Module Title: 5C4 Speech and Audio Engineering

Code: EE5C04

Level: MAI Year 5 (Optional module)

Credits: 5

Lecturer(s): Associate Professor Naomi Harte

Module Organisation

<table>
<thead>
<tr>
<th>Semester</th>
<th>Start Week</th>
<th>End Week</th>
<th>Associated Project Hours</th>
<th>Lectures</th>
<th>Tutorials and Seminar</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Per week</td>
<td>Total Per week Total</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>12</td>
<td>3</td>
<td>3</td>
<td>33 11</td>
</tr>
</tbody>
</table>

Total Contact Hours: 44

Module Description
Speech is the most important and pervasive form of communication. Speech Engineering requires an understanding of the physiology of the human vocal and auditory systems. This understanding informs the signal processing methods that characterise speech signals. These methods include lossless tube models of speech production; time and frequency domain representations of speech; and window characteristics and time/frequency resolution tradeoffs. Statistical signal processing methods such as autocorrelation and linear prediction of speech provide the basis of methods for speech synthesis and language processing. The module will introduce the role of speech in Human Computer Interaction, HCI. This role requires an understanding of approaches to speech recognition such as hidden Markov models and associated machine learning algorithms and methods for assessment of speech quality and biometric features.

Learning Outcomes
On completion of this module, the student will be able to:

1. Describe, in terms of signal processing, the functioning of the human vocal and auditory systems;
2. Explain the frequency characteristics of speech signals;
3. Explain non-uniform frequency bands and metrics such as mel scales and dBA;
4. Explain the role of binaural hearing in distinguishing the direction of an acoustic source;
5. Apply machine learning techniques to develop a speech processing application in speech recognition, speaker identification or other similar domain;
6. Analyse the function of feature extraction in speech and audio signal processing;
7. Assess and explore the privacy and ethical issues around developing a new speech technology.
Module Syllabus

1. Digital signal processing for speech and audio processing
   - Time and spectral domain filters for speech and audio
   - Window functions and time-frequency analysis
   - Linear and nonlinear operators for feature extraction
   - Pattern classification
2. Speech production, the auditory system and speech perception
   - Introduction to phonetics and phonology
   - Models of speech production
   - Ear physiology, psychoacoustics and speech perception
3. Applications to speech systems
   - Machine learning techniques for speech processing
   - Speech Recognition, speaker recognition, biometrics & forensics
   - Paralinguistics
   - Data privacy and ethics

Recommended Texts

- *Fundamentals of Speech Recognition*, Lawrence Rabiner, B H Juang

Other papers will be given as assigned reading during the course.

Teaching Strategies

This is a learning rather than a teaching based programme. As a Master’s level module it is the expectation that students learn best through a learning based program. That is, as a graduate level module, students learn on their own, with guidance from their professors. So, while teacher centred approaches are used to begin and end a series of lessons on each topic, in-between the focus is on the students with group reads, research and discussions that allow students to pro-actively seek information for themselves.

Assessment

50% of the mark for this subject will come from a number of assignments based on course lectures and reading material. The remaining 50% will be based on a formal 2-hour annual exam.