Approval: 10th Senate Meeting

Course Number: CS609

Course Name: Speech Processing

Credits: 3-0-2-4

Prerequisites: Signals and Systems (IC260), or Digital Signal Processing (EE305) or COT

Students intended for: BTech./MTech./M.S./Ph.D.

Elective or Core: Elective

Semester: Odd

Preamble:
The objective of this course is to introduce theory, techniques and algorithms for the processing of speech signals. Speech processing has many applications which include human-computer interaction (speech-based interfaces including speech-to-text and text-to-speech), storage and transmission of speech and enhancing noisy or degraded speech. An exposure to this field will introduce to the student to several applications of digital signal processing and pattern recognition, in real-world scenarios.

Course Outline:
3 lectures per week, 2 hours for lab assignments and evaluation

Course Modules :

1. **Introduction (1 lecture)**

2. **Review of digital signal processing: (3 lectures)**
   Discrete-time signals and systems, transform representation of signals and systems, fundamentals of digital filters, sampling

3. **Fundamentals of human speech production, acoustic theory of speech production (3lectures)**
   Speech production, short-time Fourier transform, acoustic phonetics

4. **Hearing and speech perception, auditory models (3 lectures)**
   Anatomy and functions of the ear, the perception of sound, auditory models, lossless tube models

5. **Time-domain methods (5 lectures)**
   Short-time analysis (energy, magnitude, zero-crossing rate, autocorrelation)

6. **Frequency-domain methods (7 lectures)**
   Discrete-time Fourier analysis, short-time Fourier analysis, spectrograms, Overlap-add method of synthesis, filter-bank summation method of synthesis

7. **Cepstrum and homomorphic speech processing (5 lectures)**
   Homomorphic analysis, computing the short-time cepstrum and the complex cepstrum, cepstrum analysis of all-pole models, cepstrum distance measures

8. **Linear predictive analysis of speech (6 lectures)**
   Basic ideas, gain computation, frequency-domain interpretation, solving LPC equations, the prediction error signal, representations of LP parameters

9. **Algorithms for estimating speech parameters (5 lectures)**
   Median smoothing, speech-background discrimination, pitch period estimation, formant estimation

10. **Digital coding of speech signals (3 lectures)**
    Sampling speech signals, statistical models for speech signals, quantization (instantaneous, adaptive), quantising speech model parameters, delta modulation, DPCM, ADPCM

11. **Applications: speech recognition, speech enhancement, speaker recognition (4 lectures)** Hidden Markov models for speech recognition, statistical methods for speech enhancement, factor analysis for speaker recognition

Text Books:

1. L.R. Rabiner, R. W. Schafer, *Theory and applications of digital speech processing*, Prentice Hall
2. L.R. Rabiner, R. W. Schafer, *Digital Processing of Speech Signals*, Pearson
3. Douglas O'Shaughnessy, *Speech Communications: Human and Machine*, Wiley India
4. Ben Gold and Nelson Morgan, *Speech and Audio Signal Processing: Processing and Perception of Speech and Music*, Wiley

Reference Books:

1. J. R. Deller, J. H. L. Hansen, J. G Proakis, *Discrete-time processing of speech signals*, Wiley.
2. A. V. Oppenheim, R. W. Schafer, R. Buck, *Discrete-time signal processing*, Pearson