

Speech Recognition

Graduate Class: EE 7730

Undergraduate: Special Topics

Syllabus

Teacher: Kay Berkling
Term: Winter 2006

Why should you take a course in speech recognition if you are not going into research into that area? Because you learn a handful of skills that can be used for any task that has the steps of taking a problem, analyzing it, collecting data, representing the data, clustering it and recognizing it automatically. Speech, image processing, customer data analysis, stock market prediction, weather prediction, and many more topics will use similar methods as we will study in this class. So whether you are in industry or academia, you will be working with data and if you want to understand what your data means then this class will teach you methodologies for achieving your goal.

In addition, if you are interested in speech recognition, we can help you to get internships or move on to graduate programs in this area because we have some connections. Finally, don't forget that we have a project right here in the department that needs help of experts in speech recognition, image recognition and DSP!

L#	Content	L#	Content
1	Introduction	7	Context
2	Speech Production	8	Language Modeling
3	Speech Manifestation in English	9	Spoken Dialogue Systems
4	Signal Processing	10	Online learning
5	Feature Preparation	11	Win Carib
6	Classification	12	Student Presentations

Lab 1 : Record and read your own speech (KTH wavesurf, SpeechFileSystem)

Lab 2 : Feature Extraction (ICSI SPRACH)

Lab 3 : Phoneme recognition (NN)

Lab 4 : Speech Recognition (HTK)

Lab 5: Speech Synthesis (Festival)

Lab 6 : Spoken Dialogue Systems (OGI)

Grading :

30% Attendance

40% Labs

10% Papers (undergrads can work in groups)

20% Project (undergrads can work in groups)

Reading Material is provided in Class and supported with website with supporting material.

Students will be expected to each read papers and present current research issues in one of these areas. Students pick one of the areas below or suggest their own – according to personal interest.

1. Speaker Identification
2. Issues in multi-media interactions (smart rooms)
3. Language Identification
4. Language translation
5. Virtual reality with Speech recognition
6. Combining image and speech recognition to identify person in room
7. Speech Synthesis
8. Lip synching
9. Generation of cartoons (voice – synch – movement)

Projects are to be chosen, discussed and defined in detail with professor. Undergraduates can work in groups. Graduates work individually unless two projects have interface with each other.

1. 3-D Speech recognition with our microphone array
2. Building a complete speech recognition system with freeware tools
3. Design and collect data for Puerto Rican English. What issues are there to take care of? (Linguistic, acoustic setting, persons, labeling, publishing, differences to American English highlighted?)
4. Learn to program with Dragon Speech system API
5. Connect face with speech synthesis system (Festival or ATT voices)

Prerequisites for the course:

Graduate:

Stochastic Processes

Mathematical Methods

Undergraduate:

DSP

Random Processes