



DIGITAL SPEECH PROCESSING

PROF. SHYAMAL KUMAR DAS MANDAL

Advanced Technology Development Centre
IIT Kharagpur

PRE-REQUISITES : Digital Signal Processing or Signals and System

INTENDED AUDIENCE : ECE, CS, EE, IE

INDUSTRY SUPPORT : Companies, Industry like Microsoft, Google , IBM who are working in the area of speech technology development

COURSE OUTLINE :

Oral Speech may be the most natural, common and direct mode of human communication. Since the middle of the last century, Speech has become an area of intense and active research and development (R&D) to become a prime means of direct Human-Computer Interactions (HCI). The pace of such R&D has further got boosted with the general abundance of cheap computing power in the form of PC, PDA or Mobile Handset. While man to machine in speech mode is yet to reach the minimum threshold level for wide-spread deployment, spoken messages directly by machine. This need research in speech science and development of speech technology. The course provides the foundation knowledge on speech production and perception along with processing of speech signal in digital domain.

ABOUT INSTRUCTOR :

Prof. S K. Das Mandal was born on October 1975; He received the B.E degree in Electronics and Telecommunication engineering in 1998 and Ph.D degree in 2007 from Jadavpur University, India and currently working in Indian institute of Technology Kharagpur as an Assistant Professor. His current research interests include automatic speech recognition, speech synthesis, and computer assisted spoken language acquisition.

COURSE PLAN :

- Week 1:** Introduction to speech processing, Digitization and Recording of speech signal, Review of Digital Signal Processing Concepts
- Week 2:** Human Speech production, Acoustic Phonetics and Articulatory Phonetics, Different categories speech sounds and Location of sounds in the acoustic waveform and spectrograms
- Week 3:** Uniform Tube Modeling of Speech Production, Speech Perception
- Week 4:** Time Domain Methods in Speech Processing, Analysis and Synthesis of Pole-Zero Speech Models
- Week 5:** Short-Time Fourier Transform, Analysis:- FT view and Filtering view, Synthesis:-Filter bank summation (FBS) Method and OLA Method
- Week 6:** Features Extraction, Extraction of Fundamental frequency
- Week 7:** Speech Prosody, Speech Prosody Modeling (Fujisaki Model)
- Week 8:** Speech based Applications (TTS, ASR and spoken language acquisition)