

EC6007 SPEECH PROCESSING

DETAILED SYLLABUS

OBJECTIVES:

- To introduce speech production and related parameters of speech.
- To show the computation and use of techniques such as short time Fourier transform, linear predictive coefficients and other coefficients in the analysis of speech.
- To understand different speech modeling procedures such as Markov and their implementation issues.

UNIT I BASIC CONCEPTS

Speech Fundamentals: Articulatory Phonetics – Production and Classification of Speech Sounds; Acoustic Phonetics – Acoustics of speech production; Review of Digital Signal Processing concepts; Short-Time Fourier Transform, Filter-Bank and LPC Methods.

UNIT II SPEECH ANALYSIS

Features, Feature Extraction and Pattern Comparison Techniques: Speech distortion measures– mathematical and perceptual – Log–Spectral Distance, Cepstral Distances, Weighted Cepstral Distances and Filtering, Likelihood Distortions, Spectral Distortion using a Warped Frequency Scale, LPC, PLP and MFCC Coefficients, Time Alignment and Normalization – Dynamic Time Warping, Multiple Time – Alignment Paths.

UNIT III SPEECH MODELING

Hidden Markov Models: Markov Processes, HMMs – Evaluation, Optimal State Sequence – Viterbi Search, Baum-Welch Parameter Re-estimation, Implementation issues.

UNIT IV SPEECH RECOGNITION

Large Vocabulary Continuous Speech Recognition: Architecture of a large vocabulary continuous speech recognition system – acoustics and language models – n-grams, context dependent sub-word units; Applications and present status.

UNIT V SPEECH SYNTHESIS

Text-to-Speech Synthesis: Concatenative and waveform synthesis methods, sub-word units for TTS, intelligibility and naturalness – role of prosody, Applications and present status.

TEXTBOOKS:

1. Lawrence Rabiner and Biing-Hwang Juang, “Fundamentals of Speech Recognition”, Pearson Education, 2003.
2. Daniel Jurafsky and James H Martin, “Speech and Language Processing – An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition”, Pearson Education, 2002.
3. Frederick Jelinek, “Statistical Methods of Speech Recognition”, MIT Press, 1997.

REFERENCES:

1. Steven W. Smith, "The Scientist and Engineer's Guide to Digital Signal Processing", California Technical Publishing, 1997.
2. Thomas F Quatieri, "Discrete-Time Speech Signal Processing – Principles and Practice", Pearson Education, 2004.
3. Claudio Becchetti and Lucio Prina Ricotti, "Speech Recognition", John Wiley and Sons, 1999.
4. Ben Gold and Nelson Morgan, "Speech and Audio Signal Processing, Processing and Perception of Speech and Music", Wiley- India Edition, 2006.