Syllabus

General information

Course	SPHSC 503
Course title	Current issues in Speech and Hearing Sciences: Speech Signal Processing
Credits	3 units
Schedule	Summer A term (June 19 – July 19), MWF 1:10-3:20 pm
Location	Savery Hall, room 149 (lectures), and room 137 (computer lab)
Instructor	Steven Schimmel
Instructor email	steven@ee.washington.edu
Course web site	http://staff.washington.edu/schmml/sphsc503/
Office Hours	MF 11:30am – 1:00pm, or by appointment

Prerequisites

This is an introductory course, so there are no formal requirements. Students are expected to have some familiarity with Matlab, but no programming experience is required.

Aims and Objectives

The aim of this course is to introduce students to the key concepts in digital signal processing and speech processing. The course will concentrate on a conceptual understanding of the techniques used in speech and signal processing, supported by some mathematical details. Students will develop practical Matlab skills to analyze and process speech signals using those techniques.

After taking this course, students should

- understand the fundamental concepts of speech and signal processing
- understand the main concepts of some key speech processing algorithms, such as pitch estimation, LTI filtering, time-frequency analysis, linear predictive coding, the cepstrum, and speech recognition
- be able to apply those algorithms to speech signals in Matlab

Course outline

- Fundamentals of Digital Signal Processing
  - Signals, sequences, systems
  - Linear, time-invariant systems
  - Impulse response
  - Convolution, correlation
  - Frequency response
  - Causality, stability
  - FIR systems, IIR systems
- Speech representations
  - Time-domain: waveform
  - Time-frequency domain: spectrogram, windows
  - Linear predictive coding
  - Cepstrum
- Time-domain speech analysis
  - End-point detection / voice activity detection
  - Voiced/unvoiced detection

  - Pitch estimation (autocorrelation)
- Frequency and time-frequency domain speech analysis
  - Formant tracking
- Speech processing
  - Sampling, up/down-sampling
  - Pre-emphasis, de-emphasis
  - Dynamic range compression (automatic gain control)
  - Filtering: FIR filters, IIR filters, filter design, filter analysis, filter implementation
- Speech enhancement
  - Noise reduction/removal: wiener filtering, spectral subtraction
- Speech production and speech perception
  - Human speech and hearing (brief, assumed familiar)
  - Engineering models for speech production (source / system)
  - Engineering understanding of speech perception
  - (filterbank model, envelope detection)
  - Higher levels of perception, auditory scene analysis
- Introduction to automatic speech recognition

Lectures, in-class exercises, homework, final project and grading

The course consists of lectures, in the lecture room, and hands-on sessions in the computer lab. The handouts for the hands-on sessions will contain in-class exercises. These exercises will not be graded, and collaboration on the in-class exercises is highly encouraged.

There will be three weekly homework assignments. Homework will be assigned on Wednesdays, and is due the next Wednesday. Homework assignments will be posted on the class website on Wednesday 6/21, 6/28, and 7/5, and are due 6/28, 7/5 and 7/12, respectively. You are free to discuss the homework assignments with others to understand the concepts, but you must write up your own solutions.

Ideally, the course ends with a final project, done individually or in pairs. This option will be discussed in class with the students, and its feasibility will be determined in the first week of class. Final projects would be due Wednesday 7/19, in the form of a write-up (~4 pages) and a presentation (~15 minutes) in class on Wednesday 7/19. Alternatively, the course could end with a larger homework, due Wednesday 7/19 and assigned no later than Wednesday 7/12.

The grade in this class will be based on the following;

- In-class participation (20%)
- Homework (40%)
- Final project (40%) or final homework (40%)

Class text

We will use notes and handouts made by the instructor for the lectures and the hands-on sessions. They will be published before the relevant lecture on the class website.

The notes, handouts and in-class exercises are based on the following texts:

- Discrete-Time Signal Processing, by Alan V. Oppenheim and Ronald W. Schafer, 2nd edition.
- Discrete-Time Speech Signal Processing: principles and practice, by Thomas F. Quatieri.

- Spoken language processing: a guide to theory, algorithm, and system development, by Xuedong Huang, Alex Acero and Hsiao-Wuen Hon
- Digital Signal Processing using Matlab, by Vinay K. Ingle and John G. Proakis
- Computer-based exercises for Signal Processing using Matlab, by James H. McClellan *et al*.