IMAGE AND SPEECH PROCESSING

V Semester: CSE(AI & ML)

Course Code	Category	Hours / Week			Credits	Maximum Marks		
		L	T	P	C	CIA	SEE	Total
ACAC05	Core	3	1	0	4	30	70	100
Contact Classes: 45	Tutorial Classes: Nil	Practical Classes: 15				Total Classes:60		

Prerequisites: Linear Algebra and Calculus

I. COURSE OVERVIEW:

This course introduces the fundamental concepts and techniques of digital image processing, concentrating on aspects of image processing in which both the inputs and outputs are images. It include image sampling and quantization, intensity transformation, spatial filtering, frequency domain filtering, image compression. It also deals with speech processing that ranges from the basic nature of the speech signal, through a variety of methods of representing speech in digital form, to applications in voice communication and automatic synthesis and recognition of speech. Speech processing in widely used is applications like google voice search, artificial intelligence voice controlled assistant like Alexia.

II. COURSE OBJECTIVES:

The students will try to learn:

I. The fundamental concepts of Digital Image Processing.
II. The algorithms to solve image processing problems and meet design specifications for industry, medicine and defense applications.
III. Methods and systems for efficient quantization and coding of speech signals.
IV. The concepts of linear predictive analysis (LPC) for speech synthesis

III. COURSE OUTCOMES:

After successful completion of the course, students should be able to:

CO 1	**Make use of** image transform techniques for analyzing images in transformation domain for image pre-processing.	Understand
CO 2	**List** the lossy and lossless compression models for achieving imagecompression.	Analyze
CO 3	**Illustate** the difference between acoustic phonetics and articulatory phonetics for speech processing	Understand
CO 4	**Utilize** digital model designed by sampled speech signal for speech processing applications like speech recognition, speech synthesis and verification.	Apply
CO 5	**Analyze** methods to estimate pitch period to designvocoders, artificial intelligence voice-controlled assistants like Alexa	Analyze
CO 6	**Apply** linear predictive coding for speech synthesis, compression and spectrographic displays	Apply

IV. SYLLABUS:

MODULE – I: DIGITAL IMAGE BASICS AND INTENSITY TRANSFORMATION (09)

Digital image basics, Brightness, Contrast and grey levels, Fundamentals steps in image processing, Intensity Transformation, Histogram equalization, Histogram matching, Smoothening and Sharpening spatial filters.

MODULE – II: IMAGE COMPRESSION (09)

Fundamental Redundancy, Image Compression Models, Coding Theorems, Entropy, Error-Free Compression, Lossy Compression, LZW coding, Transform Coding, JPEG-2000 encoding, Lossless predictive coding, Lossy predictive coding.

MODULE – III: FUNDAMENTALS OF HUMAN SPEECH PRODUCTION (09)

The Process of Speech Production, Short-Time Fourier Representation of Speech, The Acoustic Theory of Speech Production.
Lossless Tube Models of the Vocal Tract, Digital Models for Sampled Speech Signals.

MODULE – IV: TIME DOMAIN MODELS FOR SPEECH PROCESSING (09)

Time dependent processing of speech, short time energy and average magnitude, short time average zero crossing rate, discrimination between speech& silence, pitch period estimation using parallel processing, short time autocorrelation function and AMDF, pitch period estimation using autocorrelation function.

MODULE - V : SHORT TIME FOURIER ANALYSISAND LINEAR PREDICTIVE CODING (09)

Definition and properties, design of filter banks, implementation of filter bank summation method using FFT, spectrographic displays, pitch detection, analysis by synthesis phase, vocoder and channel vocoder, Basic principles of linear predictive analysis, the autocorrelation method, computation of the gain for the model, solution of LPC equations for auto correlation method, prediction error and normalized mean square error.

V. TEXT BOOKS:

1. Anil K. Jain, “Fundamentals of Digital Image Processing”, Pearson, 2002.
2. Lawrence R. Rabiner, Ronald W. Schafer, "Digital Processing of Speech Signals", Pearson Education, 2012.

VI. REFERENCE BOOKS:

1. Kenneth R. Castleman, “Digital Image Processing”, Pearson, 2006.
2. Rafael C. Gonzalez, Richard E. Woods, Steven Eddins, “Digital Image Processing using MATLAB”, Pearson Education Inc., 2011.
3. R. L. Rabiner, R.W. Schafer, “Digital Processing of speech signals”, Pearson Education.
4. B. Gold and Nelson Morgon, “Speech and audio signal processing”, Wiley India Edition,2006.
5. Dan Jurafsky and James H. Martin, “Speech and Language Processing”, 2nd Edition, Pearson, 2017.

VII. WEB REFERENCES:

1. https://nptel.ac.in/courses/106/105/106105032/
2. https://sisu.ut.ee/imageprocessing/documents
3. https://research.iaun.ac.ir/pd/mahmoodian/pdfs/UploadFile_2643.pdf
4. https://tinyurl.com/yjcmyrcd
5. http://www.speech.cs.cmu.edu/15-492/