Sardar Patel Institute of Technology

Bhavan's Campus, Munshi Nagar, Andheri (West), Mumbai-400058-India
(Autonomous Institute Affiliated to University of Mumbai)

Course Code	Course Name	Teaching Scheme (Hrs/week)			Credits Assigned			
		L	**T**	**P**	**L**	**T**	**P**	**Total**
ETE801	**Speech Processing**	4	--	--	4	--	--	4
		Examination Scheme						
		ISE	**MSE**	**ESE**				
		10	30	100 (60% Weightage)				

Pre-requisite Course Codes	ETC405 Signals and Systems ETC602 Discrete Time Signal Processing	
After successful completion of the course, student will be able to		
Course Outcomes	CO1	Demonstrate basic knowledge in speech processing production mechanism, phoneme classification, digital models for speech production, Homomorphic speech processing and LPC analysis.
	CO2	Demonstrate applications of signal processing theory for estimation of speech parameters in time and frequency domain including pitch and formants.
	CO3	Analyze application of speech processing in speech compression, speech recognition and speech synthesis.
	CO4	Enhance their written and oral technical communication skills related to speech processing subjects and will better prepared for higher study and life long learning.

Module No.	Unit No.	Topics	Ref.	Hrs.
1	**Speech Production, Acoustic Phonetics and Auditory Perception**		2,3	10
	1.1	Anatomy and physiology of speech organs, articulatory phonetics, acoustic phonetics, acoustic theory of speech production, discrete time model for speech production		
	1.2	Ear physiology and psychoacoustics		
2	**Speech Analysis in Time Domain**		1,2,3	06
	2.1	Time energy, average magnitude, and zero-crossing rate, speech vs silence discrimination		
	2.2	Short-time autocorrelation, pitch period estimation using short-time autocorrelation, median smoothing		
3	**Speech Analysis in Frequency Domain:**		4,5	06
	3.1	Time dependent Fourier representation for voiced and unvoiced speech signals, linear filtering interpretation, spectrographic displays		
	3.2	Pitch period estimation based on FFT and harmonic peak detection method, estimation of formants using log spectrum		
4	**Homomorphic Speech Processing**		1,2	08

	4.1	Cepstral analysis of speech, mel frequency cepstral coefficients (MFCC), perceptual linear prediction (PLP)		
	4.2	Pitch period estimation in cepstral domain, evaluation of formants using cepstrum		
5	**LPC and Parametric Speech Coding**		3,4,5	12
	5.1	Review of lattice structure realization, forward and backward error filters, normal equations & its solutions, levinson-durbin algorithm, covariance method, Berg's algorithm		
	5.2	Channel Vocoders, linear prediction (LP) based vocoders, residual excited LP (RELP) based Vocoders, voice Excited LP (VELP) based vocoders, multi-pulse LP (MPLP) based vocoders, code excited LP (CELP) based vocoders		
6	**Speech Processing Applications**		2,3,5	10
	6.1	Speech recognition systems, deterministic sequence recognition for ASR, statistical sequence recognition for ASR (Hidden Markov Model (HMM))		
	6.2	Text to speech system (TTS), concatenative synthesis, synthesis using formants, LPC synthesizer		
			Total	**52**

References:

1. Rabiner and Schafer, ―*Digital Processing of Speech Signals*‖ , Pearson Education, Delhi, 2004.
2. Shaila D. Apte, ―*Speech and Audio Processing*‖ , Wiley India, New Delhi, 2012.
3. Douglas O'Shaughnessy, ―*Speech Communications: Human & Machine*‖ , Universities Press, Hyderabad, Second Edition, 2001.
4. Ben Gold and Nelson Morgan, ―*Speech and Audio Signal Processing*‖ , Wiley India (P) Ltd, New Delhi, 2006.
5. Thomas F. Quatieri, ―*Discrete-Time Speech Signal Processing: Principles and Practice*‖ , Prentice Hall, 2001.
6. J. L. Flanagan, ―*Speech Analysis Synthesis and Perception*‖ , Second edition, Springer-Verlag (1972).