

# Digital Speech Processing

Professor Lawrence Rabiner  
UCSB

Dept. of Electrical and Computer  
Engineering  
Jan-March 2014

# Course Description

This course covers the basic principles of digital speech processing:

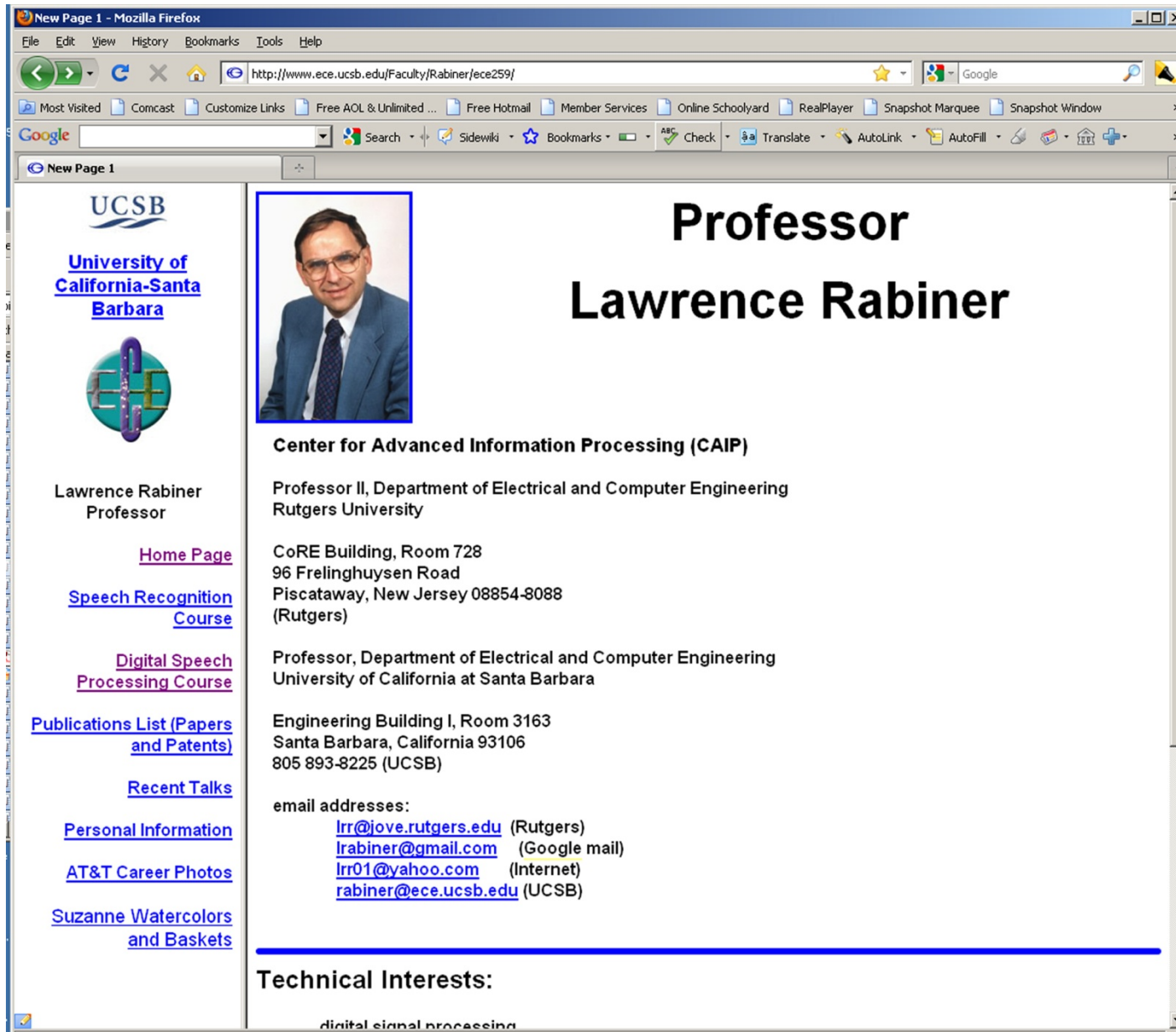
- ***Review of digital signal processing***
- ***MATLAB functionality for speech processing***
- ***Fundamentals of speech production and perception***
- ***Basic techniques for digital speech processing:***
  - short - time energy, magnitude, autocorrelation
  - short - time Fourier analysis
  - homomorphic (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

- **Textbook:** L. R. Rabiner and R. W. Schafer, Theory and Applications of Digital Speech Processing, Prentice-Hall Inc., 2011
- **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, HRH 4164.
- **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



**UCSB**  
**University of California-Santa Barbara**

[Home Page](#)  
[Speech Recognition Course](#)  
[Digital Speech Processing Course](#)  
[Publications List \(Papers and Patents\)](#)  
[Recent Talks](#)  
[Personal Information](#)  
[AT&T Career Photos](#)  
[Suzanne Watercolors and Baskets](#)

**Lawrence Rabiner**  
Professor

**Professor Lawrence Rabiner**

**Center for Advanced Information Processing (CAIP)**

Professor II, Department of Electrical and Computer Engineering  
Rutgers University

CoRE Building, Room 728  
96 Frelinghuysen Road  
Piscataway, New Jersey 08854-8088  
(Rutgers)

Professor, Department of Electrical and Computer Engineering  
University of California at Santa Barbara

Engineering Building I, Room 3163  
Santa Barbara, California 93106  
805 893-8225 (UCSB)

email addresses:  
[lrr@jove.rutgers.edu](mailto:lrr@jove.rutgers.edu) (Rutgers)  
[lrabiner@gmail.com](mailto:lrabiner@gmail.com) (Google mail)  
[lrr01@yahoo.com](mailto:lrr01@yahoo.com) (Internet)  
[rabiner@ece.ucsb.edu](mailto:rabiner@ece.ucsb.edu) (UCSB)

**Technical Interests:**  
digital signal processing

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

# Web Page for Speech Course

New Page 1 - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.ece.ucsb.edu/Faculty/Rabiner/ece259/

Most Visited Comcast Customize Links Free AOL & Unlimited ... Free Hotmail Member Services Online Schoology RealPlayer Snapshot Marquee Snapshot Window

Google Search Sidewiki Bookmarks Check Translate AutoLink AutoFill

New Page 1

**UCSB**  
**University of**  
**California-Santa**  
**Barbara**



Lawrence Rabiner  
Professor

[Home Page](#)

[Speech Recognition Course](#)

[Digital Speech Processing Course](#)

[Publications List \(Papers and Patents\)](#)

[Recent Talks](#)

[Personal Information](#)

[AT&T Career Photos](#)

[Suzanne Watercolors and Baskets](#)

## Digital Speech Processing Course (Winter 2011)

### No Cheating Policy:

No cheating declaration: ([No Cheating Policy.pdf](#))

### Monograph on Digital Speech Processing: [Introduction to Digital Speech processing.pdf](#)

### Lectures:

Introductory Material: ([basic course material\\_winter\\_2011.pdf](#)), 6-to-a-page: ([basic course material\\_winter\\_2011\\_6tp.pdf](#))

Lecture 1: Introduction to Digital Speech Processing: ([Lecture 1\\_fall\\_2010\\_6tp.pdf](#)) : 6-to-a-page: ([Lecture 1\\_fall\\_2010\\_6tp.pdf](#))

Lecture 2: Review of DSP Fundamentals: ([Lecture 2\\_fall\\_2010.pdf](#)) : 6-to-a-page: ([Lecture 2\\_fall\\_2010\\_6tp.pdf](#))

Lecture 3: Acoustic Theory of Speech Production: ([Lecture 3\\_fall\\_2010\\_6tp.pdf](#)) : 6-to-a-page: ([Lecture 3\\_fall\\_2010\\_6tp.pdf](#))

Lecture 4: Speech Perception--Auditory Models, Sound Perception Models, MOS Methods: ([Lecture 4\\_fall\\_2010.pdf](#)) : 6-to-a-page: ([Lecture 4\\_fall\\_2010\\_6tp.pdf](#))

Lectures 5-6: Sound Propagation in the Vocal Tract: ([Lectures 5-6\\_fall\\_2010.pdf](#)) : 6-to-a-page: ([Lectures 5-6\\_fall\\_2010\\_6tp.pdf](#))

Lectures 7-8: Time Domain Methods in Speech Processing: ([Lectures 7-8\\_fall\\_2010.pdf](#)) : 6-to-a-page: ([Lectures 7-8\\_fall\\_2010\\_6tp.pdf](#))

Methods of Pitch Period Estimation:

Lecture 9: Short-Time Fourier Transform (STFT) Concepts: ([Lecture9\\_fall\\_2010.pdf](#)) : 6-to-a-page: ([Lecture9\\_fall\\_2010\\_6tp.pdf](#))

Lecture 10: Short Time Fourier Analysis Methods--Filter Bank Summation and Overlap Add: ([Lecture 10\\_fall\\_2010.pdf](#)) : 6-to-a-page: ([Lecture10\\_fall\\_2010\\_6tp.pdf](#))

Lecture 11: Speech Representations Based on STFT Analysis-Synthesis Methods: ([Lecture 11\\_fall\\_2010.pdf](#)) : 6-to-a-page: ([Lecture 11\\_fall\\_2010\\_6tp.pdf](#))

Lecture 12: Homomorphic Speech Processing: ([Lecture 12\\_fall\\_2010.pdf](#)) : 6-to-a-page: ([Lecture 12\\_fall\\_2010\\_6tp.pdf](#))

Lecture 13: Linear Predictive Coding (LPC) Methods: ([Lecture 13\\_fall\\_2010.pdf](#)) : 6-to-a-page: ([Lecture 13\\_fall\\_2010\\_6tp.pdf](#))

Download  
course  
lecture slides

# Web Page for Speech Course

New Page 1 - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.ece.ucsb.edu/Faculty/Rabiner/ece259/

Most Visited Comcast Customize Links Free AOL & Unlimited ... Free Hotmail Member Services Online Schoology RealPlayer Snapshot Marquee Snapshot Window

Google Search Sidewiki Bookmarks Check Translate AutoLink AutoFill

New Page 1

UCSB  
University of California-Santa Barbara

Lawrence Rabiner  
Professor

[Home Page](#)

[Speech Recognition Course](#)

[Digital Speech Processing Course](#)

[Publications List \(Papers and Patents\)](#)

[Recent Talks](#)

[Personal Information](#)

[AT&T Career Photos](#)

[Suzanne Watercolors and Baskets](#)

12/22/2010

**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**

- **Textbook:** L. R. Rabiner and R. W. Schafer, Theory and Applications of Digital Speech Processing, Prentice-Hall Inc., 2011
- **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

**Web Page for Speech Course**

Click on Digital Speech Processing Course on left-side panel

Course  
lecture slides  
(6-to-page)



# Web Page for Speech Course

New Page 1 - Mozilla Firefox

File Edit View History Bookmarks Tools Help


http://www.ece.ucsb.edu/Faculty/Rabiner/ece259/

Most Visited Comcast Customize Links Free AOL & Unlimited ... Free Hotmail Member Services Online Schoology RealPlayer Snapshot Marquee Snapshot Window

Google Search Sidewiki Bookmarks Check Translate AutoLink AutoFill

New Page 1

**UCSB**  
**University of**  
**California-Santa**  
**Barbara**



Lawrence Rabiner  
Professor

[Home Page](#)

[Speech Recognition Course](#)

[Digital Speech Processing Course](#)

[Publications List \(Papers and Patents\)](#)

[Recent Talks](#)

[Personal Information](#)

[AT&T Career Photos](#)

[Suzanne Watercolors and Baskets](#)

Lecture 16: Speech Waveform Coding--Adaptive and Differential Quantization: ([Lecture 16 fall 2010.pdf](#))  
: 6-to-a-page: ([Lecture 16 fall 2010 6tp.pdf](#))  
Lecture 17: Speech Coding Methods--Model-Based Approaches: ([Lecture 17 fall 2010.pdf](#)) : 6-to-a-page: ([Lecture 17 fall 2010 6tp.pdf](#))

---

**Homeworks:**

Problem Set 1: ([ps1.pdf](#)); PS1 solution:  
Problem Set 2: ([ps2.pdf](#)); PS2 solution:  
Problem Set 3: ([ps3.pdf](#)); PS3 solution:  
Problem Set 4: ([ps4.pdf](#)); PS4 solution:  
Problem Set 5: ([ps5.pdf](#)); PS5 solution:  
Problem Set 6: ([ps6.pdf](#)); PS6 solution:  
Problem Set 7: ([ps7.pdf](#)); PS7 solution:  
Problem Set 8: ([ps8.pdf](#)); PS8 solution:

---

**Speech Files:**

test\_16k.wav: ([test\\_16k.wav](#))  
ah.wav: ([ah.wav](#))  
should.wav: ([should.wav](#))  
s1.wav: ([s1.wav](#)); pitch period contour for s1.wav: ([pp1.mat](#))  
s2.wav: ([s2.wav](#)); pitch period contour for s2.wav: ([pp2.mat](#))  
s3.wav: ([s3.wav](#)); pitch period contour for s3.wav: ([pp3.mat](#))  
s4.wav: ([s4.wav](#)); pitch period contour for s4.wav: ([pp4.mat](#))  
s5.wav: ([s5.wav](#)); pitch period contour for s5.wav: ([pp5.mat](#))  
s6.wav: ([s6.wav](#)); pitch period contour for s6.wav: ([pp6.mat](#))  
we\_were: ([we were away a year ago\\_lrr.wav](#))  
isolated digit training files: ([digits\\_train.zip](#))  
isolated digit testing files: ([digits\\_test.zip](#))  
isolated digit training files (raw-no endpoints marked): ([digits\\_train\\_raw.zip](#))  
isolated digit testing files (raw-no endpoints marked): ([digits\\_test\\_raw.zip](#))

---

**Matlab Files:**

loadwav.m: ([loadwav.m](#))

Download  
homework  
assignments,  
speech files

# Web Page for Speech Course

New Page 1 - Mozilla Firefox

File Edit View History Bookmarks Tools Help


http://www.ece.ucsb.edu/Faculty/Rabiner/ece259/

Most Visited Comcast Customize Links Free AOL & Unlimited ... Free Hotmail Member Services Online Schoology RealPlayer Snapshot Marquee Snapshot Window

Google Search Sidewiki Bookmarks Check Translate AutoLink AutoFill

New Page 1

UCSB  
University of  
California-Santa  
Barbara



Lawrence Rabiner  
Professor

[Home Page](#)

[Speech Recognition  
Course](#)

[Digital Speech  
Processing Course](#)

[Publications List \(Papers  
and Patents\)](#)

[Recent Talks](#)

[Personal Information](#)

[AT&T Career Photos](#)

[Suzanne Watercolors  
and Baskets](#)

test\_16k.wav: ([test\\_16k.wav](#))  
ah.wav: ([ah.wav](#))  
should.wav: ([should.wav](#))  
s1.wav: ([s1.wav](#)); pitch period contour for s1.wav: ([pp1.mat](#))  
s2.wav: ([s2.wav](#)); pitch period contour for s2.wav: ([pp2.mat](#))  
s3.wav: ([s3.wav](#)); pitch period contour for s3.wav: ([pp3.mat](#))  
s4.wav: ([s4.wav](#)); pitch period contour for s4.wav: ([pp4.mat](#))  
s5.wav: ([s5.wav](#)); pitch period contour for s5.wav: ([pp5.mat](#))  
s6.wav: ([s6.wav](#)); pitch period contour for s6.wav: ([pp6.mat](#))  
we\_were: ([we were away a year ago\\_lrr.wav](#))  
isolated digit training files: ([digits\\_train.zip](#))  
isolated digit testing files: ([digits\\_test.zip](#))  
isolated digit training files (raw-no endpoints marked): ([digits\\_train\\_raw.zip](#))  
isolated digit testing files (raw-no endpoints marked): ([digits\\_test\\_raw.zip](#))

---

Matlab Files:

loadwav.m: ([loadwav.m](#))  
savewav.m: ([savewav.m](#))  
loadraw.m: ([loadraw.m](#))  
saveraw.m: ([saveraw.m](#))  
grayscale.m: ([grayscale.m](#))  
fxquant.m: ([fxquant.m](#))  
pspect.m: ([pspect.m](#))  
spectgr.m: ([spectgr.m](#))  
LPC solutions: ([cholesky\\_full.m](#)), ([durbin.m](#)), ([lattice.m](#))

---

Project Suggestions:

General Project Suggestions: ([Digital Speech Processing Projects.pdf](#))  
LPC Vocoder Project Details: ([LPC Vocoder Project.pdf](#))  
Project Schedule (UCSB-2009):  
User Interface Example (Sound Spectrograms):([GUI\\_plot\\_spectrogram\\_ucsb.m](#)), ([select\\_dir.m](#))

Oded:Material: ([oded\\_material.zip](#))

Download  
MATLAB (.m)  
files; Examine  
Project  
Suggestions



# Course Readings

## Required Course Textbook:

- ***L. R. Rabiner and R. W. Schafer, Theory and Applications of Digital Speech Processing, Prentice-Hall Inc., 2011***

## Recommended Supplementary Textbook:

- T. F. Quatieri, *Principles of Discrete - Time Speech Processing*, Prentice Hall Inc, 2002

## Matlab Exercises:

- C. S. Burrus et al, *Computer-Based Exercises for Signal Processing using Matlab*, Prentice Hall Inc, 1994
- J. R. Buck, M. M. Daniel, and A. C. Singer, *Computer Explorations in Signals and Systems using Matlab*, Prentice Hall Inc, 2002

# Recommended References

- ***J. L. Flanagan, Speech Analysis, Synthesis, and Perception, Springer -Verlag, 2<sup>nd</sup> Edition, Berlin, 1972***
- 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
- J. Deller, Jr., J. G. Proakis, and J. Hansen, *Discrete - Time Processing of Speech Signals*, Macmillan Publishing, 1993
- 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
- R. W. Schafer and J. D. Markel, Editors, *Speech Analysis*, IEEE Press Selected Reprint Series, 1979
- D. G. Childers, *Speech Processing and Synthesis Toolboxes*, John Wiley and Sons, 1999
- K. Stevens, *Acoustic Phonetics*, MIT Press, 1998
- J. Benesty, M. M. Sondhi and Y. Huang, Editors, *Springer Handbook of Speech Processing and Speech Communication*, Springer, 2008.

# References in Selected Areas of Speech Processing

## Speech Coding:

- ***A. M. Kondo, Digital Speech: Coding for Low Bit Rate Communication Systems-2<sup>nd</sup> Edition, John Wiley and Sons, 2004***
- 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:

- ***T. Dutoit, An Introduction to Text - To-Speech Synthesis, Kluwer Academic Publishers, 1997***
- ***P. Taylor, Text-to-Speech Synthesis, Cambridge University Press, 2008***
- J. Allen, S. Hunnicutt, and D. Klatt, *From Text to Speech*, Cambridge University Press, 1987
- Y. Sagisaka, N. Campbell, and N. Higuchi, *Computing Prosody*, Springer Verlag, 1996
- J. VanSanten, R. W. Sproat, J. P. Olive and J. Hirschberg, Editors, *Progress in Speech Synthesis*, Springer Verlag, 1996
- J. P. Olive, A. Greenwood, and J. Coleman, *Acoustics of American English*, Springer Verlag, 1993

# References in Selected Areas of Speech Processing

## Speech Recognition:

- ***L. R. Rabiner and B. H. Juang, Fundamentals of Speech Recognition, Prentice Hall Inc, 1993***
- ***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
- H. A. Bourlard and N. Morgan, *Connectionist Speech Recognition-A Hybrid Approach*, Kluwer Academic Publishers, 1994
- C. H. Lee, F. K. Soong, and K. K. Paliwal, Editors, *Automatic Speech and Speaker Recognition*, Kluwer Academic Publisher, 1996



## References in Digital Signal Processing

- A. V. Oppenheim and R. W. Schaffer, *Discrete - Time Signal Processing, 3<sup>rd</sup> Ed.*, Prentice-Hall Inc, 2010
- L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*, Prentice Hall Inc, 1975
- S. K. Mitra, *Digital Signal Processing-A Computer-Based Approach*, Third Edition, McGraw Hill, 2006
- S. K. Mitra, *Digital Signal Processing Laboratory Using Matlab*, McGraw Hill, 1999

# The Speech Stack

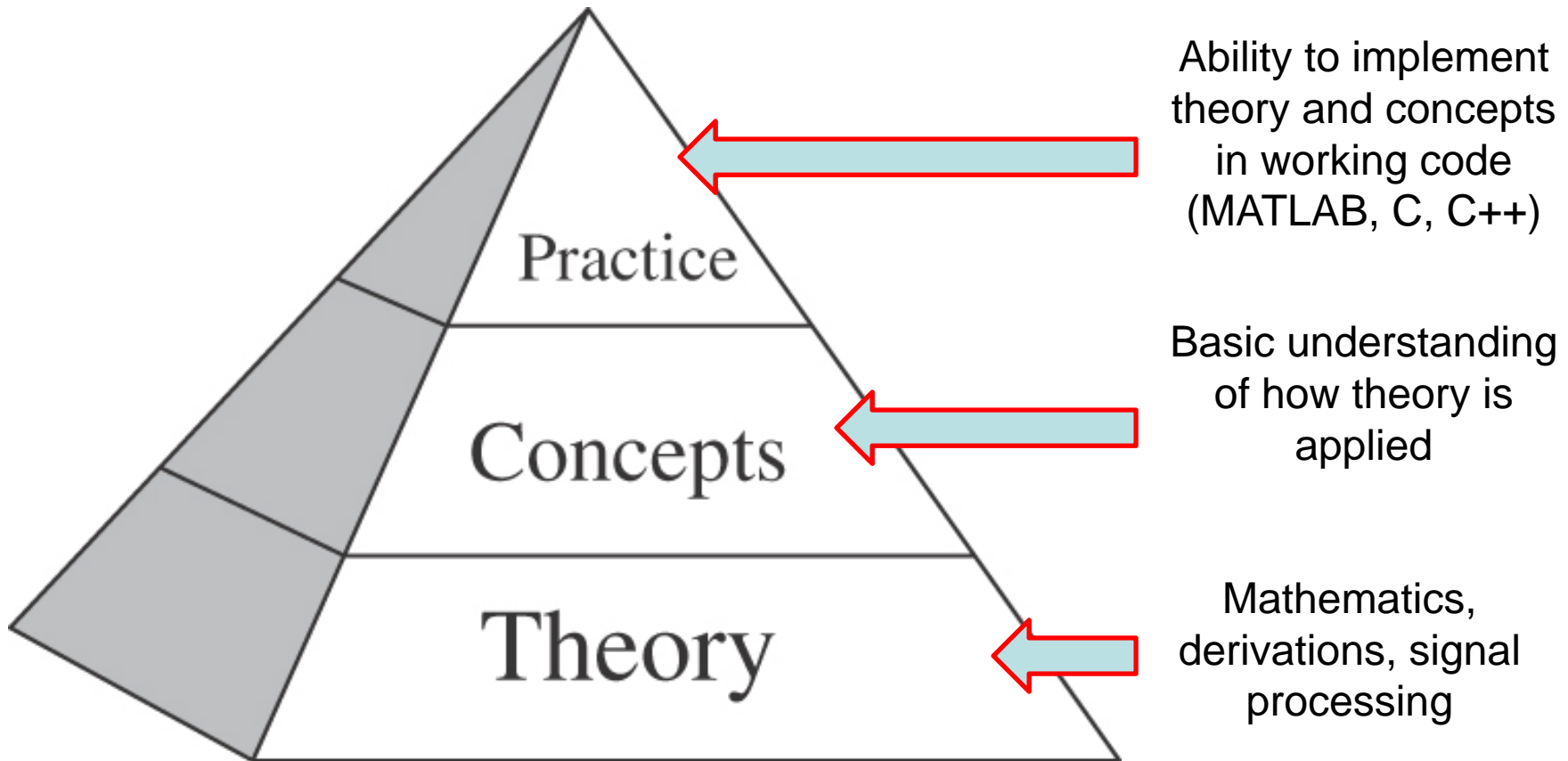
***Speech Applications*** — coding, synthesis, recognition, understanding, verification, language translation, speed-up/slow-down

***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



**Need to understand speech processing at all three levels**

# Course Outline – ECE 259A – Speech Processing

- Jan 7 - Lecture 1, Basic Course Material; Introduction to Digital Speech Processing
- Jan 9 - Lecture 2a, Review of DSP Fundamentals
- Jan 14 - Lecture 2b, Review of DSP Fundamentals
- Jan 16 - Lecture 3a, Acoustic Theory of Speech Production
- Jan 21 - Lecture 3b, Lecture 4, Speech Perception—Auditory Models
- Jan 23 - Lecture 5, Sound Propagation in the Vocal Tract -- Part 1
- Jan 28 - Lecture 6, Sound Propagation in the Vocal Tract -- Part 2
- Jan 30 - Lecture 7, Time Domain Methods -- Part 1
- Feb 4 - Lecture 8, Time Domain Methods -- Part 2
- Feb 6 - Lecture 9, Frequency Domain Methods -- Part 1
- Feb 11- Lecture 10-11, Frequency Domain Methods -- Part 2
- **Feb 13 - Mid - Term Exam**
- Feb 18 - Lecture 12a, Homomorphic Speech Processing -- Part 1
- Feb 20 - Lecture 12b, Homomorphic Speech Processing -- Part 2
- Feb 25 - Lecture 13, Linear Predictive Coding (LPC) -- Part 1
- Feb 27 - Lecture 14, Linear Predictive Coding (LPC) -- Part 2
- Mar 4 - Lecture Algorithms – Part 1
- Mar 6 - Lecture Algorithms – Part 2
- **Mar 11 - Term Project Presentations – Group 1**
- **Mar 13 - Term Project Presentations – Group 2**
- **Mar 18 - 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/group is also required to make a 10-15 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.