



EQ2321 Speech and Audio Processing 7.5 credits

Tal- och ljudsignalbehandling

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

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

Establishment

Course syllabus for EQ2321 valid from Autumn 2017

Grading scale

A, B, C, D, E, FX, F

Education cycle

Second cycle

Main field of study

Electrical Engineering

Specific prerequisites

For single course students: 120 credits and documented proficiency in

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 passing the course, students should be able to:

- (1) Qualitatively describe the mechanisms of human speech production and how the articulation mode of different classes of speech sounds determines their acoustic characteristics.
- (2) Apply programming tools (such as Matlab or Python) to analyze speech and audio signals in time and frequency domains, and in terms of the parameters of a source-filter production model and harmonic models.
- (3) Critically analyze, compare and implement methods and systems for coding of speech and audio signals, and finally engineer efficient coding solutions.
- (4) Analyze, compare and implement methods and systems for enhancement of speech and audio signals in environmental noisy conditions.

Course contents

The course considers the foundational and advanced signal and information processing methods for human speech and natural audio signal applications, such as telephone conversation and music playing. For example, what kinds of information from human speech signal need to be extracted and then transmitted through the channel for effective speech communication over phone, and how?

- (1) Preliminaries of associated digital signal processing methodologies, such as convolution, Z-transform, Fourier transform, power spectrum etc.
- (2) A source-filter model: analysis-synthesis architecture.
- (3) Source coding: scalar and vector quantization, redundancy removal, linear prediction, open loop and closed loop coding, coding noise buildup, coding noise shaping, coding gain.
- (4) Speech and audio coding: vocoders, low bit rate and high bit rate codecs, perceptual audio coding, psychoacoustic principles.
- (5) Speech and audio signal enhancement, minimum mean square error estimation, linear estimation for Gaussian distribution, Wiener filtering, power spectral subtraction methods, spectral band replication, etc

Disposition

How the course will be conducted:

- (1) The course is for one period (typically 8 weeks study).

(2) Preliminary, there will be in total 24 classes. Each class is of around 1.5 hours.
Teaching class: 14 and Tutorial class: 10.

Course literature

Will be announced on the course homepage before course start. Preliminary literature:

- (1) Digital speech transmission: Enhancement, coding and error concealment. By Peter Vary and Rainer Martin.
- (2) Perceptual coding of digital audio. By Ted Painter and Andreas Spanias.
- (3) Notes of the class teacher. This can be downloaded from the course website.
- (4) Some research papers.

Examination

- PRO1 - Project 1, 1.5 credits, grading scale: A, B, C, D, E, FX, F
- PRO2 - Project 2, 1.5 credits, grading scale: A, B, C, D, E, FX, F
- TEN1 - Exam, 4.5 credits, grading scale: A, B, C, D, E, FX, 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.

Other requirements for final grade

There are three assessment components for the course.

- (1) Master tests: There will be two master tests in the span of teaching 14 classes. Each test is of 20-30 minutes. The master tests are intended to check concepts and require sustained (or regular) study at home as the teachers cover topics in class. The tests will use short conceptual questions, and no lengthy problem. Grades for master tests: A-F.
- (2) Projects: There are two projects. Projects are examined via presentations. Projects can be performed in groups of two persons. However, the grades are on the basis of individual performance. Grades for projects: A-F.
- (3) Written exam: There is a final written exam. Grades for the final exam: A-F.

The overall grade of the course is based on collective performance. The teacher will provide weights to all tests for the overall grade.

To pass the course, master tests are not mandatory. But the projects and final test are mandatory. To achieve a good course grade, a student is expected to perform well in all the three assessment components.

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.