain

Fourier transformation.
Wavelets.
Discrete cosine transformation.
Z transformation.
Spectogram, specturm, cepstrum.
Sampling theory.
Denoising.

Linear filters

Infinite impulse response filters.
Finite impulse response filters.

Source Separation

Principal component analysis.
The cocktail party problem.
Independent Component Analyisis: FastICA, Jade, Radical algorithms.
ICA extensions: post-nonlinear ICA, ICA on AR processes.
Blind source deconvolution. Blind Subspace Deconvolution.
One-microphone source separation.

Speech Analysis

Linear predictive coding.
Pitch detection.

Speech Recognition

Hidden Markov models.
Dynamic time warping.