FDT3119 Speech and Speaker Recognition 7.5 credits

Igenkänning av tal och talare

This is a translation of the Swedish, legally binding, course syllabus.

Establishment

Course syllabus for FDT3119 valid from Spring 2019

Grading scale

P, F

Education cycle

Third cycle

Specific prerequisites

Doctoral students from EECS

Language of instruction

The language of instruction is specified in the course offering information in the course catalogue.

Intended learning outcomes

After completion of the course the students should be able to:

• implement training and evaluation methods for speech recognition

---

- train and evaluate a speech recogniser using software packages
- compare different feature extraction and training methods
- document and discuss specific aspects related to speech and speaker recognition
- with the help of the literature, review and criticise other students' work in the subject

Course contents

The course consists of lectures, three laboratory sessions with hand-in assignments, as well as writing a thesis in a subject chosen in consultation with the teacher. The thesis is furthermore presented orally during a final seminar. The laboratory sessions consist of designing different parts of a speech recognition application, train the system and evaluate its performance.

The following theoretical components are included:

- algorithms for training, recognition as well as adaptation to properties of speakers and transmissions channel, including pattern recognition, Hidden Markov Models (HMMs) and Deep Neural Networks (DNNs)
- methods to decrease the sensitivity against disturbances and deviations
- probability theory
- signal processing and parameter extraction
- acoustic modelling of the static and dynamic spectral properties of the speech sounds statistical modelling of language in spontaneous and formal speech
- search strategies- basic methods and strategies for large vocabularies
- specific analysis and decision making methods for recognition of speakers.

Furthermore, certain practical insight to build an application is given. Here, implementing certain functions based on prototypes and testing them on real speech data are included.

Disposition

12 lectures, 3 labs, final project

Course literature

Huang, X., Acero, A., Hon, H.-W. Spoken Language Processing - A Guide to Theory, Algorithm and System Development, Prentice Hall, 2001.

Automatic Speech Recognition: A deep learning approach, Dong Yu and Li Deng, Springer 2015. You can download the PDF through KTH Library.

Research articles in speech recognition

Examination

- EXA1 - Exam, 7.5 credits, grading scale: P, F

Based on recommendation from KTH's coordinator for disabilities, the examiner will decide how to adapt an examination for students with documented disability.

The examiner may apply another examination format when re-examining individual students.

If the course is discontinued, students may request to be examined during the following two academic years.

LABl - Computer Lab, 4.5, grading scale: P, F

PROl - Project, 3.0, grading scale: P, F

Other requirements for final grade

Laboratory exercises with oral presentation

Research project with written report

Ethical approach

- All members of a group are responsible for the group's work.
- In any assessment, every student shall honestly disclose any help received and sources used.
- In an oral assessment, every student shall be able to present and answer questions about the entire assignment and solution.