Course Description

This course covers the basic principles of digital speech processing:
- Review of digital signal processing
- MATLAB functionality for speech processing
- Fundamentals of speech production and perception
- Basic techniques for digital speech processing:
  - short-time energy, magnitude, autocorrelation
  - short-time Fourier analysis
  - homomorphic (convolutional) 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, 10:00 am to 11:20 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

Web Page for Speech Course

Click on **Digital Speech Processing Course** on left-side panel

Download course lecture slides

Course lecture slides (6-to-page)
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

References in Selected Areas of Speech Processing

Speech Synthesis:
• Y. Sagisaka, N. Campbell, and N. Higuchi, Computing Prosody, Springer Verlag, 1996

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
References in Selected Areas of Speech Processing

Speech Recognition:
• X. Huang, A. Acero and H-W. Hon, Spoken Language Processing, Prentice Hall Inc, 2000
• F. Jelinek, Statistical Methods for Speech Recognition, MIT Press, 1998

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, pragmatic, speech production/perception

Digital Speech Processing

Mathematics, derivations, signal processing

Basic understanding of how theory is applied

Ability to implement theory and concepts in working code (MATLAB, C, C++)

Need to understand speech processing at all three levels

Course Outline – ECE 259A – Speech Processing

Jan 8 – Lecture 1, Basic Course Material; Introduction to Digital Speech Processing
Jan 10 – Lecture 2a, Review of DSP Fundamentals
Jan 15 – Lecture 2b, Review of DSP Fundamentals
Jan 17 – Lecture 2c, Review of DSP Fundamentals
Jan 22 – Lecture 2d, Review of DSP Fundamentals
Jan 24 – Lecture 2e, Review of DSP Fundamentals
Jan 29 – Lecture 2f, Review of DSP Fundamentals
Jan 31 – Lecture 2g, Review of DSP Fundamentals
Feb 5 – Lecture 3, Time Domain Methods — Part 1
Feb 7 – Lecture 4, Time Domain Methods — Part 2
Feb 12 – Lecture 5, Frequency Domain Methods — Part 1
Feb 14 – Lecture 6, Frequency Domain Methods — Part 2
Feb 19 – Lecture 7, Frequency Domain Methods — Part 2
Feb 21 – Lecture 8, Frequency Domain Methods — Part 2
Feb 26 – Lecture 9, Frequency Domain Methods — Part 2
Feb 28 – Lecture 10, Frequency Domain Methods — Part 2
Mar 5 – Lecture 11, Frequency Domain Methods — Part 2
Mar 7 – Lecture 12a, Homomorphic Speech Processing — Part 1
Mar 12 – Lecture 13, Homomorphic Speech Processing — Part 2
Mar 14 – Lecture 14, Linear Predictive Coding (LPC) — Part 1
Mar 16 – Lecture 15, Linear Predictive Coding (LPC) — Part 2
Mar 20 – Lecture 16, Linear Predictive Coding (LPC) — Part 2
Mar 22 – 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.