

SEMESTER S8

SPEECH AND AUDIO PROCESSING

(Common to CS/CA/CM/CD/CR/AD/CC/CG)

Course Code	PECST866	CIE Marks	40
Teaching Hours/Week (L:T:P: R)	3:0:0:0	ESE Marks	60
Credits	3	Exam Hours	2 Hrs. 30 Min.
Prerequisites (if any)	PECST636	Course Type	Theory

Course Objectives:

1. To get familiarised with speech processing and audio processing concepts.
2. To equip the student to apply speech processing techniques in finding solutions to day-to-day problems

SYLLABUS

Module No.	Syllabus Description	Contact Hours
1	Speech Production :- Acoustic theory of speech production; Source/Filter model - Pitch, Formant; Spectrogram- Wide and narrow band spectrogram; Discrete model for speech production; Short-Time Speech Analysis; Windowing; STFT; Time domain parameters (Short time energy, short time zero crossing Rate, ACF); Frequency domain parameters - Filter bank analysis; STFT Analysis.	9
2	Mel-frequency cepstral coefficient (MFCC)- Computation; Pitch Estimation ACF/AMDF approaches; Cepstral analysis - Pitch and Formant estimation using cepstral analysis; <i>LPC Analysis</i> - LPC model; Auto correlation method - Levinson Durbin Algorithm	9
3	Speech Enhancement :- Spectral subtraction and Filtering, Harmonic filtering, Parametric resynthesis; Speech coding - fundamentals, class of coders : Time domain/spectral domain/vocoders, Sub band coding, adaptive transform coding, phase vocoder; Speaker Recognition :- Speaker verification and speaker identification, log-likelihood; Language identification - Implicit and explicit models; Machine learning models in Speaker Recognition.	9

4	Signal Processing models of audio perception - Basic anatomy of hearing System, Basilar membrane behaviour; Sound perception - Auditory Filter Banks, Critical Band Structure, Absolute Threshold of Hearing; Masking - Simultaneous Masking, Temporal Masking; Models of speech perception.	9
----------	--	----------

**Course Assessment Method
(CIE: 40 marks, ESE: 60 marks)**

Continuous Internal Evaluation Marks (CIE):

Attendance	Assignment/ Microproject	Internal Examination-1 (Written)	Internal Examination- 2 (Written)	Total
5	15	10	10	40

End Semester Examination Marks (ESE)

In Part A, all questions need to be answered and in Part B, each student can choose any one full question out of two questions

Part A	Part B	Total
<ul style="list-style-type: none"> ● 2 Questions from each module. ● Total of 8 Questions, each carrying 3 marks <p style="text-align: center;">(8x3 =24 marks)</p>	<ul style="list-style-type: none"> ● Each question carries 9 marks. ● Two questions will be given from each module, out of which 1 question should be answered. ● Each question can have a maximum of 3 subdivisions. <p style="text-align: center;">(4x9 = 36 marks)</p>	60

Course Outcomes (COs)

At the end of the course students should be able to:

Course Outcome		Bloom's Knowledge Level (KL)
CO1	To recall various steps in the speech production process	K2
CO2	To summarise various speech processing approaches	K2
CO3	To develop speech-processing applications in various domains	K3
CO4	To analyse the speech processing model for audio perception	K4

Note: K1- Remember, K2- Understand, K3- Apply, K4- Analyse, K5- Evaluate, K6- Create

CO-PO Mapping Table (Mapping of Course Outcomes to Program Outcomes)

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	3	3	3	2		2	2					3
CO2	3	3	3									3
CO3	3	3	3									3
CO4	3	3	3	2			2					3

Note: 1: Slight (Low), 2: Moderate (Medium), 3: Substantial (High), -: No Correlation

Text Books				
Sl. No	Title of the Book	Name of the Author/s	Name of the Publisher	Edition and Year
1	Speech Communications: Human & Machine	Douglas O'Shaughnessy	IEEE Press	2/e, 1999
2	Discrete-Time Speech Signal Processing: Principles and Practice	Thomas F. Quatieri	Prentice Hall	1/e, 2001
3	Fundamentals of Speech Recognition	Lawrence Rabiner, Biing-Hwang Juang, B. Yegnanarayana	Pearson	1/e, 2008

Reference Books				
Sl. No	Title of the Book	Name of the Author/s	Name of the Publisher	Edition and Year
1	Theory and Application of Digital Processing of Speech Signals	Rabiner and Schafer	Prentice Hall	1/e, 2010
2	Speech and Audio Signal Processing: Processing and Perception Speech and Music	Nelson Morgan and Ben Gold	John Wiley & Sons	2/e, 2011

Video Links (NPTEL, SWAYAM...)	
No.	Link ID
1	https://youtu.be/Xjzm7S_kBU?si=j11bk3F7gocYjhfg