Audio & Speech Processing
EC802D
Contacts: 3L
Credits: 3

Objective: The course provides fundamentals in human speech and music analysis, modeling and processing using digital filters and Pattern Recognition techniques, and an overview of Hidden Markov Models for speech encoding . The different representations of digitized human speech, the importance of adequate voiced and unvoiced speech sounds grouped into phonemes, are used along with spectrograms for speech recognition, articulation and understanding. Also covered are - how the dominant features of speech may be analyzed to form significant abstractions for speaker identification and speaker-independent linguistic comprehension. Prerequisites: Audio Systems, Analog Filters, Digital Signal Processing.

Module -1
Introduction : Production and transmission of acoustic signals : articulation of human speech. Acoustic-phonetic structure of Speech and Music : music synthesis and speech synthesis. A history of Voders & Vocoders and early speech recognition methods.   [4]

Module -2
Acoustic-Phonetic classification : Phonemes, Auto-spectra. Review of Digital Signal Processing and FFT. Short-term Spectral Analysis and STFT, the ARPA and DARPA projects, Pattern matching, introduction to Hidden Markov (HMM) Models. Adaptive segmentation of speech.   [6]

Module -3
The stochastic parameters of human speech, Gaussian densities and statistical model training, voiced and unvoiced speech, voice-box modeling, resonance. Acoustic travelling waves. Psycho-acoustics, Physiological exploration of periodicity, audio-spectrograms and sonograms, pitch-perception models.   [7]

Module -4
Physiology of the ear and hearing mechanism, the Auditory System modeled as a Filter-bank, Gamma-tone and Roex filters, Spectrum and Complex Cepstrum analysis of speech as perceived by detectors, Automatic Speech Recognition (ASR), Linear Prediction analysis   [7]

Module -5
Phonetic and phonemic alphabets, phonological models of ASR, Linear and Dynamic Time-warping, connected word recognition, Statistical sequence recognition and model training in speech pattern recognition, HMM training, Viterbi training, MLP architecture and training,   [8]

Module -6
Speech Synthesis and coding, Formant synthesizers, Vocoders, Speech transformation, Speaker verification, Music synthesizers, speech-assisted applications in industry, defence and medicine.   [5]

Text Books :
14. B.Gold & N.Morgan :- Speech & Audio Signal Processing -*Processing and Perception of Speech & Music* (Wiley Student edition)
15. L.R. Rabiner & B.H.Juang :- Fundamentals of Speech Recognition (Prentice-Hall Signal Processing series)
16. B.Plannerer : An Introduction to Speech Recognition [Freely downloadable e-Book]
17. F.Mihelic & J.Zibert : Speech Recognition (InTech) [Freely downloadable e-Book]
18. I. Mcloughlin :Applied Speech and Audio Processing with MATLAB examples (Cambridge University Press)

Reference Books :
viii) G. Young :-The Application of Hidden Markov Models in Speech Recognition [freely downloadable e-Book]
ix) M.Grimm & K.Kroschel :-Robust Speech Recognition & Understanding (Intech)[Freely downloadable]
x) L. R.Rabiner & R.W.Schafer : Theory and Applications of Digital Speech Processing (Hewlett-Packard Labs/Pearson Pub)
xi) C. Schmandt :- Voice Communication with Computers-Conversational Systems (Van Norstrand Reinhold Computers Series)
xii) SOUND FORGE software package (SONY) for practice sessions [freely downloadable]