Course Code	Course Name	L	T	P	C
1152BM108	**Speech Processing**	**3**	**0**	**0**	**3**

Course Category:

Program Elective

Preamble:

This course helps the students to learn about the fundamentals of speech analysis and their applications. Also this course gives the detail information about the different speech synthesis techniques.

Prerequisite:

Digital Signal Processing

Related Course:

Outcome:

1.	**Model Speech Production System and describe the fundamentals of speech**	**K2**
2.	**Compare the different feature extraction methods for speech analysis based on speech parameters**	**K4**
3.	**Choose an appropriate Statistical speech model for a given application**	**K3**
4.	**Design a speech recognition system**	**K3**
5.	**Use different speech synthesis techniques**	**K3**

	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	**M**		**H**	**L**								
CO2	**H**	**L**	**H**	**M**								
CO3		**L**	**M**	**M**								
CO4			**H**			**M**	**M**		**M**			**M**
CO5		**M**	**M**									

Unit I: Basic Concepts 10

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 **10**

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, PLP and MFCC Coefficients, Time Alignment and Normalization – Dynamic Time Warping, Multiple Time – Alignment Paths.

Unit III: Modeling **8**

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

Unit IV: Speech Recognition **8**

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

Unit V: Speech Synthesis: **9**

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

Total: 45 Hours

TEXT BOOKS

1.LawrenceRabinerandBiing-Hwang Juang, “Fundamentals of Speech Recognition”, Pearson Education, 2003.

2.DanielJurafsky and James H Martin, “Speech and Language Processing – An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition”, Pearson Education.

REFERENCES

1.Steven W. Smith, “The Scientist and Engineer’s Guide to Digital Signal Processing”, California Technical Publishing.

2.Thomas F Quatieri, “Discrete-Time Speech Signal Processing – Principles and Practice”, Pearson Education.

3.ClaudioBecchetti and LucioPrinaRicotti, “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 Edition.

5.FrederickJelinek, “Statistical Methods of Speech Recognition”, MIT Press.