CSE 5539-0010 - Spring 2019

Computational Audition

Syllabus

Time/Place: Monday 12:40-2:30pm, 357 Dreese Lab

Instructor: Prof. DeLiang Wang; Office: 598 DL, 292-6827; dwang@cse.ohio-state.edu

Http://www.cse.ohio-state.edu/~dwang/

Office Hours: 2:30-3:00 M & 2-3:30 R

Course Description: Computational models and deep learning algorithms for auditory perception and processing. Topics include auditory scene analysis, pitch analysis, sound localization, speech and music processing (enhancement, segregation, recognition), modeling of the auditory system, and temporal pattern processing.

Prerequisites: Grad standing or permission of the instructor

Suggested Text: "Computational Auditory Scene Analysis," edited by D.L. Wang
and G.J. Brown, Wiley/IEEE Press, 2006

Course Material: Papers from the literature

Grading Plan: Paper presentation - 25%; Midterm programming - 15%
Participation - 10%; Term project - 50%

Tentative Schedule

Week/Sequence and Topics

Week 1:   Class canceled due to travel
Monday, 1/7

Week 2:   Introduction, acoustics, and physiological basis of audition
Monday, 1/14

Week 3:   No Class – MLK Day
Monday, 1/21

Week 4:   Psychoacoustic basis of audition
Monday, 1/28

Week 5:   Fundamentals of signal processing
Monday, 2/4

Week 6:   Real-world audition, and fundamental auditory representations
Monday, 2/11

Student presentation starts

Week 7:   Pitch detection and multipitch tracking
Monday, 2/18   Readings: de Cheveigne (2006), Han and Wang (2014)

Week 8:   Classification based speech separation
Monday, 2/25   Readings: Wang and Wang (2013), Healy et al. (2013)

Week 9:   Masking-based speaker separation
Monday, 3/4   Readings: Healy et al. (2017), Luo et al. (2018)

Spring Break Week: 3/11-15

Week 10:   Reverberant signal processing
Monday, 3/18   Readings: Brown and Palomäki (2006), Hazrati et al. (2013)

Week 11:   Sound localization
Monday, 3/25   Readings: Stern et al. (2006), Wang et al. (2019)

Week 12:   Binaural segregation
Monday, 3/25   Readings: Feng and Jones (2006), Zhang et al. (2017)

Week 13:   Automatic speech recognition (ASR) and robust ASR
Monday, 4/8   Reading: Hinton et al. (2012), Wang and Wang (2016)

Week 14:   Music signal analysis and separation
Monday, 4/15   Readings: Goto (2006), Rafii and Pardo (2013)

Week 15:   Auditory attention and neural modeling
Monday, 4/22   Readings: Shamma et al. (2011), Wrigley and Brown (2004)

Term Project

Each student is required to do a term project (counted as 50% of your final grade). Topics for a term project may be, but not limited to, a complete piece of original research, a proposal and some initial results of a research plan, or a detailed analysis or an in-depth review on a specific topic covered in the class. The term paper summarizing the project must be 2000-4000 (5-10 pages) words in length. **Written project report is due by 5 pm, April 24 (Wednesday).**

Selecting a topic

A topic proposal needs to be submitted prior to conducting the project, and **the proposal is due in class on February 11 (Monday)**. Check with the instructor before deciding on a topic and writing the proposal. Your proposal should be one page long and should include

1. Your background: technical area, relevant coursework, and pertinent research you have done if any.
2. Proposed project
   - General area that the project falls in.
   - Technical description of your proposed project. What are you going to do?
   - List of a few references that you have read or scanned on the subject.

Class Presentation

Each enrolled student (BOTH regular and audit) is required to be responsible for at least one class presentation (counted as 25% of your final grade). The content of each class consists typically of a couple of papers. Please make three choices, in the order of preference, of the topic you want to present. However, your choices may not be honored due to conflict among student choices. Students responsible for a presentation need to check with the instructor beforehand to go over presentation details. It is required that students make powerpoint slides to facilitate presentations.

Course Materials

(Class website: www.cse.ohio-state.edu/~dwang/teaching/cse5539/Wang5539-19.html)

Week 7

de Cheveigne A. (2006): "Multiple F0 estimation," In Wang D.L. & Brown G.J. (eds.): *Computational auditory scene analysis: Principles, algorithms, and applications*. IEEE Press/Wiley, Hoboken NJ, Chapter 2, pp. 45-79.

Han K. and Wang D.L. (2014): "Neural network based pitch tracking in very noisy speech," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, pp. 2158-2168.

Week 8

Wang Y. and Wang D.L. (2013): "Towards scaling up classification-based speech separation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, pp. 1381-1390.

Healy E.W., Yoho S.E., Wang Y., and Wang D.L. (2013): "An algorithm to improve speech recognition in noise for hearing-impaired listeners," *Journal of the Acoustical Society of America*, vol. 134, pp. 3029-3038.

Week 9

Healy E.W., Delfarah M., Vasko J.L., Carter B.L., and Wang D.L. (2017): "An algorithm to increase intelligibility for hearing-impaired listeners in the presence of a competing talker," *Journal of the Acoustical Society of America*, vol. 141, pp. 4230-4239.

Luo Y., Chen Z., Mesgarani N. (2018): "Speaker-independent speech separation with deep attractor network," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, pp. 787-796.

Week 10

Brown G.J. and Palomäki K.J. (2006): "Reverberation," In Wang D.L. & Brown G.J. (eds.): *Computational auditory scene analysis: Principles, algorithms, and applications*. IEEE Press/Wiley, Hoboken NJ, Chapter 7, pp. 209-250.

Hazrati O., Lee J., and Loizou P.C. (2013): "Blind binary masking for reverberation suppression in cochlear implants," *Journal of the Acoustical Society of America*, vol. 133, pp. 1607-1614.

Week 11

Stern R.M., Wang D.L., and Brown G.J. (2006): "Binaural sound localization," In Wang D.L. & Brown G.J. (eds.): *Computational auditory scene analysis: Principles, algorithms, and applications*. IEEE Press/Wiley, Hoboken NJ, Chapter 5, pp. 147-185.

Wang, Z.-Q., Zhang X., and Wang D.L. (2019): "Robust speaker localization guided by deep learning based time-frequency masking," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, pp. 178-188.

Week 12

Feng A.S. and Jones D.L. (2006): "Location-based grouping," In Wang D.L. & Brown G.J. (eds.): *Computational auditory scene analysis: Principles, algorithms, and applications*. IEEE Press/Wiley, Hoboken NJ, Chapter 6, pp. 187-207.
Zhang X. and Wang D.L. (2017): "Deep learning based binaural speech separation in reverberant environments," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, pp. 1075-1084.

Week 13

Hinton G. et al. (2012): "Deep neural networks for acoustic modeling in speech recognition," *IEEE Signal Processing Magazine*, November, pp. 82-97.
Wang Z.-Q. and Wang D.L. (2016): "A joint training framework for robust automatic speech recognition," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, pp. 796-806.

Week 14

Goto M. (2006): "Analysis of musical audio signals," In Wang D.L. & Brown G.J. (eds.): *Computational auditory scene analysis: Principles, algorithms, and applications*. IEEE Press/Wiley, Hoboken NJ, Chapter 8, pp. 251-295.
Rafii Z. and Pardo B. (2013): "REpeating Pattern Extraction Technique (REPET): A simple method for music/voice separation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 21, pp. 71-82.

Week 15

Shamma S., Elhilali M., and Micheyl C. (2011): "Temporal coherence and attention in auditory scene analysis," *Trends in Neuroscience*, vol. 34, pp. 114-123.
Wrigley S. N. and Brown G. J. (2004): "A computational model of auditory selective attention," *IEEE Transactions on Neural Networks*, vol. 15, pp. 1151-1163.

Form for Student Information and Choosing Presentation Topic

Due: January 28 (Monday)

Name: _____

Email Address and Phone No: _____

Department: _____

Thesis advisor (if any): _____

Level: _____-th year Graduate or Undergraduate

Your Areas of Interest: _____

Selecting three topics for possible class presentation in the order of preference, according to the class schedule. Student presentations cover Week 7 through Week 15. Take a look at the titles of the papers to be discussed in each class.

1. Topic _____

    Week _____

2. Topic _____

    Week _____

3. Topic _____

    Week _____