

Course Description:

Speech is the most natural means of communication among humans. It also plays a critical role in enhancing human-machine communication. In this course, we attempt to cover all aspects of digital speech processing, including both theoretical and practical topics. The linkage with acoustics and language processing will also be discussed.

Course Outline:

- Introduction to Speech Communications
- Mathematical Foundations
- Speech and Acoustics
- Speech Analysis and Speech Modeling
- Speech Production and Speech Synthesis
- Speech Coding
- Speech Recognition
- Speaker Recognition
- Language Identification
- Spoken Dialogue Processing
- Speech Data Mining
- Natural Language Processing

Teaching Philosophy: The textbook by Quatieri is the main source of knowledge. You are expected to read assignments before you come to each class. From time to time, we will go beyond the textbook in our class lectures to supplement your learning. Office hours are best time for you to ask questions. It is important for you to study the textbooks, attend the lectures and take advantage of the office hours. Digital speech processing requires also hand-on exposure to real speech data. Your homework assignment will include some programming exercises. Your term project is expected to go deeper into a subject of your choice.

Grading Policy: Students are graded based on homework (30%), examinations (40%), and a team project report at the end of the semester (30%).

- **Homework** (30%)
- **Examinations** (40%):
 - Quiz #1 (10%):
 - Quiz #2 (10%):
 - Final (20%):
- **Project** (30%):
 - A Team Project on topics discussed in the course is expected.
 - A team should have no more than two people on the project. Teams will be formed first and course projects will be commented before finalizing.
 - A list of projects will be posted in February and discussed in class.
 - MATLAB and other programming tools (C/C++) are essential for your projects.
 - A collection of speech samples will be posted for download.
 - Project term papers are due before the Finals Week.

Prerequisites: ECE4270 or ECE6250

Main Text: T. F. Quatieri, *Discrete-Time Digital Speech Processing*, Prentice Hall, 2002

Supplements:

- (1) Manning & Shutze, *Foundations of Statistical Natural Language Processing*, MIT Press, 1999.
- (2) Rabiner & Schafer, *Digital Processing of Speech Signals*, Prentice Hall, 1978.

Readings:

- (1) C. Cherry, *On Human Communications*, MIT Press, 1968.
- (2) D. G. Stork (ed.), *HAL's Legacy*, MIT Press, 1997.

Research Papers: Will be provided as needed.

Important Class News: Watch for update of important news on the course web page