Digital Speech Processing

Professor Lawrence Rabiner
UCSB
Dept. of Electrical and Computer Engineering
Jan-March 2011

Course Description

This course covers the basic principles of digital speech processing:

- Review of digital signal processing
- Fundamentals of speech production and perception
- Basic techniques for digital speech processing:
  - short-time energy, magnitude, autocorrelation
  - short-time Fourier analysis
  - homomorphic methods
  - linear predictive methods
- Speech estimation methods
  - speech/non-speech detection
  - voiced/unvoiced/non-speech segmentation/classification
  - pitch detection
  - formant estimation
- Applications of speech signal processing
  - Speech coding
  - Speech synthesis
  - Speech recognition/natural language processing

A MATLAB-based term project will be required for all students taking this course for credit.

Course Information

- **Grading**:
  - Homework 20%
  - Term Project 20%
  - Mid - Term Exam 20%
  - Final Exam 40%
- **Prerequisites**: Basic Digital Signal Processing, good knowledge of MATLAB
- **Time and Location**: Tuesday, Thursday, 10:30 am to 11:50 am, Phelps 1437.
- **Course Website**: [www.ece.ucsb.edu/Faculty/Rabiner/ece259](http://www.ece.ucsb.edu/Faculty/Rabiner/ece259)
- **Office Hours**: Tuesday, 1:00-3:00 pm
Course Readings

Required Course Textbook:

Recommended Supplementary Textbook:

Matlab Exercises:

References in Selected Areas of Speech Processing

Speech Coding:
• W. B. Kleijn and K. K. Paliwal, Editors, Speech Coding and Synthesis, Elsevier, 1995
• P. E. Papamichalis, Practical Approaches to Speech Coding, Prentice Hall Inc, 1987
• N. S. Jayant and P. Noll, Digital Coding of Waveforms, Prentice Hall Inc, 1984

Recommended References

• J. D. Markel and A. H. Gray, Jr., Linear Prediction of Speech, Springer-Verlag, Berlin, 1976
• B. Gold and N. Morgan, Speech and Audio Signal Processing, J. Wiley and Sons, 2000
• D. O'Shaughnessy, Speech Communication, Human and Machine, Addison-Wesley, 1987
• S. Furui and M. Sondhi, Advances in Speech Signal Processing, Marcel Dekker Inc, NY, 1991
• D. G. Childers, Speech Processing and Synthesis Toolboxes, John Wiley and Sons, 1999
• K. Stevens, Acoustic Phonetics, MIT Press, 1998

Speech Synthesis:
• Y. Sagisaka, N. Campbell, and N. Higuchi, Computing Prosody, Springer Verlag, 1996
**References in Selected Areas of Speech Processing**

**Speech Recognition:**

**References in Digital Signal Processing**


---

**The Speech Stack**

*Speech Applications* — coding, synthesis, recognition, understanding, verification, language translation, speed-up/slow-down

*Speech Algorithms* — speech-silence (background), voiced-unvoiced, pitch detection, formant estimation

*Speech Representations* — temporal, spectral, homomorphic, LPC

*Fundamentals* — acoustics, linguistics, pragmatics, speech production/perception

---

**Digital Speech Processing**

- Ability to implement theory and concepts in working code (MATLAB, C, C++)
- Basic understanding of how theory is applied
- Mathematics, derivations, signal processing

**Course Outline – ECE 259A – Speech Processing**

- Jan 4 - Lecture 1, Introduction to Digital Speech Processing
- Jan 6 - Lecture 2a, Review of DSP Fundamentals
- Jan 11 - Lecture 2b, Review of DSP Fundamentals
- Jan 13 - Lecture 3a, Acoustic Theory of Speech Production
- Jan 15 - Lecture 3b, Lecture 4, Speech Perception—Auditory Models, Sound Perception, MOS Methods
- Jan 20 - Lecture 5, Sound Propagation in the Vocal Tract—Fundamentals, Solutions of the Wave Equation
- Jan 22 - Lecture 6, Sound Propagation in the Vocal Tract—Lossless Tube Models, Digital Filters
- Feb 1 - Lecture 8, Time Domain Methods—Short-Time Energy, Magnitude, Zero Crossings, Autocorrelation
- Feb 3 - Lecture 9, STFT Methods—Introduction, FBS, OLA, Modifications
- Feb 8 - Lecture 10-11, STFT Methods—Speech Representations Using Analysis-Synthesis Methods
- Feb 12 - Lecture 12a, Homomorphic Speech Processing—Analysis, Synthesis Methods
- Feb 15 - Lecture 12b, Homomorphic Speech Processing—Practical Implementations
- Feb 22 - Lecture 13, Linear Predictive Coding (LPC)—Introduction, Autocorrelation Method, Covariance Method
- Feb 24 - Lecture 14, LPC—Lattice Implementation, Frequency Domain Interpretations
- Mar 1 - Lecture 15, Algorithms—Speech Detection, VQ/Classification, Pitch/Formant Estimation Algorithms
- Mar 3 - Lecture 15, Speech Waveform Coding—Uniform and Non-Uniform Quantization
- Mar 8 - Lecture 16, Speech Waveform Coding—Adaptive and Differential Quantization
- Mar 16 - Final Exam (8 am-11 am)

---

**Other Potential Topics for Discussion/Term Projects**

- Sinusoidal modeling of speech
- Speech modification and enhancement—slowing down and speeding up speech, noise reduction methods
- Speaker verification methods
- Music coding including MP3 and AAC standards-based methods
- Pitch detection methods
Term Project

• All registered students are required to do a term project. This term project, implemented using Matlab, must be a speech or audio processing system that accomplishes a simple or even a complex task—e.g., pitch detection, voiced-unvoiced detection, speech/silence classification, speech synthesis, speech recognition, speaker recognition, helium speech restoration, speech coding, MP3 audio coding, etc.
• Every student is also required to make a 10-minute Power Point presentation of their term project to the entire class. The presentation must include:
  – A short description of the project and its objectives
  – An explanation of the implemented algorithm and relevant theory
  – A demonstration of the working program – i.e., results obtained when running the program.

Suggestions for Term Projects
1. Pitch detector – time domain, autocorrelation, cepstrum, LPC, etc.
2. Voiced/Unvoiced/Silence detector
3. Formant analyzer/tracker
4. Speech coders including ADPCM, LDM, CELP, Multipulse, etc.
5. N-channel spectral analyzer and synthesizer – phase vocoder, channel vocoder, homomorphic vocoder
6. Speech endpoint detector
7. Simple speech recognizer – e.g. isolated digits, speaker trained
8. Speech synthesizer – serial, parallel, direct, lattice
9. Helium speech restoration system
10. Audio/music coder
11. System to speed up and slow down speech by arbitrary factors
12. Speaker verification system
13. Sinusoidal speech coder
14. Speaker recognition system
15. Speech understanding system
16. Speech enhancement system (noise reduction, post filtering, spectral flattening)

MATLAB Computer Project

The requirements for this project are a short description of the problem containing relevant mathematical theory and objectives of the project, a listing (with sufficient documentation and comments) of the program, and a demonstration that the program works properly.