Speech Processing

(Elective)

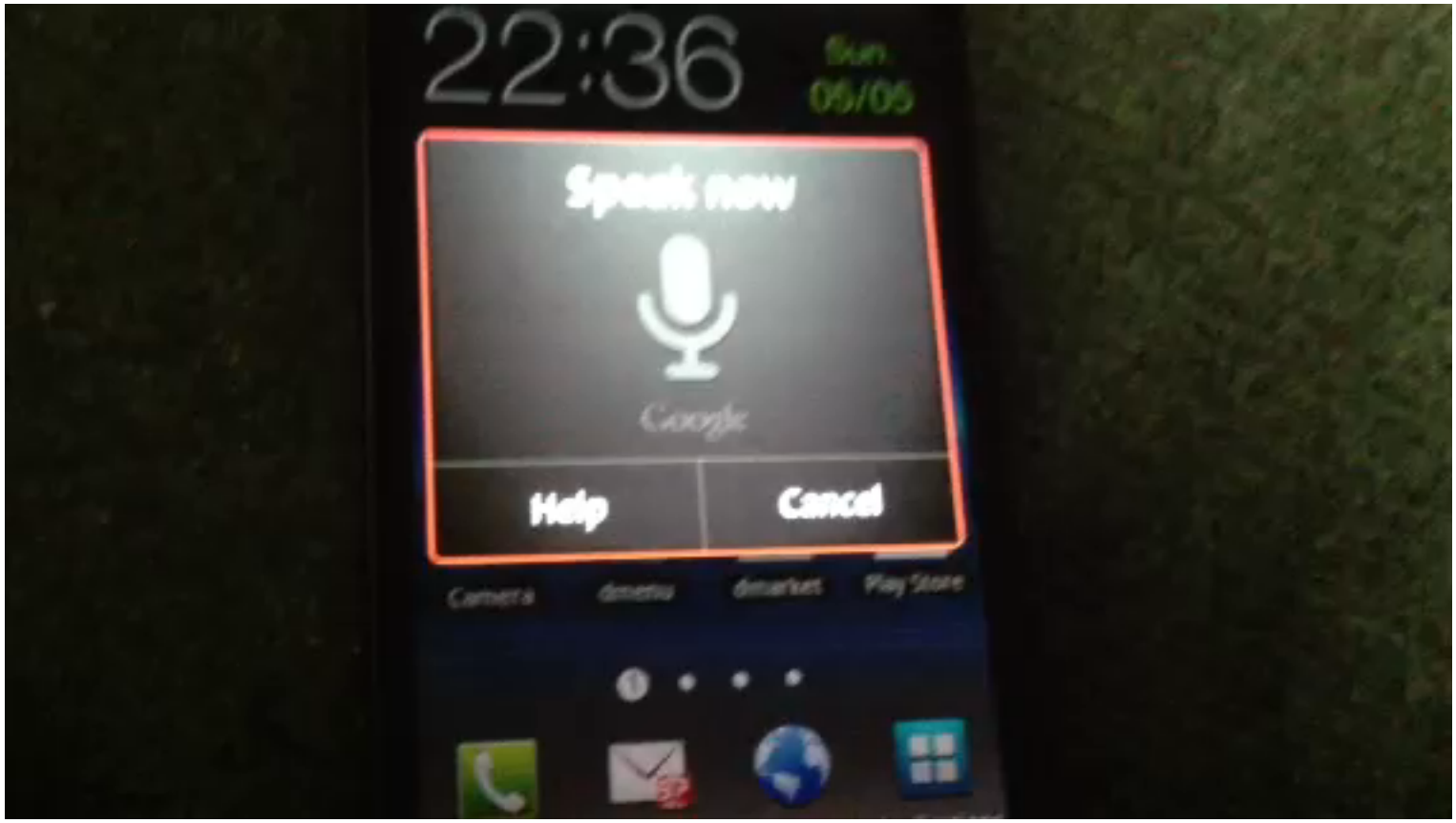
Year : IV
Lecture : 3
Practical : 1.5

Part : II
Tutorial : 0

Basanta Joshi, PhD
basanta@ioe.edu.np
www.basantajoshi.com.np

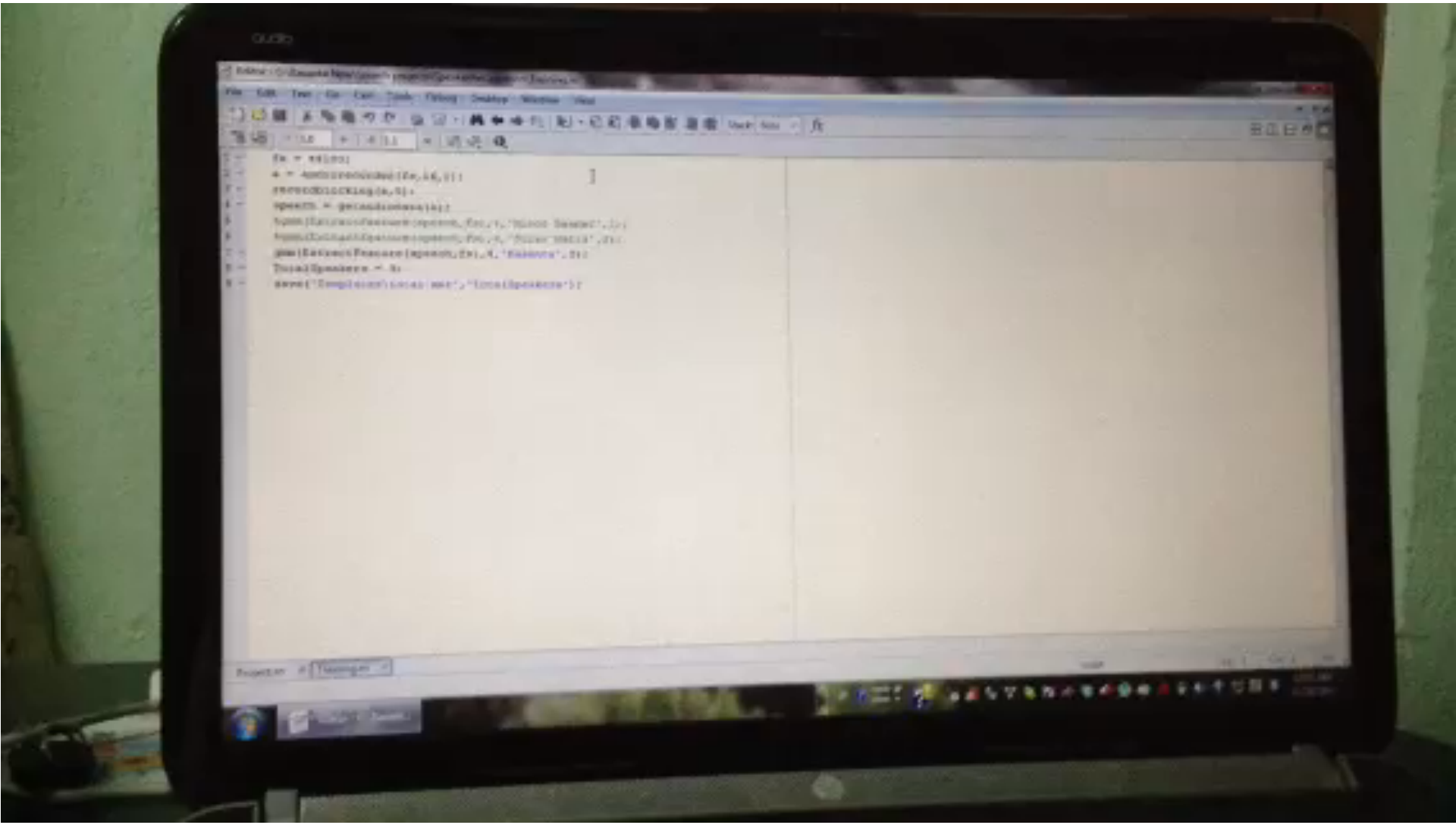
Applications

Applications



Applications

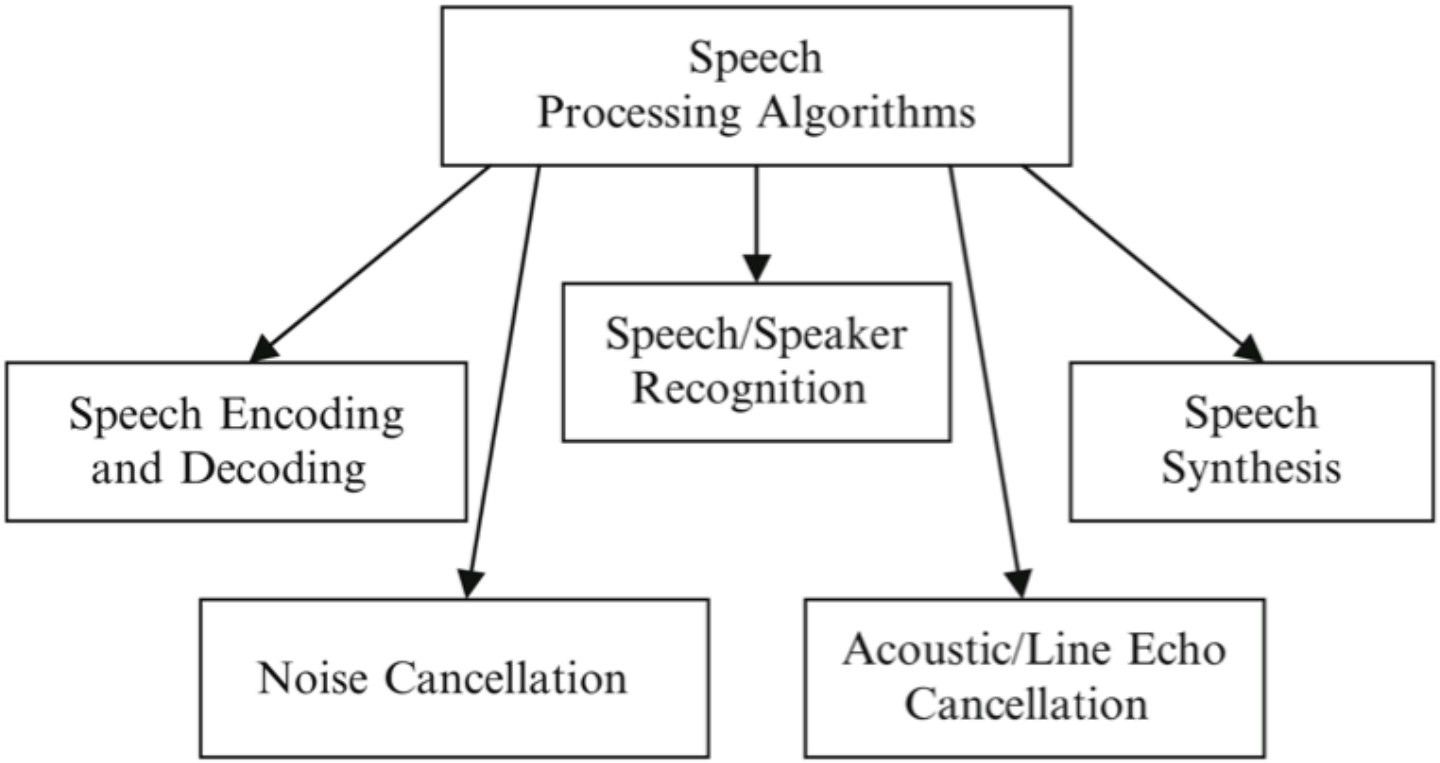
Applications



Speech Processing

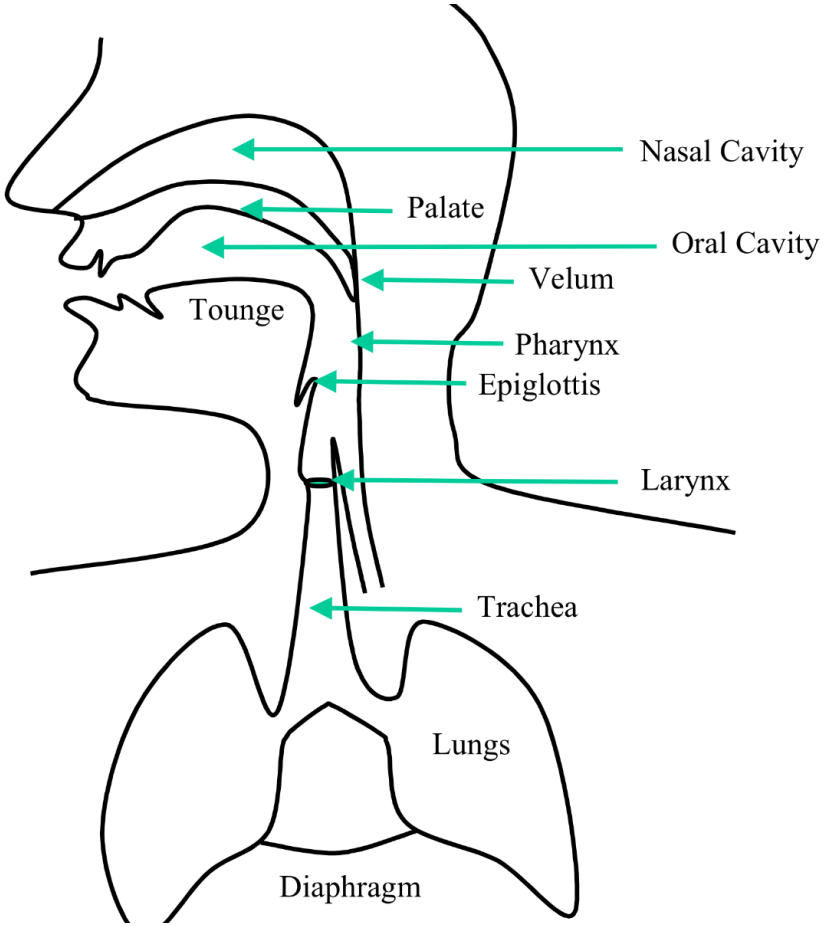
- Speech is means of transferring thoughts, ideas, directions, and emotions and information reach signal.
- Speech processing is the study of speech signals and the processing methods of these signals.
- Speech processing aids in perceiving the unknown signal.

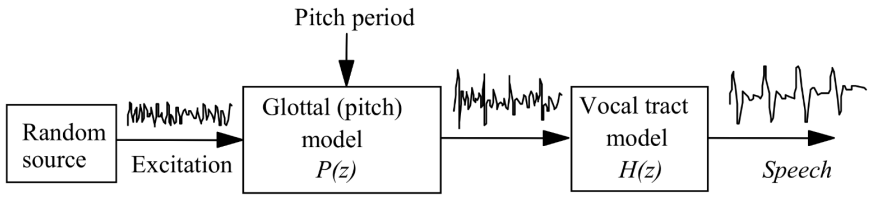
Signal processing tasks



Speech

- Speech sounds are sensations of air pressure vibrations produced by air exhaled from the lungs and
- modulated and shaped by the vibrations of the glottal cords and
- the resonance of the vocal tract as the air is pushed out through the lips and nose





Applications

Table 1.1 Speech processing application examples in various market segments

Telecom	Automotive	Consumer/medical	Industrial/military
Intercom systems	Car mobile hands-free kits	Talking toys	Test equipment with spoken instructions
Speakerphones	Talking GPS units	Medical emergency phones	Satellite phones
Walkie-talkies	Voice recorders during car service	Appliances with spoken instructions	Radios
Voice-over-IP phones	Voice activated dialing	Recorders for physician's notes	Noise cancelling helmets
Analog telephone adapters	Voice instructions during car service	Appliances with voice record and playback	Public address systems
Mobile phones	Public announcement systems	Bluetooth headsets	Noise cancelling headsets
Telephones	Voice activated car controls	Dolls with customized voices	Security panels

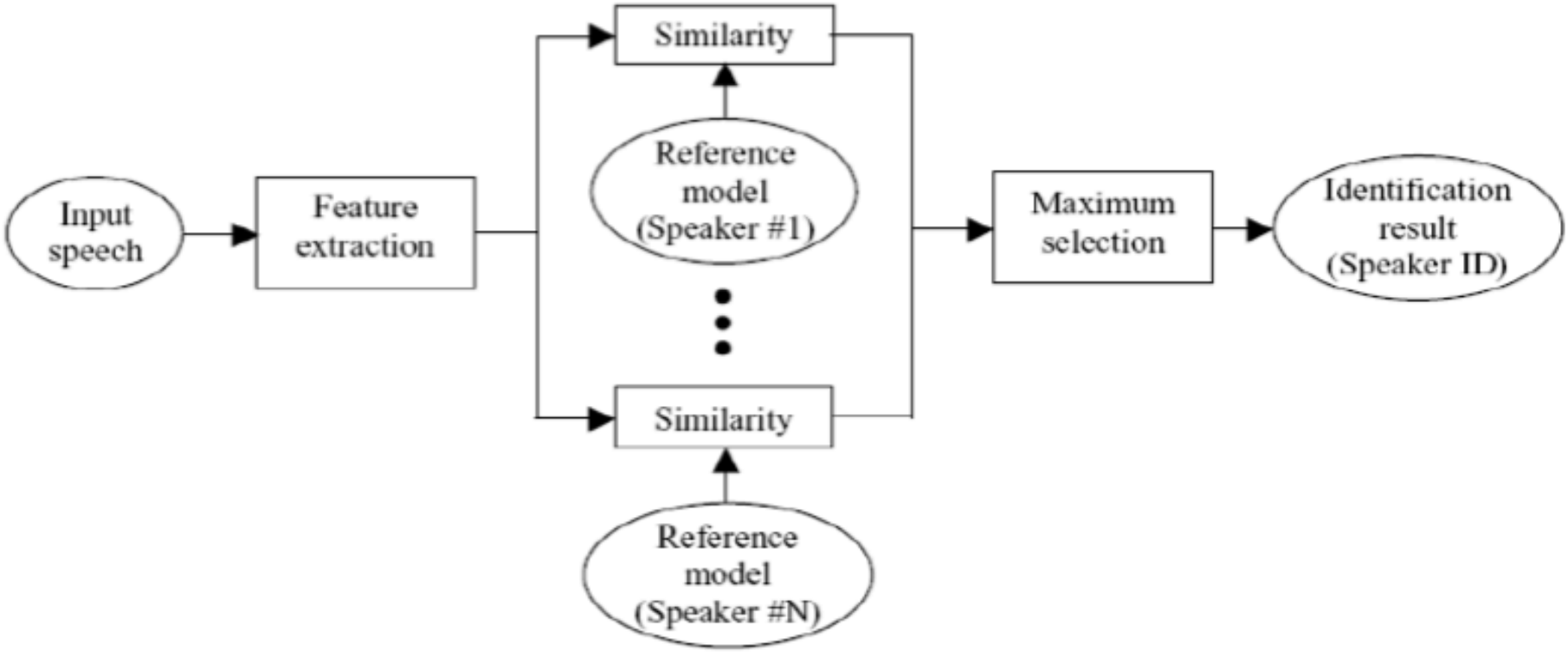
Current Popular Applications

- Siri
- Google voice Search
- TTS
- Voice to Text translators
- Song Identifications (Soundhound)
- Speech recognition
- Speaker Identification

Still there are lots of area where we need to explore.

Application for NEPALI language

Speaker Identification



Course Objectives

- To introduce the characteristics of Speech signals and the related time and frequency domain methods for speech analysis and speech compression
- To introduce the models for speech production
- To develop time and frequency domain techniques for estimating speech parameters
- To introduce a predictive technique for speech compression
- To understand basic NLP, speech recognition and speaker identification.

Syllabus

- 1. Nature of speech signal
- 2. Time domain methods for speech processing
- 3. Frequency domain method for speech processing
- 4. Linear predictive analysis of speech
- 5. An introduction to natural language processing.
- 6. Speech & Speaker Recognition

Syllabus in Detail

1. Nature of speech signal

- Speech production: Mechanism of speech production
- Acoustic phonetics
- Digitalmodels for speech signals
- Representations of speech waveform
  - Sampling speechsignals
  - Basics of quantization
  - Delta modulation
  - Differential PCM
- Auditoryperception: psychoacoustics.

2. Time domain methods for speech processing

- Time domain parameters of Speech signal
- Methods for extracting the parameters
  - Short-time Energy
  - Average Magnitude
  - Short-time average Zero crossing Rate
- Silence Discrimination using ZCR and energy
- Short Time Auto Correlation Function
- Pitch period estimation using AutoCorrelation Function

Syllabus in Detail

3. **Frequency domain method for speech processing**
   - Short Time Fourier analysis
     - Fourier transform and linear filtering interpretations
     - Sampling rates
   - Spectrographic displays
   - Pitch and formant extraction
   - Analysis bySynthesis
   - Analysis synthesis systems
     - Phase vocoder
     - Channel Vocoder
   - Homomorphic speech analysis
     - Cepstral analysis of Speech
     - Formant and PitchEstimation
     - Homomorphic Vocoders

Syllabus in Detail

4. Linear predictive analysis of speech

- Basic Principles of linear predictive analysis
- Auto correlation method
- Covariance method
- Solution of LPC equations
- Cholesky method
- Durbin's Recursive algorithm
- Application of LPC parameters
  - Pitch detection using LPC parameters
  - Formant analysis
  - VELP
  - CELP

Syllabus in Detail

5. Application of speech & audio signal processing

- Algorithms:
  - Dynamic time warping
  - K-means clustering and Vector quantization
  - Gaussian mixture modeling
  - Hidden Markov modeling
- Basics of NLP
- Automatic Speech Recognition
  - Feature Extraction for ASR
  - Deterministic sequence recognition
  - Statistical Sequence
  - Recognition
  - Language models
- Speaker identification and verification
- Voice response system
- Speech synthesis
  - Basics of articulatory
  - Source-filter
  - Concatenative synthesis

Practical

- There should be at 4-6 experiments based on following topics
- Spectral analysis
- Time-Frequency analysis
- Pitch extraction
- Formant tracking
- Speech enhancement
- Audio coding
- Speaker recognition

All these lab works may be performed in Matlab or similar softwares capable of processing speech signals.

Students must implement an algorithm for speech processing and submit it at the end.

References

- Thomas F. Quatieri, “Discrete-Time Speech Signal Processing”, Prentice Hall /Pearson Education.
- Ben Gold and Nelson Morgan, “Speech and Audio Signal Processing”, John Wiley and Sons Inc.
- L.R.Rabiner and R.W.Schaffer, “Digital Processing of Speech signals”, Prentice Hall
- L.R. Rabiner and B. H. Juang, “Fundamentals of Speech Recognition”, Prentice Hall.
- J.R. Deller, J.H.L. Hansen and J.G. Proakis, “Discrete Time Processing of SpeechSignals”, John Wiley, IEEE Press.
- J.L Flanagan, “Speech Analysis Synthesis and Perception”,Springer, Verlag.

Thank you !!!

Contact me
basanta@ioe.edu.np
Download lecture slides via
www.basantajoshi.com.np