Computer Science
CSC 83060: Speech and Audio Understanding
Fall, 2016

Instructor: Prof. Michael Mandel

Email mim@sci.brooklyn.cuny.edu
Phone 718-951-5600 x2053
Web http://mr-pc.org
Office Graduate Center: Room 4410, Brooklyn College: 2232 Ingersoll
Office hours Fri 10–11:30am at the Graduate Center and by appointment

Rationale In October, 2014, the number of mobile devices surpassed the number of humans in the world. In addition to data communications, almost all of these devices are equipped with at least one microphone and speaker. This course gives a thorough foundation in analyzing, understanding, and manipulating audio signals recorded by these devices. It will cover traditional algorithms that provide human-human and human-machine communication along with promising solutions to some of their limitations. It will also cover recent applications to musical audio and environmental signals.

Description "Machine listening" is a multidisciplinary field at the intersection of signal processing, machine learning, and psychoacoustics. This course will begin by introducing necessary material from those fields to provide a foundation for the rest of the course. Machine listening is primarily concerned with analyzing and understanding three types of signals: speech, music, and environmental sounds and these will be the focus of the course. We will also consider additional applications that require the creation or manipulation of these sounds in speech and music.

Topic list Topics may include, but are not limited to:

1. Fundamentals of
   (a) Digital signal processing
   (b) Acoustics
   (c) Auditory perception
   (d) Machine learning
2. Core machine listening topics
   (a) Speech models and speech synthesis
   (b) Speech recognition features and acoustic modeling, including noise robustness issues
   (c) Speech recognition language modeling, search, and weighted finite state transducers
   (d) Music analysis and modeling
   (e) Source separation and spatial sound
   (f) Environmental sound analysis

Learning goals Students will be able to demonstrate working knowledge of the theory, algorithms, and software involved in

1. Speech recognition
2. Speech synthesis
3. Music information retrieval
4. Environmental audio analysis

Assessment Assessment will be based on two main components: weekly paper discussions and a final project.

Each week, the class will read one review paper giving an overview of a topic and one or two papers presenting novel contributions in that area. A student will present each contribution paper, receiving feedback on their presentation from the class and from me. Everyone will write two paragraphs on Blackboard in response to the review paper, one summarizing it and one identifying gaps in current knowledge as presented in the paper and describing future directions the work could take. I will lead the discussion of the review paper.

In order to encourage final projects that are novel, well thought out, and well executed, activities related to them will be spread through the semester. Weekly writing assignments will also include a third paragraph discussing progress on the project. There will be a midterm project proposal presentation, where each group will present their current plan and receive feedback on it. There will be final project presentations at the end of the semester, with additional feedback. Final paper submissions slightly after the presentations should incorporate this feedback.

The above components will be weighted as follows:

Weekly writing	20%
Paper presentation	10%
Participation / attendance	10%
Project proposal presentation	20%
Final project presentation	10%
Final project paper	30%

Online Resources Slides, assignments, and readings will be posted on the course website:

http://mr-pc.org/t/csc83060/

The course will also have a **Blackboard** site with a dropbox for each assignment, grades, and announcements.

Grading All homeworks and projects should be turned in via blackboard at least 30 minutes prior to the beginning of the corresponding class period. Homeworks turned in late will be penalized 10% for each day they are late. A project that is turned in two days late and would have received a 100% will instead receive an 80%.

Key Dates Project proposal presentations will be on **October 7, 2016** and final project presentations will be on **December 9, 2016**.

University policy on Academic Integrity The faculty and administration of CUNY support an environment free from cheating and plagiarism. Each student is responsible for being award of what constitutes cheating a plagiarism and for avoiding both. The complete text of the CUNY Academic Integrity Policy can be found at http://web.cuny.edu/academics/info-central/policies/academic-integrity.pdf. CUNY guidelines for avoiding and detecting plagiarism can be found at http://www.gc.cuny.edu/CUNY_GC/media/CUNY-Graduate-Center/PDF/Policies/General/AvoidingPlagiarism.pdf. If a faculty member suspects a violation of academic integrity and, upon investigation, confirms that violation, or if the student admits the violation, the faculty member MUST report the violation.

Course policy on Academic Integrity While you are encouraged to discuss the course material and assignments with your classmates and anyone else you might like, all submitted individual assignments must be *strictly your own work*. If you include any work from other sources, including existing web pages, publications, books, or conversations, it should be explicitly cited with proper credit given to the original author. In the case of copying, both the copier and the copy-ee may be equally guilty.

You may work in small groups for the course project, in which case the groups will be established early in the term and each group will turn in a single final project report and present a single midterm progress presentation and final project presentation. Project-related journal entries are required from each individual describing your personal progress on the project.