New Elective

Advanced topics in speech processing (IT 60116)

Semester: Spring (Jan – Apr)   **Credits: 3-0-0**
Slot: F   **Prerequisites: None**
Instructor: K. Sreenivasa Rao, Assistant Professor, School of Information Technology, IIT Kharagpur.

Detailed course contents

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. D. O'Shaughnessy, *Speech Communication: Human and Machine*, 2nd edition, IEEE Press, NY, USA, 1999.
2. J. R. Deller, Jr., J. H. L. Hansen and J. G. Proakis, *Discrete-time Processing of Speech Signals*, IEEE Press, NY, USA, 1999.
3. T.F. Quateri, Discrete-Time Speech Signal Processing: Principles and Practice, Pearson Education, 2004.

References:
1. L. R. Rabiner and B. H. Juang, *Fundamentals of Speech Recognition*, Pearson Education, Delhi, India, 2003.
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, 2nd ed, Wiley publication, 2004.
4. IEEE Trans. on Speech and Audio Processing.
5. Speech Communication (Elsiver)
6. Computer, Speech and Language (Elsiver)

Copy to: HOD (CSE, ECE, EE, CET and GSSST)