TEACHING

Sten Ternström
Director of
undergraduate studies

Undergraduate Level

The undergraduate courses given by the department are all aimed at the last two years of the Master of Science (civilingenjör). All courses belong to the pool of optional courses that undergraduate students choose among for their specialisation. The students that choose the department's courses usually follow the M.Sc. programmes Electrical engineering, Computer science and engineering or Technical physics.

The department has had a very low teaching activity on the undergraduate level but that is changing. Two of the courses are new from this year, namely the courses in *Pattern recognition* and *Source coding theory* (this explains their low number of students). In the spring of 2000, a further course *in Digital speech coding* will start. The department will also be involved in the new M.Sc. programme *Media Technology*. To complete their M.Sc. degree the undergraduates are required to do a thesis that corresponds to five months full time work. During the last year, 17 students chose to do their thesis at the department.

2F1111 Speech technology (4 cr)

14 students

Inger Karlsson

Elements of linguistics and phonetics applied to speech transmission and communication. Speech signal analysis and processing. Acoustic theory of speech production, implementation in speech synthesis. Auditory functions and perception of speech. Automatic speech recognition and text-to-speech conversion. Speech quality and intelligibility assessments. Automatic speaker verification. Text-to-speech and speech-to-text in systems for human-computer interaction, especially multi-modal dialogue systems.

2F1212 Music acoustics (4 cr)

19 students

Erik Jansson

Elements of acoustics including wave propagation, electro-mechanical analogies, strings, membranes, tubes. Fourier transform and elements of auditory theory. Design and function of musical instruments. Musical scales. Music generation and performance rules. Elements of room acoustics. Synthesis of instrument and voice.

2F1400 Electro acoustics (4 cr)

30 students

Johan Liljencrants

General theory of sound waves. Auditory functions and perceptual limitations. Analogies between electrical, mechanical, and acoustic systems. Acoustical impedance. Basic equations for electro-mechanical four-terminal systems. Loudspeakers, microphones. Sound-recording techniques. Ultrasound. Measurement techniques.

2F1510 Pattern recognition (4 cr)

32 students

Arne Leijon

Classical Bayesian decision theory, signal classification in additive white Gaussian noise, matched filter design, parameter estimation and supervised learning, non-parametric classification techniques, linear discriminant functions, and an introduction to neural network design. These concepts and methods are extended to the classification of pattern sequences using the Hidden Markov Model. Human communication by speech and hearing in the framework of signal classification theory.

2F1520 Digital speech compression

4 students

Bastiaan Kleijn

Abstract: The course treats current speech-coding technology by means of laboratory exercises, projects, and lectures. It provides hands-on experience with the application of signal processing methods.

Goal: To provide a good understanding of current speech coding technology and to provide practical experience in signal processing.

Syllabus: Classification and overview of speech coders, overview of speech-coding standards. Uniform, nonuniform, and adaptive quantization. Pulse code modulation and adaptive pulse code modulation. Vocoders which use models of the vocal tract and its excitation. Waveform coding techniques including differential PCM, adaptive DPCM and delta modulation. Analysis-by-synthesis coding methods including multi-pulse LPC, RPLPC, and CELP. Subband and transform coding. Sinusoidal coding and waveform interpolation. Detailed description of several current speech-coding standards including the GSM speech-coding algorithm.

2F1531 Information theory and source coding (5 cr)

13 students

Bastiaan Kleijn

The course treats the principles of the encoding of speech, audio, video, and images at low bit rates. Source coding techniques such as scalar and vector quantization, orthogonal transforms, and linear prediction are introduced and their performance is analyzed theoretically. The theoretical bounds on the performance of source coders are discussed.

2D1574 Media technology, basic course (12 cr)

26 students

Sten Ternström

An introduction to media engineering, including text and print, still photography, moving pictures, and sound. The course provides basic knowledge on human visual and auditory perception, techniques and production methods relevant to each medium, requirements for data rates and volumes, and an ability to communicate effectively with specialists in the respective areas. Several departments cooperate on this course. TMH is coordinator and supplies the sound section.

Professor Rolf Carlson
Director of
graduate studies

Graduate Level

Graduate students comprise about one third of the personnel at the department. Graduate studies towards the Doctor of Science degree require a minimum of four years after the M.Sc. graduation. Since most students are financed by research projects this time is generally exceeded.

The requirements include theoretical studies and a thesis. The thesis may be composed of a number of publications. The theoretical studies are individually tailored within the domain of graduate courses. Requirements include participation in research seminars and attending special lectures which supplement literature assignments. Credits are also given for certain undergraduate courses on top of the undergraduate requirements such as courses in linguistics and phonetics taken at Stockholm University. In addition to the teaching arranged by the staff at the department, special "bullet" courses are organised every year. At such an event a well-known researcher is invited to give a course during a limited period of time, typically a week. Several students at the department have participated in European Summer School and similar summer schools in Europe.

The graduate studies are organised according to two main programmes, with two subtopics each.

1. Speech and Music Communication

The Speech and Music Communication programme includes studies of human communication primarily with the help of acoustic signals such as speech and music. Communication with visual signals such as facial gestures during speech production is also included. The programme contains descriptions, theories, models and applications covering all parts of the communication chain: production - acoustic transmission - perception - understanding or impression.

The programme has two subtopics**: *Speech Communication*** and ***Music Acoustics.***

2. Acoustic Signal Processing

The Acoustic Signal Processing programme covers theory and application in the field of acoustic signal processing, signal coding and information transmission, related to human sound production and signal processing by the human senses.

The programme has two subtopics***: Hearing Technology*** and ***Speech Signal Processing.***

Graduate Level Courses during Year 2000

2F5501 Signal Processing (6 cr)

Bastiaan Kleijn

Hilbert spaces; time-frequency analysis; frame theory; filter banks; wavelets and wavelet packets; best cosine bases; Gabor transforms; Karhunen-Loeve transforms; matching pursuit; thresholding estimators; transform coding. Emphasis on items relevant for the processing of speech.

Reading course in Speaker Characteristics (5 cr)

Mats Blomberg

Selected articles on speaker verification and recognition. The course includes: system overviews and tutorials, probability theory, Hidden Markov Models, Gaussian Mixture Models, the EM training algorithm, maximum likelihood linear regression, Bayesian adaptation, maximum a posteriori adaptation. At the end of the course, each participant writes a short report on a topic given in the course.

Reading course in Speech Recognition (5 cr)

Kjell Elenius

An overview of speech recognition with examples from current techniques. The course includes: acoustic modelling, Hidden Markov Models, Artificial Neural Networks, speech corpora; language models, automata theory, transducers, data representation; decoding, dynamic programming, Viterbi search, A*-search, token-passing and finally some unconventional techniques for speech recognition.

Reading course in Dialog and Discourse (5 cr)

Rolf Carlson

Theoretical foundations of dialog modelling and implications for dialog systems. The course includes selected articles on centering in discourse, speech acts, attention, intention and the structure of dialogue, logic and conversation, lexical choice, representation and use of focus.

Reading course in Dialog systems (5 cr)

Rolf Carlson

Analysis of a number of dialog systems, both research systems and commercial applications, with special focus on the relation between dialog models and application characteristics, error handling, robustness and domain data.

Effective communication - practical rhetoric

Per Nygren

This special course on aural presentation technique was organised together with the department of Signals, Sensors and Systems (S3), KTH. It included both lectures and training sessions.

Bullet course

Julia Hirschberg, AT&T Labs–Research

Intonational variation in spoken dialogue systems: Generation and understanding.

Further Education Courses

2F4212 Acoustics for violin and guitar builders

10 students

Erik Jansson

The course was arranged for the guitarmaking school of Carl Malmsten Center of Wood Technology and Design, professional and amateur violin makers. The lectures consisted of a fundamental theoretical part on sound, hearing, resonators and room acoustics. In the applied instrument acoustics part the played string, the wood, the function and quality measures of guitars and violins were introduced. TeknD Urban Fröjel was invited to present the string maker's view. The course also included laboratory experiments on plate tuning and the function of assembled instruments.

Grunderna i talteknologi för näringslivets intressenter,
Nov 22-24, 2000
Centrum för talteknologi (CTT),
http://www.speech.kth.se/ctt/kortkurs_web.htm

David House: Språk och fonetik

Björn Granström & David House: Multimodal talsyntes

Kåre Sjölander m.fl. Laborationer, talanalys och talsyntes

Kjell Elenius & Mats Blomberg: Taligenkänning

Mats Blomberg: Talarkaraktärisering och talar-verifiering

Rolf Carlson: Dialogsystem

Botond Pakucs m.fl.: Laborationer, dialogsystem

Samtliga föreläsare: Vad händer inom talteknologi i världen, en överblick.

Avslutande diskussion