CSCE 630: Speech Processing
Fall 2014

Time: MWF 11:30am-12:20pm, **Room**: HRBB 204

Instructor: Ricardo Gutierrez-Osuna
Office: 506A HRBB
Office Hours: By appointment
Phone: (979) 845-2942
Email: rgutier@cse.tamu.edu
URL: http://faculty.cs.tamu.edu/rgutier/

Catalog Description: Speech production and perception (speech apparatus, articulatory/auditory phonetics); mathematical foundations (sampling, filtering, probability, pattern recognition); speech analysis and coding (short-time Fourier analysis, linear prediction, cesptrum); speech recognition (dynamic time warping, hidden Markov models, language models); speech synthesis (frontend, back-end); speech modification (overlap-add, enhancement, voice conversion). Prerequisites: ECEN 314 or equivalent or approval of instructor. Basic knowledge of signals and systems, linear algebra, probability and statistics. Programming experience in a high-level language is required.

Prerequisites: ECEN 314 or equivalent, or permission of the instructor. Basic knowledge of signals and systems, linear algebra, probability and statistics. Programming experience in a high-level language is required.

Textbook:
The course will be based on lecture slides developed by the instructor from several sources.

References:
- B. Gold, N. Morgan and D. Ellis, Speech and Audio Signal Processing: Processing and perception of speech and music, 2nd Ed., Wiley, 2011
- J. Holmes and W. Holmes. Speech Synthesis and Recognition, 2nd Edition. CRC Press, 2001
- P. Taylor, Text-to-speech synthesis, Cambridge University Press, 2009
- X. Huang, A. Acero and H.-W. Hon, Spoken Language Processing, Prentice Hall, 2001
- L. R. Rabiner and R. W. Schafer, Introduction to Digital Speech Processing, Foundations and Trends in Signal Processing 1(1–2), 1–194, 2007
- J. Benesty, M. M. Sondhi and Y. Huang (Eds.), Springer Handbook of Speech Processing, 2008
- T. Dutoit and F. Marques, Applied Signal Processing, a Matlab-based proof-of-concept, Springer, 2009

Course Objectives: This course seeks to familiarize students with
- Fundamental concepts of speech production and speech perception
- Mathematical foundations of signal processing and pattern recognition
- Computational methods for speech analysis, recognition, synthesis, and modification

Course Outcomes: Upon satisfactory completion of the course, the student will be able to:
- Manipulate, visualize, and analyze speech signals
- Perform various decompositions, codifications, and modifications of speech signals
- Build a complete speech recognition system using state of the art tools

Course Outline

- Introduction (3 lectures)
  - Course introduction
  - Speech production and perception
  - Organization of speech sounds
- Mathematical foundations (4 lectures)
  - Signals and transforms
  - Digital filters
  - Probability, statistics and estimation theory
  - Pattern recognition principles
- Speech analysis and coding (4 lectures)
  - Short-time Fourier analysis and synthesis
  - Linear prediction of speech
  - Source estimation
  - Cepstral analysis
- Speech and speaker recognition (6 lectures)
  - Template matching
  - Hidden Markov models
  - Refinements for HMMs
  - Large vocabulary continuous speech recognition
  - The HTK speech recognition system
  - Speaker recognition
- Speech synthesis and modification (4 lectures)
  - Text-to-speech front-end
  - Text-to-speech back-end
  - Prosodic modification of speech
  - Voice conversion

Grading: The course grade will be the weighted sum of four grades. Grading will be straight scale (90-100 A, 80-89 B, 70-79 C, 60-69 D, below 60 F). These numeric thresholds may be lowered due to clustering, but will not be raised.

- **Homework**: There will be three homework assignments, distributed every 2-3 weeks during the first part of the semester. Homework assignments will emphasize the implementation (programming) of material presented in class. *Homework assignments must be done individually*.
- **Tests**: There will be a midterm exam and a final exam. All tests will be closed-books, closed--notes. One double-sided, hand-written sheet (8.5"×11") will be allowed. Tests will have an emphasis on new material from the class notes.
- **Project**: The last part of the semester will be dedicated to a term project. Students are encouraged to propose projects related to their own research. The projects must be performed in groups of up to three people. Projects may emphasize the application of existing tools, the development of new tools, or the design of new algorithms. Projects will be graded by their content (75%) and the quality of a classroom presentation (25%) at the end of the semester. G*rading criteria for the project presentation and final report are available in the course webpage.*

	Weight (%)
Homework	40
Project	30
Midterm	15
Final Exam	15

Homework submissions. Homework assignments are due at *11:00am* on the due date. Electronic material will be submitted with the "*turnin*" utility at https://csnet.cs.tamu.edu; hardcopies will be submitted directly to the instructor. Email submissions will not be accepted. Note that '*turnin*' has a maximum file size that can be submitted.

Late submissions. Late submissions (i.e., as flagged by *csnet*) will receive a 15% penalty on the total grade of the assignment; the penalty will increase by an additional 15% every 24 hours. Hardcopies of late submissions must be *date and time stamped* by the staff in the Computer Science main office. An assignment is considered submitted when ALL components of the assignment have been submitted; e.g., late submission of one problem in a homework will cause your entire homework to be considered as a late submission.

Missed Tests: Missed tests can only be made up in case of emergency or work conflicts, and will require supporting documentation. Whenever possible, these issues should be discussed with the instructor prior to the conflicting date.

Collaboration vs. Academic Dishonesty: Students are encouraged to exchange ideas and form study groups to discuss the course material, and prepare for homework assignments and tests. However, discussions regarding homework assignments should be kept at the conceptual level (i.e., sharing code is not allowed). Scholastic dishonesty will not be tolerated in homework assignments, tests or projects. For a list of examples of scholastic dishonesty see Section 20 of the TAMU Student Rules (http://student-rules.tamu.edu/).

Academic Integrity Statement
"*An Aggie does not lie, cheat, or steal or tolerate those who do.*" Please review the Aggie Honor Code and Honor Council Rules and Procedures at http://www.tamu.edu/aggiehonor.

Course Schedule

Week	Date	Classroom meeting	Materials due
1	9/1	Course introduction	
	9/3	Speech production and perception	
	9/5	Organization of speech sounds	
2	9/8	Signals and transforms	
	9/10	Signals and transforms	
	9/12	Digital filters	
3	9/15	Digital filters	
	9/17	Short-time Fourier analysis and synthesis	
	9/19	Short-time Fourier analysis and synthesis	HW1 assigned
4	9/22	Linear prediction of speech	
	9/24	Linear prediction of speech	
	9/26	Source estimation	
5	9/29	Source estimation	
	10/1	Cepstral analysis	
	10/3	Cepstral analysis	HW1 due
6	10/6	Probability, statistics, and estimation theory	HW2 assigned
	10/8	Probability, statistics, and estimation theory	
	10/10	Pattern recognition principles	
7	10/13	Pattern recognition principles	
	10/15	Template matching	
	10/17	Hidden Markov models	
8	10/20	Hidden Markov models	
	10/22	Review/catch-up day	HW2 due
	10/24	Midterm exam	
9	10/27	Refinements for HMMs	
	10/29	Refinements for HMMs	
	10/31	**HTK speech recognition system**	HW3 assigned
10	11/3	**HTK speech recognition system**	
	11/5	Large vocabulary continuous speech recognition	
	11/7	Large vocabulary continuous speech recognition	
11	11/10	Speaker recognition	
	11/12	Speaker recognition	
	11/14	Speech synthesis	
12	11/17	Speech synthesis	HW3 due
	11/19	Speech synthesis	
	11/21	Speech modification	
13	11/24	Proposal presentations	Project proposal
	11/26	Proposal presentations	
	11/28	Thanksgiving holiday	
14	12/1	Speech modification	
	12/3	Speech modification	
	12/5	Review/catch-up day	
15	12/8	Final exam	
	12/10	Reading day	
	12/12	No class	
16	12/17	Project presentations: 10:30am-12:30pm	Project report