Macintosh HD:Users:ers:Dropbox:ACTIVE:JHU-logos:WSE:small:PDF:whiting.small.horizontal.black.pdf

**Electrical and Computer Engineering 520.680**

Speech and Auditory Processing by Humans and Machines

**Spring Semester, 2022 (3 credits)**

**Instructor**

Professor Hynek Hermansky, hynek@jhu.edu

Office: 324F

Office hours: by appointment

**Teaching Assistants**

Samik Sadhu <ssadhu1@jhu.edu>

Office: 321

Office hours: by appointment

**Meetings**

Tuesday, Thursday 10:30-11:45

Homewood Campus, Shaffer 304

**Textbooks**

No textbook is required, lecture slides and relevant reading materials will be put on Blackboard **after** each lecture. However, if you find that you want to learn more, beyond what we cover in the class or if you find the provided materials to be not sufficient, you may consult some sections from

1. Rabiner, Lawrence, and Ronald Schafer. *Theory and applications of digital speech processing*. Prentice Hall Press, 2010.
2. Gold, Ben, Nelson Morgan, and Dan Ellis. *Speech and audio signal processing: processing and perception of speech and music*. John Wiley & Sons, 2011.
3. Deng, Li, and Douglas O'Shaughnessy. *Speech processing: a dynamic and optimization-oriented approach*. CRC Press, 2003.
4. Miller, George Armitage. "Language and communication." (1951).
5. Cherry, Colin. "On human communication." (1966): 169-178
6. Moore, Brian CJ. *An introduction to the psychology of hearing*. Brill, 2012.

**Online Resources**

Please log in to Blackboard for all materials related to this course

**Course Information**

* The course relevant to building advanced systems for information extraction from speech and auditory signals. It introduces some relevant historical efforts for information processing of speech and audio signals and basic concepts of human auditory perception and human production and perception of speech. The main goal of the course is in implementation of relevant knowledge of human speech information processing in engineering systems for information extraction from speech signals, emphasizing power of the modern data-guided machine learning techniques. Basic knowledge of signal processing is assumed and the previous completion of the EN.520.445 or EN.520.645 is beneficial.
* **Prerequisites** None
* **Elective or Selective Elective**

**Course Goals**

* Introducing engineering students to principles of processing of audio and visual signals by human listeners and by machines, aiming at applications of this knowledge in designing engineering systems for processing of auditory signals.

Specific Outcomes for this course are that

* Students will learn to appreciate the power of applying knowledge about human audio-visual perception in design of engineering systems for processing auditory signals

This course will address the following Criterion 3 Student Outcomes

* an ability to identify, formulate, and solve complex engineering problems by applying principles of engineering, science, and mathematics
* an ability to apply engineering design to produce solutions that meet specified needs with consideration of public health, safety, and welfare, as well as global, cultural, social, environmental, and economic factors
* an ability to communicate effectively with a range of audiences
* an ability to recognize ethical and professional responsibilities in engineering situations and make informed judgments, which must consider the impact of engineering solutions in global, economic, environmental, and societal contexts
* an ability to function effectively on a team whose members together provide leadership, create a collaborative and inclusive environment, establish goals, plan tasks, and meet objectives
* an ability to develop and conduct appropriate experimentation, analyze and interpret data, and use engineering judgment to draw conclusions
* an ability to acquire and apply new knowledge as needed, using appropriate learning strategies

**Course Topics**

Introduction

Why speech?

Speech and hearing connection

Human speech communication chain

Basics of information theory

Measuring information

Entropy

Information in printed text

Channel capacity

Information transfer through communication channel

Information in speech signal and in speech messages

Implications in engineering

Human speech communication chain again

Short-time spectral analysis

Review of Fourier transform

Sampling and quantization

Short-time Fourier analysis

Uncertainty principle in spectral analysis

Basic principles of speech production

Linear model of speech production

Propagation of sound in air

Quarter-wave resonator

Half-wave resonators

Perturbation theory

Constricting acoustic tube – introducing redundancies in frequency

Effect of resonance frequency of front cavity of speech spectrum

Introduction of anti-resonances

Basic properties of hearing

Physiology of hearing periphery

Physiology of higher hearing stages

Simultaneous and temporal masking

Critical bands of hearing (Bark scale)

Perception of pitch (Mel scale)

Time in hearing

Perception of modulations

History of speech processing

Humans speech communication chain and speech engineering

Newton, von Helmholtz, von Kempelen, Fant, ….

Radio Rex

Voder and Vocoder

Spectrograph and spectrogram

The first “knowledge based” speech recognizer (R. Galt)

The first “data-based” speech recognizers (Davis et al, Smith,..)

Fundamentals of speech recognition

Template matching

Zerocrossings of high-pass and low-pass speech signal

Distance measures

Dynamic time warping

Principles of stochastic recognition

Hidden Markov Model training and recognition

Estimating spectral envelope

Cepstral analysis

Pole-zero decomposition

Spectral envelope and source estimation

Linear Predictive Analysis

Autocorrelation and covariance methods

Spectral envelope approximation by LP

Spectral transform LP

Perceptual techniques for spectral envelope estimation

Mel cepstrum

Perceptual linear prediction

Effect of PLP model order

Suppressing speaker-dependent element of speech

Hypothetical speech production - speech perception link

Estimating speech source parameters

Artificial neural nets

Deriving posterior probabilities of speech sounds

DNN/HMM hybrid

Using ANNs for deriving speech features (TANDEM

Classifier of TemporAl PatternS (TRAPS)

Speech dynamics

Coarticulation and introducing temporal redundancies in speech

Measuring vocal tract movements

Correlations between tract movements and speech envelopes dynamics

Modulation spectrum of speech

Perception of modulations by human hearing

Intelligibility of speech with modified dynamics

Employing spectral dynamics in machine recognition of speech

Formant-less vowel experiment

Delta features

RASTA filter

Data-guided processing

Speech-hearing link again

Mutual information between speech feature and labels

Mutual information in frequency and in time

Linear discriminant analysis and design of spectral projections

Linear discriminant analysis and design of temporal RASTA filters

Linear discriminant analysis and design of 2-D spectro-temporal filters

Human speech recognition

Effect of vocabulary size

Words in-context and out-of-context (parallel context channel)

Recognition of nonsense syllables (Fletcher et al)

Product of recognition probabilities of phonemes

Recognition of high-pass and low-pass filtered noisy syllables

Recognition accuracy and articulation index

Product of error probabilities in sub-bands

Multi-stream recognition of speech

Consistency with human processing of sensory signals

Main engineering issues

**Course Expectations & Grading**

Requirements for the successful completion of the course are:

* Interest, curiosity, class activity
* Regular weekly or bi-weekly home-works
* Mid-term exam (within the second month of the course)
* Second late mid-term exam (within the third month of the course)
* Class project

Grading will be based on

25 % homeworks

25 % mid-term exam

25 % second mid-term exam

25% final project (its choice and its mastering)

Your interest, curiosity and class activity will not be formally graded in either course but will

be used as an additional criterion in the final grading

**Key Dates**

This will be on Blackboard

**Assignments & Readings**

Will be posted on the Blackboard site for this course.

Personal Wellbeing

Because of the ongoing COVID-19 pandemic special requirements will be in effect this term, and these may vary during the term. Please keep updated with these at the following sites: o University information: https://covidinfo.jhu.edu/

o Whiting School of Engineering information: https://engineering.jhu.edu/covid-19/

· As of the start of the term all students, instructors and staff must complete health screening daily using the Prodensity app before coming to campus. Masks must be worn properly at all times while in the classroom and other indoor spaces. Vaccination is required unless an exception has been granted by the university for health or religious reasons. Periodic asymptomatic testing may be required. Please follow the university guidance faithfully.

· The Johns Hopkins COVID-19 Call Center (JHCCC), which can be reached at 443-287-8500 seven days a week from 7 a.m. to 7 p.m., supports all JHU students, faculty, and staff experiencing COVID-19 symptoms. Primarily intended for those currently within driving distance of Baltimore, the JHCCC will evaluate your symptoms, order testing if needed, and conduct contact investigation for those affiliates who test positive. More information on the JHCCC and testing is on the coronavirus information website.

· If you are sick please notify me by email so that we can make appropriate accommodations should this affect your ability to attend class, complete assignments, or participate in assessments. The Student Health and Wellness Center is open and operational for primary care needs. If you would like to speak with a medical provider, please call 410-516-8270, and staff will determine an appropriate course of action. See also https://studentaffairs.jhu.edu/student-life/student-outreach-support/absences-from-class/illness-note-policy/

· All students with disabilities who require accommodations for this course should contact me at their earliest convenience to discuss their specific needs. If you have a documented disability, you must be registered with the JHU Office for Student Disability Services (101 Shaffer Hall; 410-516-4720; http://web.jhu.edu/disabilities/) to receive accommodations.

· Students who are struggling with anxiety, stress, depression or other mental health related concerns, please consider connecting with resources through the JHU Counseling Center. The Counseling Center will be providing services remotely to protect the health of students, staff, and communities. Please reach out to get connected and learn about service options based on where you are living this fall at 410-516-8278 and online at http://studentaffairs.jhu.edu/counselingcenter/.

· Student Outreach & Support helps students manage physical and mental health concerns, personal and family emergencies, financial issues, and other obstacles that may arise during their college experience. Students can self-refer or refer a friend who may need extra support or help getting connected to resources. To connect with SOS, please visit this website: https://studentaffairs.jhu.edu/student-life/student-outreach-support/ or email deanofstudents@jhu.edu, call 410-516-7857, or students can schedule to meet with a Case Manager by visiting the Student Outreach & Support website and filling out a referral form online.

Classroom Climate

As your instructor, I am committed to creating a classroom environment that values the diversity of experiences and perspectives that all students bring. Everyone here has the right to be treated with dignity and respect. I believe fostering an inclusive climate is important because research and my experience show that students who interact with peers who are different from themselves learn new things and experience tangible educational outcomes. Please join me in creating a welcoming and vibrant classroom climate. Note that you should expect to be challenged intellectually by me, the TAs, and your peers, and at times this may feel uncomfortable. Indeed, it can be helpful to be pushed sometimes in order to learn and grow. But at no time in this learning process should someone be singled out or treated unequally on the basis of any seen or unseen part of their identity. If you ever have concerns in this course about harassment, discrimination, or any unequal treatment, or if you seek accommodations or resources, I invite you to share directly with me or the TAs. I promise that we will take your communication seriously and to seek mutually acceptable resolutions and accommodations. Reporting will never impact your course grade. You may also share concerns with the department/center chair/head/director ([Chair’s Name and Email]), the Director of Undergraduate Studies ([DUS Name and Email]), the Assistant Dean for Diversity and Inclusion (Darlene Saporu, dsaporu@jhu.edu), or the Office of Institutional Equity (oie@jhu.edu). In handling reports, people will protect your privacy as much as possible, but faculty and staff are required to officially report information for some cases (e.g. sexual harassment).

Family Accommodations Policy

You are welcome to bring a family member to class on occasional days when your responsibilities require it (for example, if emergency child care is unavailable, or for health needs of a relative). In fact, you may see my children in class on days when their school is closed. Please be sensitive to the classroom environment, and if your family member becomes uncomfortably disruptive, you may leave the classroom and return as needed.

University Policy on Incompletes

Students who are confronted with compelling circumstances beyond their control which interfere with the ability to complete their semester's work during the normal course of a term may request an incomplete grade from the instructor. This must be requested by the last day of class. Approval of such a request is neither automatic nor guaranteed, but it is expected that faculty will make every effort to accommodate students dealing with illness in the family and other pandemic-related hardships. The instructor and student must establish a timetable for submitting the unfinished work with a final deadline no later than the end of the third week of the subsequent semester. Exceptions to this deadline require a petition from the instructor to the student's academic advising office before this date. When entering an Incomplete grade in SIS, faculty must include a reversion grade which represents the grade the student will receive if they do not complete the missing work by the agreed-upon deadline.

Deadlines for Adding, Dropping and Withdrawing from Courses

Students may add a course up to February 4, 2022. They may drop courses up until March 6, 2022 provided they remain registered for a minimum of 12 credits. Between March 7, 2022 and April 15, 2022 a student may withdraw from a course with a W on their academic record. A record of the course will remain on the academic record with a W appearing in the grade column to indicate that the student registered and then withdrew from the course.

For more information on these and other academic policies, see https://e-catalogue.jhu.edu/engineering/full-time-residential-programs/undergraduate-policies/academic-policies/grading-policies/