DT2118 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 DT2118 valid from Autumn 2009

Grading scale

P, F

Education cycle

Second cycle

Main field of study

Computer Science and Engineering, Information Technology, Information and Communication Technology

Specific prerequisites

Single course students: 90 university credits including 45 university credits in Mathematics or Information Technology. English B, or equivalent

Language of instruction

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

Intended learning outcomes

After the course, the student will be able to:

- use the described methods to recognise speech and speaker
- design a system for a given application
- adapt and modify existing algorithms for speech and speaker recognition
- evaluate the performance of speech and speaker recognition systems
- pursue research in the domain.

Course contents

The course consists of lectures, one practical laboratory exercise, computational exercises with assignments and a written essay on a chosen title. The essay will be presented orally during a closing seminar.

Included topics:

- algorithms for training, recognition and adaptation to speaker and transmission channel, mainly based on Hidden Markov Models (HMM)
- methods for reducing the sensitivity to external noise and distortion
- probability theory
- signal processing and feature extraction
- acoustic modelling of static and time-varying spectral properties of speech
- statistic modelling of language in spontaneous speech and written text
- search strategies – basic methods and algorithms for large vocabularies
- specific analysis and decision techniques for speaker recognition.

The laboratory exercise is intended to give practical experience of designing a speech recognition application with the software package HTK. It consists of writing of a simple grammar, specification of the acoustic model structure, recording of training and test data, model training, recognition experiments and performance evaluation.

Course literature

- Huang, X., Acero, A., Hon, H.-W. Spoken Language Processing – A Guide to Theory, Algorithm and System Development, Prentice Hall, 2001,
- copies of articles on speaker recognition and other topics not included in the course book.

Examination

- INL1 - Assignment, 3.0 credits, grading scale: P, F
- INL2 - Assignment, 3.0 credits, grading scale: P, F

- LAB1 - Laboratory Work, 1.5 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.

In this course all the regulations of the code of honor at the School of Computer science and Communication apply, see: http://www.kth.se/csc/student/hederskodex/1.17237?l=en_UK.

Other requirements for final grade

Practical laboratory exercise
Computational exercises (assignment)
A written essay with oral presentation in the closing seminar
Reviewing the essays of two other course participants and acting as opponent to their presentations.

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.