

RAJIV GANDHI PROUDYOGIKI VISHWAVIDYALAYA, BHOPAL

New Scheme Based On AICTE Flexible Curricula

Electronics & Communication Engineering, VIII-Semester

Open Elective EC 803 (C) Speech Processing

Course Objectives

To study speech production and related parameters of speech and to understand different speech modeling procedures such as Markov and their implementation issues.

Course Outcomes:

1. Model speech production system and describe the fundamentals of speech.
2. Extract and compare different speech parameters.
3. Choose an appropriate statistical speech model for a given application.
4. Design a speech recognition system.
5. Use different speech synthesis techniques.

Unit-I

Basic Concepts of Speech Processing

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 – ngrams, context dependent sub-word units; Applications and present status.

Unit-V

Speech Synthesis

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

References:

2. Lawrence Rabiner and Biing-Hwang Juang, “Fundamentals of Speech Recognition”, Pearson Education, 2003.
3. Daniel Jurafsky and James H Martin, “Speech and Language Processing – An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition”, Pearson Education.
4. Steven W. Smith, “The Scientist and Engineer’s Guide to Digital Signal Processing”, California Technical Publishing.
5. Thomas F Quatieri, “Discrete-Time Speech Signal Processing – Principles and Practice”, Pearson Education.
6. Claudio Becchetti and Lucio Prina Ricotti, “Speech Recognition”, John Wiley and Sons, 1999.
7. Ben Gold and Nelson Morgan, “Speech and audio signal processing”, processing and perception of speech and music, Wiley- India Edition, 2006 Edition.
8. Frederick Jelinek, “Statistical Methods of Speech Recognition”, MIT Press.