

Digital Speech Processing

Last updated: 2/12/2015

Author Information

Professor Lawrence Rabiner
Rutgers University

Professor Ronald Schafer
Stanford University

Course Details

Description

This course covers the basic principles of Digital Speech Processing (DSP):

- 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 (algorithms)
 - 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.

Prerequisites

- Basic Digital Signal Processing
- Knowledge of MATLAB

Original Course Documents

[Source file URL](#)

Course Contents

Course contents can be downloaded [here](#).

Lectures

- Introductory Material
- Introduction to MATLAB Speech Processing/Exercises
- MATLAB Speech Processing Apps
- Lecture 1: Introduction to Digital Speech Processing
- Lecture 2: Review of DSP Fundamentals
- Lecture 3: Acoustic Theory of Speech Production
- Lecture 4: Speech Perception--Auditory Models, Sound Perception Models, MOS Methods
- Lectures 5-6: Sound Propagation in the Vocal Tract
- Lectures 7-8: Time Domain Methods in Speech Processing
- Methods of Pitch Period Estimation
- Lecture 9: Short-Time Fourier Transform (STFT) Concepts
- Lecture 10: Short Time Fourier Analysis Methods--Filter Bank Summation and Overlap Add
- Lecture 11: Speech Representations Based on STFT Analysis-Synthesis Methods
- Lecture 12: Homomorphic Speech Processing
- Lecture 13: Linear Predictive Coding (LPC) Methods
- Lecture 14: LPC--Frequency Domain Interpretations, Methods for Synthesis and Vocoding
- Lecture Algorithms
- Lecture 15: Speech Waveform Coding--Uniform and Non-Uniform Quantization
- Lecture 16: Speech Waveform Coding--Adaptive and Differential Quantization
- Lecture 17: Speech Coding Methods--Model-Based Approaches

Homework Assignments

- Problem Set 1
- Problem Set 2
- Problem Set 3
- Problem Set 4
- Problem Set 5
- Problem Set 6
- Problem Set 7
- Problem Set 8

Speech Files

- test_16k.wav: (test_16k.wav)
- ah.wav: (ah.wav)
- beep_fs_10000.wav: (beep_fs_10000.wav)
- beep_fs_16000.wav: (beep_fs_16000.wav)

- should.wav: (should.wav)
- s5_synthetic.wav: (s5_synthetic.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

- MATLAB Speech Processing and GUI Files: (matlab_speech_2011.pdf)
- 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)
- User Interface Example (Sound Spectrograms):(GUI_plot_spectrogram_ucsb.m), (select_dir.m)

MATLAB DSP Apps

[Over 60 speech processing apps](#) developed in MATLAB available for download on MATLAB Central.

These apps are designed to give students and instructors hands-on experience with digital speech processing basics, fundamentals, representations, algorithms, and applications.

In addition, a [webinar](#) describes the set of speech processing apps and shows how they can be used to enhance the teaching and learning of digital speech processing.

Textbooks

Required Course Textbook:

- L. R. Rabiner and R. W. Schafer, Theory and Applications of Digital Speech Processing, PrenticeHall 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

Resources

Recommended References:

- J. L. Flanagan, Speech Analysis, Synthesis, and Perception, Springer -Verlag, 2nd 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. Kondoz, Digital Speech: Coding for Low Bit Rate Communication Systems-2 nd 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
- **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
- **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, 3rd 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

Learn more about MathWorks academic resources:

- [MATLAB Courseware](#)
- [Hardware Resources](#)
- [Classroom Resources](#)
- [MATLAB Examples](#)
- [Books](#)
- [Tutorials](#)
- [Webinars](#)
- [Technical Articles](#)