Welcome to

Computer Audition

(ECE 277/477, AME 277/477, CSC 264/464, TEE 477)

Zhiyao Duan
Human Audition

- Understanding the environment
- Communication
- Entertainment
Computer Audition

- Understanding the environment
- Communication
- Entertainment – entertain human
Some Key Problems

- Sound source identification

- Source localization

- Content understanding
  - Speech, event, melody, rhythm

- Source separation
Tools for Sound Interaction

Create: Bone Flutes (7000 B.C.)

Modify: Delphi Theater (300 B.C.)

Record: Cylinder Phonograph (1899)

Transmit: Crystal Radio (1914)
Impact on Many Fields

- Computer Audition
- Psycho-acoustics
- Information Retrieval
- Machine Learning
- Speech Science
- Music Cognition
- Signal Processing
Many Applications

- SoundHound
  Instant Music Search and Discovery

- Siri

- Microsoft Research Songsmith
Some Demos

- Automatic music accompaniment
  - [http://www.music.informatics.indiana.edu/~craphael/music_plus_one/movies/movies.html](http://www.music.informatics.indiana.edu/~craphael/music_plus_one/movies/movies.html)

- Multimedia synchronization
  - [https://www.audiolabs-erlangen.de/fau/professor/mueller/demos](https://www.audiolabs-erlangen.de/fau/professor/mueller/demos)
Some Demos

• Source Separation
  – https://www.youtube.com/watch?v=b07ty1jNcs
  – https://www.youtube.com/watch?v=C8aZYVaFcjE

• Singing-informed source separation
  – http://paris.cs.illinois.edu/demos/ai/user-guide.mp4
Some Demos

- **Soundprism**

  Single-channel polyphonic music

  ![Score](image1)

  Score follower

  Source separator

  Source 1
  Source 2
  ...
  Source N

J. Brahms, Clarinet Quintet in B minor, op.115.
3rd movement

---

ECE 477 - Computer Audition, Zhiyao Duan 2018
Some Demos

- **Automatic music transcription**

```
<table>
<thead>
<tr>
<th>Time (Second)</th>
<th>Algorithm Transcription</th>
<th>Ground-truth Transcription</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

- **Acoustic event detection and localization**

  - [https://www.youtube.com/watch?v=iImkV6oK](https://www.youtube.com/watch?v=iImkV6oK)
Some Demos

• Voice conversion
  – https://www.youtube.com/watch?v=RB7upq8nzIU

• Audio morphing
  – https://www.audiolabs-erlangen.de/resources/MIR/2015-ISMIR-LetItBee
Some Demos

• Automatic song writing
  – http://www.youtube.com/watch?v=3oGFogwc
    x-E

• Music Generation
  – https://www.youtube.com/watch?v=BfrNiqvK
    bLQ
Course Topics

• Fundamentals of human audition
• Auditory models
• Audio features (pitch, timbre, etc.)
• Audio modeling techniques
• State-of-the-art research topics
  – Polyphonic pitch analysis
  – Source separation
  – Sound identification
  – ......
Course Objectives

• General understanding of the field
• Deep understanding and hands-on research experience in a sub-field

• Gain experience of the full cycle of research
• Able to think critically
• Improve presentation and writing skills
Assignments

• Total (110 points)
  – Homework (50 points)
  – Class paper review (14 points)
  – Presentation of research (10 points)
  – Course project (30 points)
  – Peer feedback (6 points)

• No exams
Grading

- No extra credit
- No curve

---

C- C C+ B- B B+ A- A

70 73 77 80 83 87 90 93

110
Important Policies

• No late homework

• Do your own work

• Attendance is not taken, but class discussions are very important for learning
Prerequisites

• Signal Processing
  – ECE 246/446 or ECE 272/472 or equivalent

• Matlab programming

• Preferred but not required
  – Machine learning such as SVM, Markov models, neural networks, clustering, etc.
Three Websites

• Course website
  – All materials (lecture notes, readings, assignments, etc.)
  – http://www.ece.rochester.edu/~zduan/teaching/ece477

• Blackboard:
  – Only for announcements and homework submissions

• Piazza
  – Only for discussions