

## **Advanced topics in speech processing (IT 60116) (3-0-0)**

**Introduction:** Speech production and perception mechanisms, Speech Signal Processing Methods (6 hours)

**Knowledge sources in speech:** Time domain and frequency domain, Spectrograms, Knowledge sources at segmental, sub-segmental and supra-segmental (prosodic) levels, excitation source, vocal tract system and higher level knowledge sources and linguistic and semantic knowledge. (6 hours)

**Modeling techniques for developing speech systems:** Vector quantization, Hidden Markov models, Gaussian mixture models, Support vector machines and Neural networks (8 hours)

**Speech Coding:** Coding of speech signals, Waveform coding, Speech-specific coders (4 hours)

**Speech Recognition:** Issues in speech recognition, Isolated word recognition, Connected word recognition, Continuous speech recognition, Large vocabulary continuous speech recognition. (4 hours)

**Speech Synthesis:** Issues in speech synthesis, Models for speech synthesis, Different speech synthesis systems, Prosodic aspects in speech synthesis, Development of speech synthesis system. Evaluation methodologies for speech synthesis systems. (4 hours)

**Speaker Recognition:** Issues in speaker recognition, Speaker verification vs identification, Text-dependent vs text-independent speaker recognition, Development of speaker recognition systems. (4 hours)

**Speech Enhancement:** Enhancement of noisy speech, Enhancement of reverberant speech, Enhancement of multi-speaker speech. (4 hours)

### **Text books:**

1. L. R. Rabiner and B. H. Juang, *Fundamentals of Speech Recognition*, Pearson Education, Delhi, India, 2003.
2. D. O'Shaughnessy, *Speech Communication: Human and Machine*, 2<sup>nd</sup> edition, IEEE Press, NY, USA, 1999.
3. J. R. Deller, Jr., J. H. L. Hansen and J. G. Proakis, *Discrete-time Processing of Speech Signals*, IEEE Press, NY, USA, 1999.
4. T.F. Quateri, *Discrete-Time Speech Signal Processing: Principles and Practice*, Pearson Education, 2004.

### **References:**

1. J. Benesty, M. M. Sondhi and Y. Huang, "*Springer Handbook on Speech Processing*", Springer publishers, 2008.
2. B. Gold and N. Morgan, *Speech and Audio Signal Processing*, Wiley Student Edition, Singapore, 2004.
3. Ahmet M. Kondoz, *Digital Speech: Coding for Low Bit Rate Communication*, 2<sup>nd</sup> ed, Wiley publication, 2004.
4. X. Huang, A. Acero and H. W. Hon, "*Spoken Language Processing*", Printice-Hall, Inc., 2001
5. IEEE Trans. on Speech and Audio Processing.
6. *Speech Communication* (Elsiver)
7. *Computer, Speech and Language* (Elsiver)

**Instructor:** K. Sreenivasa Rao, Assistant Professor, School of Information Technology, IIT Kharagpur.