GUJARAT TECHNOLOGICAL UNIVERSITY

SPEECH SIGNAL PROCESSING
SUBJECT CODE: 2720504
SEMESTER: II

Type of course: Fundamentals with practical applications

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)		
						ESE	OEP	PA	RP	
3	2#	2	5	70	30	20	10	10	10	150

Content:

Sr. No.	Content	Total Hrs	% weightage
1	Speech Communication: Introduction, discrete-time speech signal processing, speech communication, review of signals and linear systems	04	08
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	06	12
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),	08	17
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	08	17
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	08	17

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	08	17
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	06	12

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:

On successful completion of the course, the students should be able to:

- Understand the process of human speech production and its relation to basic speech signal characteristics
- Get familiar with time- and frequency-domain methods of speech signal processing
- Understand basic algorithms applied to many common applications of speech processing
- Extract speech features for applications like voice activity detector, pitch detection etc.

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.

Open Ended Problems:

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.

Major equipments and software:

MATLAB signal processing toolbox
Praat: doing phonetics by computer (version 5.4.01)

List of Open Source Software/learning website:

Sakshat Virtual Labs, IIT Guwahati

Review Presentation (RP): The concerned faculty member shall provide the list of peer reviewed Journals and Tier-I and Tier-II Conferences relating to the subject (or relating to the area of thesis for seminar) to the students in the beginning of the semester. The same list will be uploaded on GTU website during the first two weeks of the start of the semester. Every student or a group of students shall critically study 2 papers, integrate the details and make presentation in the last two weeks of the semester. The GTU marks entry portal will allow entry of marks only after uploading of the best 3 presentations. A unique id number will be generated only after uploading the presentations. Thereafter the entry of marks will be allowed. The best 3 presentations of each college will be uploaded on GTU website