Speech and Audio Processing: Theory and Applications (3 Credits)

Synopsis
Speech is inarguably the most preferred and natural way of communication for humans. Speech is transmitted from a speaker to a listener in the form of an acoustic signal. The signal carries abundant information, including the linguistic content, the speaker’s voice characteristic, health and emotional conditions, and the ambient environment. Speech signals have many distinctive features that are not found in other signals from the natural world. In the first part of this course, students will study the fundamental theory of digital processing of speech signals. Important time-domain and frequency-domain properties of speech signals will be investigated. Other types of audio signals, namely music and noise, will also be covered in our discussion. The second part of this course will be focused on a few selected applications of speech and audio processing, which include automatic speech recognition, music classification, hearing and speaking aids. The basic principles of system design will be introduced and the major technological challenges will be discussed. Students who take this course are expected to have fundamental knowledge in signals and systems and experience in using MATLAB.

Offering
2015 Julmester

Audience
Year 3 & 4 Undergraduate and Graduate Students

Classroom
Room xxx, Teaching Bldg. No. XX, Peking University

Schedule
Class: 8-11 AM, M-F, July 6–24, 2015; Final Exam: 8-11 AM, July 25, 2015

Objective
- To understand human speech communication from signal processing perspective.
- To gain hands-on experience in analyzing and manipulating different types of audio signals.
- To study existing and potential applications of speech and audio processing.
- To experience the design and implementation process of a computer-based speech and audio processing system.

Topics
1. Digital signal processing [6 hr]: discrete Fourier transform, short-time Fourier transform, digital filters and filter bank.
2. Speech communication [6 hr]: human speech production, human auditory perception, classification of speech sounds.
4. Music analysis [6 hr]: pitch and harmonics, notes, tempo, rhythm, melody and timbre.
5. Selected applications [15 hr]: hearing aids, cochlear implants, automatic speech recognition, music transcription.

References

Grading

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assignments</td>
<td>20%</td>
</tr>
<tr>
<td>Project</td>
<td>30%</td>
</tr>
<tr>
<td>Midterm Exam</td>
<td>10%</td>
</tr>
<tr>
<td>Final Exam</td>
<td>40%</td>
</tr>
</tbody>
</table>

Total 100%