Automatic speech processing

Magimai Doss Mathew

Summary
The goal of this course is to provide the students with the main formalisms, models and algorithms required for the implementation of advanced speech processing applications (involving, among others, speech coding, speech analysis/synthesis, and speech recognition).

Content
1. Introduction: Speech processing tasks, language engineering applications.
2. Basic Tools: Analysis and spectral properties of the speech signal, linear prediction algorithms, statistical pattern recognition, dynamic programming.
5. Automatic Speech Recognition: Temporal pattern matching and Dynamic Time Warping (DTW) algorithms, speech recognition systems based on Hidden Markov Models (HMMs).
6. Speaker recognition and speaker verification: Formalism, hypothesis testing, HMM based speaker verification.
7. Linguistic Engineering: state-of-the-art and typical applications

Keywords
speech processing, speech coding, speech analysis/synthesis, automatic speech recognition, speaker identification, text-to-speech

Learning Prerequisites
Required courses
Basis in linear algebra, signal processing (FFT), and statistics

Important concepts to start the course
Basic knowledge in signal processing, linear algebra, statistics and stochastic processes.

Learning Outcomes
By the end of the course, the student must be able to:
• speech signal properties
• Exploit those properties to speech coding, speech synthesis, and speech recognition

Transversal skills
• Use a work methodology appropriate to the task.
• Access and evaluate appropriate sources of information.
• Use both general and domain specific IT resources and tools

Teaching methods
Lecture + lab exercises

Expected student activities
Attending courses and lab exercises. Read additional papers and continue lab exercises at home if necessary. Regularly answer list of questions for feedback.

Assessment methods
Written exam without notes

Supervision
Office hours No
Assistants Yes
Forum No

Resources
Bibliography
Fundamentals of Speech Recognition / Rabiner and Juang

Ressources en bibliothèque
• Fundamentals of Speech Recognition / Rabiner and Juang

Websites
• http://lectures.idiap.ch/

Moodle Link
• https://go.epfl.ch/EE-554