Course Name: Speech Processing
Course Code: IT443

I. Basic Course Information

Major or minor element of program: Major
Department offering the course: Information Technology Department

Academic level: 400 Level
Semester in which course is offered: Second (spring) Semester
Course pre-requisite(s): Signals and Systems (IT241)

Credit Hours: 3
Contact Hours Through:

Lecture	Tutorial*	Practical*	Total
2.5	0.0	1.5	4.0

* 1.5 hours for **either** Tutorial or Practical

Approval date of course specification: January 2015

II. Overall Aims of Course

Course aims to reinforce concepts learned in prerequisite courses, to introduce new tools needed to deal with time-varying signals and to have students apply what they have learned to their own voices. Speech processing methodologies will be covered in lectures, computer-lab sessions and practical lab work. A semester project is a large component of this course. The course's major target is to provide students with the knowledge of basic characteristics of speech signal in relation to production and hearing of speech by humans, describe basic algorithms of speech analysis common to many applications, give an overview of applications (recognition, coding) and to inform about practical aspects of speech algorithms implementation.

III. Program ILOs covered by course

Program Intended Learning Outcomes (By Code)			
Knowledge & Understanding	Intellectual Skills	Professional Skills	General Skills
K1,K17,K20	I2,I14,I18,I19	P12,P13,P14,P15	G2,G7,G9

IV. Intended Learning Outcomes of Course (ILOs)

a. Knowledge and Understanding

On completing the course, students should be able to:

- K.1 Recall the basic concepts of digital speech processing.
- K.2 Define concepts of phonemics and phonetics.
- K.3 Recognise the human speech production and auditory sytems.
- K.4 Discuss and explain the different short-term processing methodologies of speech.
- K.5 Explain the difference between different sound classes.

b. Intellectual/Cognitive Skills

On completing the course, students should be able to:

- I.1 Utilize the normal intellectual skills developed in any study of signals and systems at this level.
- I.2 Assess different speech processing algorithms based on their outcomes during practical application.
- I.3 Apply native digital signal processing methodologies to solve speech processing problems.
- I.4 Justify unexpected results obtained through applications.
- I.5 Choose the appropriate signal processing algorithm based on the problem presented.

c. Practical/Professional Skills

On completing the course, students should be able to:

- P.1 Practice applying acquired knowledge of signals and systems to the specific problems that arise in speech communication.
- P.2 Use the Matlab tool to build different speech processing algorithms using Matlab's Signal Processing Tool Box.
- P.3 Integrate all algorithms implemented into a single working GUI framework using the Matlab tool.
- P.4 Use the HTK toolbox to build a digit recogniser as the semester's project.
- P.5 Develop a speech processing system as real world application.

d. General and Transferable Skills

On completing the course, students should be able to:

- G.1 Demonstrate ability in time management, organization skills, communication skills, report writing skills, and presentation skills for a variety of audiences (e.g., management, technical, academic).
- G.2 Demonstrate ability to work as a team member.
- G.3 Show the ability to efficiently use IT resources and general computing facilities.
- G.4 Demonstrate independent critical thinking and problem solving skills.

V. Course MatrixContents

	Main Topics / Chapters	Duration (Weeks)	Course ILOs Covered by Topic (By ILO Code)			
			K & U	I.S.	P.S.	G.S.
1-	Introduction to Speech Processing and Related Technologies. • Interaction with speech files	0.5	K1		P1	G2,G3
2-	Fundamentals of DSP (Revision): z-Transform, Fourier Transform, Digital Filters, Sampling Theorem. • Speech Spectrogram	0.5	K1	I2,I4	P1, P2	G2,G3
3-	Fundamentals of Speech Science: Speech Production Mechanism, Sound Units, Acoustic Theory, Digital Modeling • Energy and zero crossing computation	1.5	K1,K2	I1,I2,I3,I4	P1,P4	G2,G3
4-	Time domain analysis of speech signal: Short-time analysis, frame, short time energy, short time zero-crossing, short time average magnitude, short time auto-correlation, silence removing, pitch detection, voice/voiceless classification. • Classification of voiced and unvoiced speech segments	1.5	K4	I1,I2,I3,I4	P2,P4	G2,G3
5-	Frequency domain analysis of speech signal: short Fourier transform, Short-time Spectrum • End point detection	1	K1,K2, K3	I1,I2,I3,I4	P1,P4	G2,G3
6-	Linear Prediction Analysis: what is it good for ?, Prediction of a sample from past samples, linear prediction (LP), Error of LP, Determination of vocal tract characteristics using LP analysis, Spectrum estimated by LP. Features derived from LP. • Pitch detection	1.5	K4	I1,I2,I3,I4	P1,P2	G1,G2,G3
7-	Cepstrum analysis: what is it good for? Basic idea of homomorphic analysis, type of cepstrum analysis, complex cepstrum, real cepstrum, Features derived from cepstrum, cepstrum derived from LP. • Linear predictive coefficients	1	K4	I1,I2,I3,I4, I5	P1,P2	G2,G3
8-	Coding: Aims of coding. Bit-rate, objective and subjective	1.5	K2,K4	I1,I2,I3,I4	P3,P4	G1,G2,G3,G4

	measurements of quality, Classification of coders according to bit-rate, Waveform Coding, LP Coding, CELP Coding, Vector Quantization . • Building a digit recognizer based on HTK]					
9-	[Introduction to speech recognition: the task, classification of recognizers: isolated words, connected words, continuous speech, speaker dependent, and speaker independent, Basic function blocks, Feature Extraction, Template Matching, Statistical Modeling, Design of Recognition Systems Recognition using DTW, Recognition based on distance of speech frames, various definitions of distance, Timing: linear modification, dynamic programming (Dynamic Time Warping DTW) .]	[2.5]	[K1,K4, K5]	[I1,I2,I3,I4, I5]	[ ]	[ ]
10-	[Hidden Markov models: Introduction, motivations and relation to DTW, Structure of the model, type of the HMM, solutions of the three HMM problems: training, scoring and decoding problems.]	[1.5]	[K1,K4]	[I1,I2,I3,I4, I5]	[P5]	[G1,G4]
	Net Teaching Weeks	**13**				

VI. Course Weekly Detailed Topics / hours / ILOs

Week No.	**Sub-Topics**	**Total Hours**	**Contact Hours**	
			Theoretical Hours	Practical Hours*
1	[Introduction to Speech Processing and Related Technologies.]	[2.5]	[2.5]	
2	[Fundamentals of DSP (Revision): z-Transform, Fourier Transform, Digital Filters, Sampling Theorem.]	[4]	[2.5]	[4]
3	[Fundamentals of Speech Science: Speech Production Mechanism, Sound Units, Acoustic Theory, Digital Modeling]	[4]	[2.5]	[4]
4	[Time domain analysis of speech signal: Short-time analysis, frame, short time energy, short time zero-crossing, short time average magnitude, short time auto-correlation, silence removing, pitch	[4]	[2.5]	[4]

	detection, voice/voiceless classification.			
5	Frequency domain analysis of speech signal: short Fourier transform, Short-time Spectrum	4	2.5	4
6	Linear Prediction Analysis: what is it good for ?, Prediction of a sample from past samples, linear prediction (LP), Error of LP, Determination of vocal tract characteristics using LP analysis, Spectrum estimated by LP. Features derived from LP.	4	2.5	4
7	**Midterm Exam**			
8	Cepstrum analysis: what is it good for? Basic idea of homomorphic analysis, type of cepstrum analysis,	4	2.5	4
9	Cepstrum analysis: Complex cepstrum, real cepstrum, Features derived from cepstrum, cepstrum derived from LP.	4	2.5	4
10	Coding: Aims of coding. Bit-rate, objective and subjective measurements of quality.	4	2.5	4
11	Coding: Classification of coders according to bit-rate, Waveform Coding, LP Coding, CELP Coding, Vector Quantization.	4	2.5	4
12	Introduction to speech recognition: the task, classification of recognizers: isolated words, connected words, continuous speech, speaker dependent, and speaker independent, Basic function blocks, Feature Extraction, Template Matching, Statistical Modelling.	4	2.5	4
13	Design of Recognition Systems Recognition using DTW, Recognition based on distance of speech frames, various definitions of distance, Timing: linear modification, dynamic programming (Dynamic Time Warping DTW).	4	2.5	4
14	Hidden Markov models: Introduction, motivations and relation to DTW, Structure of the model, type of the HMM, solutions of the three HMM problems: training, scoring and decoding problems.	4	2.5	4
15	**Final Exam**			
Total Teaching Hours		**51**	**33**	**18**

* **No Practical/Tutorial during the first week of the semester**

VII. Teaching and Learning Methods

Teaching/Learning Method	Selected Method	Course ILOs Covered by Method (By ILO Code)			
		K & U	Intellectual Skills	Professional Skills	General Skills
Lectures & Seminars	[]	K1:K5	I1,I2,I4	[]	G3,G4
Tutorials	[]	[]	[]	[]	[]
Computer lab Sessions	[]	K1:K5	I1,I3,I5	P1,P2	G1,G3,G4
Practical lab Work	[]	K3,K4,K5	I1,I3,I5	P1,P2,P4,P5	G1,G2,G3,G4
Reading Materials	[]	K1:K5	[]	P2	G2
Web-site Searches	[]	K1:K5	[]	P2	G2,G3
Research & Reporting	[]	[]	[]	[]	[]
Problem Solving / Problem-based Learning	[]	K1:K5	I1,I2,I3,I4,I5	P1:P5	G2,G4
Projects	[]	[]	[]	[]	[]
Independent Work	[]	[]	[]	[]	[]
Group Work	[]	K1:K5	I2,I3,I4,I5	P1:P5	G2,G4
Case Studies	[]	[]	[]	[]	[]
Presentations	[]	K1,K2	I2,I4	P4	G1,G2,G3,G4
Simulation Analysis	[]	[]	[]	[]	[]
Others (Specify):	[]	[]	[]	[]	[]

VIII. Assessment Methods, Schedule and Grade Distribution

Assessment Method	Selected Method	Course ILOs Covered by Method (By ILO Code)				Assessment Weight / Percentage	Week No.
		K & U	I.S.	P.S.	G.S.		
Midterm Exam	X	K1:K4	[]	[]	[]	15 %	7
Final Exam	X	K1:K5	[]	[]	[]	60%	15
Quizzes	[]	[]	[]	[]	[]	[]	[]
Course Work	[]	[]	[]	[]	[]	[]	[]
Report Writing	[]	[]	[]	[]	[]	[]	[]
Case Study Analysis	[]	[]	[]	[]	[]	[]	[]
Oral Presentations	X	K1,K2,K3	I2	P4	G1,G2,G3,G4	5 %	[]
Practical	X	K1,K2,K3	I1	P1:P5	G2,G4	15 %	[]
Group Project	X	K1:K5	I2,I3	P1:P5	G2,G3,G4	5 %	[]
Individual Project	[]	[]	[]	[]	[]	[]	[]
Others (Specify):	[]	[]	[]	[]	[]	[]	[]

IX. List of References

Essential Text Books	• Rabiner, L. Juang, B.H.: Fundamentals of speech recognition, • Signal Processing, Prentice Hall, Engelwood Cliffs, NJ, 1993 • Rabiner, L.R., Schaeffer, L.W.: Digital processing of speechsignals, Prentice Hall, 1978
Course notes	• They are distributed during the course progress.
Recommended books	• None
Periodicals, Web sites, etc.…	• None

X. Facilities required for teaching and learning

- Matlab tool
- C#.NET
- Java
- HTK toolbox
- Recording Devices.

Course coordinator: Prof. Mahmoud Shoman

Head of Department: Prof. Hesham El Mahdy

Date: January 2015