Speech and Audio signal Processing (404211)

Teaching Scheme:
Lectures: 3 Hrs/ Week

Examination Scheme:
In Semester Assessment:
Phase I : 30
End Semester Examination:
Phase II: 70

Course Objectives:

- To understand the basic concepts of speech processing techniques and its applications.
- To teach about practical aspects of speech processing algorithms implementation.
- To provide good practical knowledge and signal processing concepts applied to various applications of speech.

Course Outcomes:

After successfully completing the course students will be able to

- Implement speech processing algorithms/techniques for speech signal analysis and various speech applications.
- Acquire skills to work in different domain of speech processing like ASR, Speech synthesis, speech enhancement, speech coding etc.

Unit I : **Speech Production and Acoustic Phonetics** 6L

Introduction, Anatomy and Physiology of speech organs, Articulatory phonetic, Acoustic Phonetics, Acoustic theory of speech production, LTI and LTV model, coarticulation, prosody.

Unit II : **Time and Frequency domain processing of speech** 8L

Time-domain processing of speech: Short-time energy, pitch estimation using autocorrelation and AMDF, formant estimation, voiced/unvoiced classification.
Frequency domain analysis: Short time analysis of speech, narrow and broad band spectrogram, cepstral domain analysis, mfcc, Homomorphic processing of speech signal, pitch detection and formant extraction.

Unit III : **Coding of speech signals** 6L

Introduction: Quantization: Quantization error, SNR, Non-uniform quantization, Measures to evaluate speech quality, Time domain waveform coding, spectral coders, Vocoders: Phase, channel, homomorphic, vector quantization coders.

Unit IV : **Linear Predictive coding of Speech** 6L

Basic principles of linear predictive analysis: Autocorrelation and covariance method, Solution

of LPC equation: Cholesky decomposition solution for covariance method and Durban's recursive solution for Autocorrelation equations. Applications of LPC parameters: Pitch detection, Formant analysis LPC Vocoders, voiced excited LPC Vocoders

Unit V : **Speech enhancement** 6L

Introduction, Nature of interfering signals, Speech enhancement techniques: spectral subtraction and filtering, harmonic filtering parametric resynthesis, filtering and adaptive noise cancellation

Unit VI : **Applications of speech and audio signal processing** 6L

Speech recognition, speech synthesis, speaker recognition and verification: Basic principles, specific features and state of the art systems. Fundamentals of Template matching, Pattern classification, statistical methods like DTW, GMM, HMM.

Text Books

3. L. R. Rabiner and R. W. Schaffer, "Digital Processing of Speech signals", Pearson
4. Thomas F. Quatieri, "Discrete-time Speech Signal Processing: Principles and Practice", Pearson Education Asia, 2003.
5. S.D. Apte, "Speech and Audio processing", Wiley India

Reference Books

1. Ben Gold and Nelson Morgan, "Speech and Audio Signal Processing: Processing and Perception of Speech and Music", John Wiley, 2002.
2. Douglas O'Shaughnessy, " Speech communication: Human and Machine, Universities Press

List of experiments:

1. Study of frame format for a .wav file. Record audio signal save as "wav" file format (16 bit, 16 kHz mono format using wave recorder interface (Audacity, Cool Edit Pro etc.). Write a program to read wave file in frame by frame manner and plot the speech segment.
2. Acoustic study of speech sounds using PRATT tool: Record and analyze speech sounds like consonants, vowels, semivowels, diphthongs, nasals, fricatives etc.
3. Classification and voiced and unvoiced part of signal using short time energy and zero crossing rate.
4. Pitch detection using Average Magnitude Difference Function method. Compare pitch with PRAAT pitch contour.
5. Pitch detection using Autocorrelation method. Compare pitch with PRAAT pitch contour.

6. Write a program to compute narrow band and wide band spectrogram and plot 2D and 3D spectrogram.
7. Write a program to draw a Cepstrum of speech segment from the speech utterance.
8. Write a program to find MFCC for a speech segment from the speech utterance.
9. Write a program to find LPC for a speech segment from the speech utterance. Use Levinson Durbin algorithm.
10. Write a program to find first 4 Formants for a speech segment from the speech utterance using a Cepstral domain window.

Group activity: Max. 3 to 4 students per group.

Application assignment: Implement small speech processing related application i.e. digit recognition, voice operated telephone dialing system, speech coding: Vocoders, spectral subtraction, vowel synthesis (using source filter model) etc.