### SUBJECT DESCRIPTION FORM

<table>
<thead>
<tr>
<th>Subject title:</th>
<th>Speech Processing and Recognition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject code:</td>
<td>EIE558</td>
</tr>
<tr>
<td>Credit value:</td>
<td>3</td>
</tr>
<tr>
<td>Responsible staff and department:</td>
<td>Dr M.W. Mak, EIE</td>
</tr>
<tr>
<td>Pre-requisite:</td>
<td>Nil</td>
</tr>
<tr>
<td>Recommended background knowledge:</td>
<td>Students are expected to have some background knowledge of digital signal processing.</td>
</tr>
<tr>
<td>Mutual exclusions:</td>
<td>Nil</td>
</tr>
</tbody>
</table>

#### Learning approach:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Lecture/Seminar/Tutorial</td>
<td>33 hours</td>
</tr>
<tr>
<td>Laboratory/Demonstration</td>
<td>9 hours</td>
</tr>
</tbody>
</table>

#### Assessment:

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuous Assessment</td>
<td>50%</td>
</tr>
<tr>
<td>Examination</td>
<td>50%</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
</tr>
</tbody>
</table>

#### Objectives:

This subject aims to enable students to master the state-of-the-art theories and technologies behind various speech related products, such as mobile phones, voice over IP, Internet phones, directory services, and speaker recognition systems. The subject focuses on practical issues of speech processing and recognition systems. After completion of the subject, students will have an in-depth understanding of what the current speech technologies can offer and be able to apply speech processing techniques to industrial and commercial applications. The subject is also suitable for students who are preparing to carry out research in speech processing and recognition.
Keyword syllabus:

Part I: Fundamental Concepts

1. Speech Production and Modelling
   1.1 Physiology of speech generation and perception; phonemics and phonetics; acoustic characteristic of speech sounds
   1.2 Acoustic theory of speech production; discrete-time speech production model; discrete-time filter model for speech production

2. Speech Analysis and Feature Parameters
   2.1 Short-term processing; short-term Fourier analysis; spectrograms
   2.2 Speech enhancement
   2.3 Linear prediction; cepstrum; LPCCs; MFCCs

Part II: Advanced Topics and Applications

3. Speech Coding
   3.1 Attributes of speech coders; waveform coding; linear prediction vocoder
   3.2 Linear prediction analysis-by-synthesis coders: CELP, MPLPC, G.729, G.723.1
   3.3 Mixed excitation coding: MELP
   3.4 Speech coding standards

4. Speech Recognition
   4.1 Feature parameters for speech recognition
   4.2 Large-vocabulary continuous speech recognition: hidden Markov models (HMM), acoustic modeling, language modeling, phone models, and decoding
   4.3 Speaker adaptation: MAP, MLLR, and Eigenvoices

5. Speaker Recognition
   5.1 Identification vs. verification; text-dependent vs. text-independent
   5.2 Components of speaker verification systems; features for speaker recognition; speaker models; background models; decision strategy
   5.3 Robust speaker verification: feature and model transformation, feature mapping, feature warping, score normalization, factor analysis, and nuisance attribute projection
   5.4 High-level speaker recognition: prosodic features, lexical features, and phonetic features
   5.5 Discriminative techniques: GMM supervectors, sequence kernels, and SVM speaker verification
   5.6 Evaluation of speaker verification systems; applications of speaker recognition technologies

Indicative reading list and references:


June 2009