Speech and Audio Signal Processing(404190)

Teaching Scheme:
Lectures: 3 Hrs/ Week

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

Course Objectives:

- To understand basic concepts and methodologies for the analysis and modeling of speech signal.
- To characterize the speech signal as generated by a speech production model
- To understand the mechanism of speech and audio perception
- To understand the motivation of short-term analysis of speech and audio
- To perform the analysis of speech signal using LPC
- To extract the information of the speech or audio signals in terms of cepstral features
- To provide a foundation for developing applications in this field.

Course Outcomes:

After successfully completing the course students will be able to

- Design and implement algorithms for processing speech and audio signals considering the properties of acoustic signals and human hearing.
  Analyze speech signal to extract the characteristic of vocal tract (formants) and vocal cords (pitch).
  Write a program for extracting LPC Parameters using Levinson Durbin algorithm
  Formulate and design a system for speech recognition and speaker recognition

Unit I: Fundamentals of speech productio 6L

Anatomy and physiology of speech production, Human speech production mechanism, LTI model for speech production, Nature of speech signal, linear time varying model, articulatory phonetics, acoustic phonetics, Voiced and Unvoiced speech.

Unit II : Human auditory system 6L

Human auditory system, simplified model of cochlea. Sound pressure level and loudness. Sound intensity and Decibel sound levels. Concept of critical band and introduction to auditory system as a filter bank, Uniform, non uniform filter bank, mel scale and bark scale. Speech perception: vowel perception.

Unit III : Time and frequency domain methods for audio processing 8L

Time-dependent speech processing. Short-time energy, short time average magnitude, Short-time average zero crossing rate. Speech Vs. silence discrimination using energy and zero crossing rate. Short-time autocorrelation function, short-time average magnitude difference function. Pitch period estimation using autocorrelation method. Audio feature extraction, Spectral centroid, spectral spread, spectral entropy, spectral flux, spectral roll-off. Spectrogram: narrow band and wide band spectrogram.

Unit IV : Linear prediction analysis 6L

Basic principles of linear predictive analysis. Autocorrelation method, covariance method. Solution of LPC equations: Cholesky decomposition, Durbin's recursive solution, lattice formulations and solutions. Frequency domain interpretation of LP analysis. Applications of LPC parameters as pitch detection and formant analysis.

Unit V : Cepstral Analysis 6L

Homomorphic speech processing, Real Cestrum: Long-term real cepstrum, short-term real cepstrum, pitch estimation, format estimation, Mel cepstrum. Complex cepstrum: Long-term complex cepstrum, short-term complex cepstrum.

Unit VI : Speech and Audio processing applications 6L

Speech recognition: complete system for an isolated word recognition with vector quantization /DTW. Speaker recognition: Complete system for speaker identification, verification. Introduction to speech enhancement, Speech enhancement using spectral subtraction method, Introduction to Text to speech conversion, Introduction to Musical instrument classification, Musical Information retrieval.

Text Books :

1. Deller J. R. Proakis J. G. and Hanson J. H., "Discrete Time Processing of Speech Signals", Wiley Interscience
2. Ben Gold and Nelson Morgan, "Speech and audio signal processing" Wiley

Reference Books :

1. L. R. Rabiner and S.W. Schafer, "Digital processing of speech signals" Pearson Education.
2. Thomas F. Quateri , "Discrete-Time Speech Signal Processing: Principles and Practice" Pearson
3. Dr. Shaila Apte, "Speech and audio processing", Wiley India Publication
4. L. R. Rabiner and B. H. Juang, "Fundamentals of speech recognition"
5. Theodoros Giannakopoulos and Aggelos pikrakis, " Introduction to audio analysis : A MATLAB Approach : Eleseiver Publication.

List of Experiments

NOTE: To perform the experiments softwares like MATLAB, SCILAB etc. can be used. For analysis of speech signals tools like PRAAT, Audacity can be used. Free source software's are encouraged.

1. Record speech signal and find Energy and ZCR for different frame rates and comment on the result.
2. Record different vowels as /a/, /e/, /i/, /o/ etc. and extract the pitch as well as first three formant frequencies. Perform similar analysis for different types of unvoiced sounds and comment on the result.
3. Write a program to identify voiced, unvoiced and silence regions of the speech signal.
4. Record a speech signal and perform the spectrographic analysis of the signal using wideband and narrowband spectrogram. Comment on narrowband and wide band spectrogram.
5. Write a program for extracting pitch period for a voiced part of the speech signal using autocorrelation .
6. Write a program to design a Mel filter bank and using this filter bank write a program to extract MFCC features.
7. Write a program to perform the cepstral analysis of speech signal and detect the pitch from the voiced part using cepstrum analysis.
8. Write a program to find LPC coefficients using Levinson Durbin algorithm.
9. Write a program to enhance the noisy speech signal using spectral subtraction method.
10. Write a program to extract frequency domain audio features like SC, SF and Spectral roll off.