Subject Description Form

Subject Code	EIE4105
Subject Title	Multimodal Human Computer Interaction Technology
Credit Value	3
Level	4
Pre-requisite	EIE3312 Linear Systems or EIE3103 Digital Signals and Systems
Co-requisite/ Exclusion	Nil
Objectives	This course aims at providing students with a basic understanding of the theory and applications of multimodal human computer interaction (HCI) technologies.
Intended Subject Learning Outcomes	**Upon completion of the subject, students will be able to:** Category A: Professional/academic knowledge and skills 1. Understand the benefits of using multimodal HCI. 2. Understand the basic theories in statistical learning. 3. Have basic understanding of speech recognition, speaker recognition, handwriting recognition and face recognition. Category B: Attributes for all-roundedness 4. Understand the creative process when designing solutions to a problem.
Subject Synopsis/ Indicative Syllabus	1. Multimodal HCI Inputs and Their Applications Applications of multimodal HCI interfaces in daily life. Advantages of multimodal input interfaces. Understanding multimodal input behaviour. Trends in HCI technologies. 2. Fundamental of Statistical Learning Probability and random variables. Probability densities and distributions. Sampling distributions. Expectations and covariance. Bayes rule and Bayes decision theory. Curse of dimensionality. 3. Statistical Machine Learning for HCI Structure of pattern recognition systems. Unsupervised Learning: principal component analysis, K-means; Gaussian mixture models; hidden Markov models. Supervised Learning: linear regression; linear discriminant analysis; support vector machines. Deep Learning: deep neural networks (DNN); restricted Boltzmann machines; backpropagation. Applications to handwriting recognition. Applications to face recognition. Eigenface. Fisherface. 4. Speech Recognition and Its Applications Acoustic feature extraction. HMM for acoustic modelling. DNN for acoustic modelling. Language modelling. Applications: voice search, voice conversion, spoken dialog, speech emotion recognition 5. Speaker Recognition and Its Applications Acoustic features for speaker recognition. GMM-UBM systems. Factor analysis and I-vectors. Probabilistic linear discriminant analysis (PLDA). DNN for speaker recognition. Applications: biometric authentication; forensic.

Teaching/Learning Methodology

Lectures: The subject matters will be delivered through lectures. Students will be engaged in the lectures through Q&A, discussions and specially designed classroom activities.

Tutorials: During tutorials, students will work on/discuss some chosen topics. This will help strengthen the knowledge taught in lectures.

Laboratory and assignments: During laboratory exercises, students will perform hands-on tasks to practice what they have learned. They will evaluate performance of systems and design solutions to problems. The assignments will help students to review the knowledge taught in class.

While lectures and tutorials will help to achieve the professional outcomes, the open-ended questions in laboratory exercises and assignments will provide the chance to students to exercise their creatively in problem solving.

Assessment Methods in Alignment with Intended Subject Learning Outcomes

Specific Assessment Methods/Tasks	**% Weighting**	**Intended Subject Learning Outcomes to be Assessed (Please tick as appropriate)**			
		1	**2**	**3**	**4**
1. Continuous Assessment (total: 50%)					
• Homework and assignments	15%	✓	✓	✓	✓
• Tests and Quizzes	20%	✓	✓	✓	
• Laboratory exercises	15%			✓	✓
2. Examination	50%	✓	✓	✓	✓
Total	100%				

Explanation of the appropriateness of the assessment methods in assessing the intended learning outcomes:

Assignment, homework and laboratory exercises will require students to apply what they have learnt to solve problems. There will be open-ended questions that allow students to exercise their creativity in making design.

Examination and tests: They assess students' achievement of the learning outcomes in a more formal manner.

Student Study Effort Expected

Class contact (time-tabled):	
• Lecture	24 Hours
• Tutorial/Laboratory/Practice Classes	15 Hours
Other student study effort:	
• Lecture: preview/review of notes; homework/assignment; preparation for test/quizzes/examination	36 Hours
• Tutorial/Laboratory/Practice Classes: preview of materials, revision and/or reports writing	30 Hours
Total student study effort:	**105 Hours**

Reading List and References	**Reference Materials:** 1. S.Y. Kung, M.W. Mak and S.H. Lin, *Biometric Authentication: A Machine Learning Approach*, Prentice Hall, 2005. 2. Spoken Language Technology, *IEEE Signal Processing Magazine*, vol. 25, No. 3, May 2008. 3. C.M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006. 4. J.P. Thiran, F. Marques and H. Bourlard, *Multimodal Signal Processing, Theory and Applications for Human Computer Interaction*, Elsevier, 2010. 5. S.B. Wan and M.W. Mak, *Machine Learning for Protein Subcellular Localization Prediction*, De Gruyter, 2015. 6. M.W. Mak, "Lecture Notes on Factor Analysis and I-Vectors", *Technical Report and Lecture Note Series, Department of Electronic and Information Engineering, The Hong Kong Polytechnic University*, Feb. 2016. 7. M.W. Mak, "Lecture Notes on Backpropagation", *Technical Report and Lecture Note Series, Department of Electronic and Information Engineering, The Hong Kong Polytechnic University*, July 2015.
Last Updated	March 2016
Prepared by	Dr M.W. Mak