GUJARAT TECHNOLOGICAL UNIVERSITY

Master of Engineering
Subject Code: 3720516
SPEECH SIGNAL PROCESSING
SEMESTER: II

Type of course: PE-III

Prerequisite: Digital signal processing

Rationale: For humans, speech is a natural way of communicating the ideas. This course is a fundamental course on how to process digital speech signal to extract useful information. The course builds upon the theory of digital signal processing and extends the concepts applied to speech signal in particular. The course also discusses the applications of speech signal processing.

Teaching and Examination Scheme:

Teaching Scheme			Credits	Examination Marks				Total Marks
L	T	P	C	Theory Marks		Practical Marks		
				ESE (E)	PA (M)	ESE (V)	PA (I)	
3	0	2	4	70	30	30	20	150

Sr. No.	Content	Total Hrs
1	Speech Communication: Introduction, discrete-time speech signal processing, speech communication, review of signals and linear Systems	4
2	Speech Production and acoustic phonetics: Anatomy and physiology of speech organs, speech sounds and classification, International Phonetic Alphabet (IPA), Articulatory Phonetics: Manner of articulation and place of articulation, vowel triangle, Acoustic Phonetics: spectrograms, wide-band and narrow-band spectrograms, acoustic characteristics of speech sounds, coarticulation and prosody	6
3	Time-domain models for speech processing: Introduction to short-time speech analysis, windowing, short-time energy and average magnitude, short-time Zero-Crossing Rate (ZCR), speech vs. silence discrimination using energy and zero crossings, short-time autocorrelation function, short-time Average Magnitude Difference Function (AMDF)	8
4	Short-time Fourier analysis: Short-time Fourier transform (STFT), spectral displays, time-frequency resolution tradeoffs, Linear filtering interpretation, short-time synthesis, filter bank summation method	8
5	Linear Predictive Analysis: Basic principles of Linear predictive analysis, autocorrelation method and covariance method, computation of gain for the model, prediction error signal, frequency domain interpretation of LP analysis, frequency domain interpretation of mean squared prediction error, applications of LPC parameters	8
6	Homomorphic Signal Processing: Concept of Homomorphic processing, Homomorphic systems for convolution, properties of complex cepstrum, Homomorphic filtering, complex cepstrum of voiced speech, complex cepstrum of unvoiced speech, Mel-scale cepstrum	8
7	Speech Coding: Fundamentals of coding, liner prediction and harmonic noise models in speech coding, modeling excitation for voiced and unvoiced speech, Code-Excited linear prediction coding	6

	Total	48

Suggested Specification table with Marks (Theory):

Distribution of Theory Marks					
R Level	U Level	A Level	N Level	E Level	C Level
10	**30**	**20**	**20**	**10**	**10**

Legends: R: Remembrance; U: Understanding; A: Application, N: Analyze and E: Evaluate C: Create and above Levels (Revised Bloom's Taxonomy)

Note: This specification table shall be treated as a general guideline for students and teachers. The actual distribution of marks in the question paper may vary slightly from above table.

Reference Books:

1. Speech Communication: Human and machine, D. O'Shaughnessy , Uniiversity Press
2. Digital Processing of Speech Signals, L. Rabiner and R. Schafer, Pearson Education
3. Discrete-time Speech Signal Processing, T. Quatieri, Pearson Education

Course Outcomes☹COs)

A student who successfully completes this course should be able to:

Sr. No.	**CO statement**	**Marks % weightage**
CO-1	To apply basic principles of problem solving in speech signal processing for the society and environment	20
CO-2	To analyze and design speech signal processing systems using acoustic phonetics,time domain methods, short time fourier analysis, linear predictive analysis, homomorphic signal processing for the society and environment in ethical way	30
CO-3	To prepare post graduates with the knowledge, ethics and skills so that they can be applied to various speech processing applications in environment friendly manner for the society	20
CO-4	To build projects individually or in a group consisting of speech processing system as per the need of the society in a professional ethical and environment friendly manner	20
CO-5	To apply the knowledge of speech processing to troubleshoot the speech related products in ethical way and constructively useful for the society and environment	10

List of Experiments:

1. To study the effects of windowing.
2. To understand the difference between stationary and non-stationary signals.
3. To extract a slice of speech signal and compute its spectrum for different window length.
4. To simulate periodic glottal pulse train.
5. To synthesize vowel using source filter model.
6. To compute wideband and narrowband spectrogram of a given speech signal.
7. To compute short-time energy and ZCR of a given speech signal.
8 To compute short-time autocorrelation function and plot pitch contour for given utterance.

9. To compute short-time AMDF and plot pitch contour for given utterance.
10. To detect pitch using harmonic product spectrum.
11. To study LPC and cepstral analysis method.

1. An important pre-processing step in many speech processing tasks is to discriminate between a speech and a silence. In this problem you should come up with an algorithm to segment a given speech signal into speech and silence parts.
2. For a given speech signal, classify speech segments into two parts: voiced and unvoiced speech segments
3. Given a speech signal, determine whether it contains an adult voice or a child voice.
4. Determine pitch of a given speech signal.
5. Determine the locations of vowels in the given speech signal.

List of Open Source Software/learning website:

- Scilab
- http://www.vlab.co.in/ (Virtual labs at IIT Guwahati)
- NTPEL
- Signal Processing Toolbox
- Praat: doing phonetics by computer (version 5.4.01)