AP5074 SPEECH AND AUDIO SIGNAL PROCESSING

DETAILED SYLLABUS

OBJECTIVES:

- To study basic concepts of processing speech and audio signals
- To study and analyse various M-band filter-banks for audio coding
- To understand audio coding based on transform coders.
- To study time and frequency domain speech processing methods

UNIT I MECHANICS OF SPEECH AND AUDIO

Introduction - Review of Signal Processing Theory-Speech production mechanism – Nature of Speech signal – Discrete time modelling of Speech production – Classification of Speech sounds – Phones – Phonemes – Phonetic and Phonemic alphabets – Articulatory features. Absolute Threshold of Hearing - Critical Bands- Simultaneous Masking, Masking-Asymmetry, and the Spread of Masking- Non simultaneous Masking - Perceptual Entropy - Basic measuring philosophy -Subjective versus objective perceptual testing - The perceptual audio quality measure (PAQM) - Cognitive effects in judging audio quality.

UNIT II TIME-FREQUENCY ANALYSIS: FILTER BANKS AND TRANSFORMS

Introduction - Analysis-Synthesis Framework for M-band Filter Banks- Filter Banks for Audio Coding: Design Considerations - Quadrature Mirror and Conjugate Quadrature Filters - Tree-Structured QMF and CQF M-band Banks - Cosine Modulated "Pseudo QMF" M-band Banks -Cosine Modulated Perfect Reconstruction (PR) M-band Banks and the Modified Discrete Cosine Transform (MDCT) - Discrete Fourier and Discrete Cosine Transform - Pre-echo Distortion- Pre-echo Control Strategies

UNIT III AUDIO CODING AND TRANSFORM CODERS

Lossless Audio Coding – Lossy Audio Coding - ISO-MPEG-1A, 2A, 2A-Advaned, 4A Audio Coding - Optimum Coding in the Frequency Domain - Perceptual Transform Coder – Brandenburg – Johnston Hybrid Coder - CNET Coders - Adaptive Spectral Entropy Coding – Differential Perceptual Audio Coder - DFT Noise Substitution -DCT with Vector Quantization - MDCT with Vector Quantization

UNIT IV TIME AND FREQUENCY DOMAIN METHODS FOR SPEECH PROCESSING

Time domain parameters of Speech signal – Methods for extracting the parameters: Energy, Average Magnitude – Zero crossing Rate – Silence Discrimination using ZCR and energy Short Time Fourier analysis – Formant extraction – Pitch Extraction using time and frequency domain methods Homomorphic Speech Analysis: Cepstral analysis of Speech – Formant and Pitch Estimation – Homomorphic Vocoders

UNIT V PREDICTIVE ANALYSIS OF SPEECH

Formulation of Linear Prediction problem in Time Domain – Basic Principle – Auto correlation method – Covariance method – Solution of LPC equations – Cholesky method – Durbin's Recursive algorithm – lattice formation and solutions – Comparison of different methods – Application of LPC parameters – Pitch detection using LPC parameters – Formant analysis – VELP – CELP

REFERENCES:

1. B. Gold and N. Morgan, “Speech and Audio Signal Processing”, Wiley and Sons, 2000.

2. L. R. Rabiner and R. W. Schaffer, "Digital Processing of Speech Signals", Prentice Hall, 1978.

3. Mark Kahrs, Karlheinz Brandenburg, Kluwer Applications of Digital Signal Processing to Audio and Acoustics, Academic Publishers,

4. Udo Zölzer, "Digital Audio Signal Processing", Second Edition A John Wiley& sons Ltd