UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Last update: 04-07-2019

230622 - DSAP - Digital Speech and Audio Processing

Coordinating unit: 230 - ETSETB - Barcelona School of Telecommunications Engineering

Teaching unit: 739 - TSC - Department of Signal Theory and Communications

Academic year: 2019

Degree: MASTER'S DEGREE IN INFORMATION AND COMMUNICATION TECHNOLOGIES (Syllabus 2009). (Teaching unit Optional)
MASTER'S DEGREE IN TELECOMMUNICATIONS ENGINEERING (Syllabus 2013). (Teaching unit Optional)
MASTER'S DEGREE IN ADVANCED TELECOMMUNICATION TECHNOLOGIES (Syllabus 2019). (Teaching unit Optional)

ECTS credits: 5 Teaching languages: English

Teaching staff

Coordinator: Climent Nadeu

Opening hours

Timetable: Tuesday and Thursday from 10:00 to 13:00

Prior skills

Advanced knowledge of Signals, Systems, and Transforms
Basic knowledge of Probability, Random variables and Stochastic processes
Experience with Matlab programming
Recommended:
- Basic knowledge of Machine Learning
- Python language

Requirements

Two courses of the area Signals, Systems, and Transforms
At least one course about Probability, Random variables and Stochastic processes

Degree competences to which the subject contributes

Specific:

1. Ability to apply information theory methods, adaptive modulation and channel coding, as well as advanced techniques of digital signal processing to communication and audiovisual systems.

Transversal:

2. TEAMWORK: Being able to work in an interdisciplinary team, whether as a member or as a leader, with the aim of contributing to projects pragmatically and responsibly and making commitments in view of the resources that are available.
3. EFFECTIVE USE OF INFORMATION RESOURCES: Managing the acquisition, structuring, analysis and display of data and information in the chosen area of specialisation and critically assessing the results obtained.
4. FOREIGN LANGUAGE: Achieving a level of spoken and written proficiency in a foreign language, preferably English, that meets the needs of the profession and the labour market.

230622 - DSAP - Digital Speech and Audio Processing

Teaching methodology

- Lectures (50%)
- Application classes (with Matlab or similar) (50%)
- Team work: course project, application classes
- Individual work: preparation and completion (out classroom) of application activities

Learning objectives of the subject

Learning objectives of the subject
Understanding and being competent on a relevant set of concepts and techniques in the field of digital audio processing, and their application to problems arising from real applications. Signals and applications related to speech and music will be particularly considered.

Learning results:
Ability to digitally process audio signals, in an application-oriented context, in order to analyze, model, extract information, clean, modify, and generate/synthesize them.

Study load

Total learning time: 125h	Hours large group:	39h	31.20%
	Hours medium group:	0h	0.00%
	Hours small group:	0h	0.00%
	Guided activities:	0h	0.00%
	Self study:	86h	68.80%

Content

Introduction

Learning time: 12h

Theory classes: 6h
Self study : 6h

Description:
Course presentation
Audio diversity
Characteristics of speech and music. Production model
Hearing and auditory modeling
Short-time Fourier transform (STFT) and spectrogram
The short-time Fourier transform

Modeling and representation of audio signals

Learning time: 12h

Theory classes: 6h
Self study : 6h

Description:
Production-based all-pole modeling
Pitch determination for speech and music
LPC-based coding used in mobile telephony

Enhancement of voice and audio signals

Learning time: 12h

Theory classes: 6h
Self study : 6h

Description:
Denoising: spectral subtraction, Wiener-based filtering, neural nets
Blind source separation: ICA, NMF
Cancellation: echo, interference

Recognition and detection of audio and speech	Learning time: 12h Theory classes: 6h Self study : 6h

Description:
Pattern-matching dynamic approaches.
Statistical and deep learning approaches
Audio activity detection
Application to speech and sound recognition

Multi-microphone audio processing	Learning time: 12h Theory classes: 6h Self study : 6h

Description:
Room acoustics
Array beamforming
Acoustic source localization and tracking

Specific objectives:

Analysis and synthesis of audio signals	Learning time: 12h Theory classes: 6h Self study : 6h

Description:
Short-term analysis-synthesis of (cuasi)periodic signals. Time-scale and pitch modification
Spatial audio synthesis with HRTF functions

Project realization and presentation	Learning time: 55h Theory classes: 4h Self study : 51h

Description:
Design, implementation and test of a audio processing system for a specific application
Oral presentation of 1) project proposal, 2) critical review, and 3) project realization and conclusions
Written report, conference paper style

Planning of activities

Practical exercises with Matlab or similar, about 50% of classroom time	Hours: 20h Theory classes: 20h

Description:
Most weeks the students work for almost 2 hours with a Matlab file and data provided by the teacher. They also do some complementary work at home.

Support materials:
Theory slides, Matlab code, data (signals, etc)

Descriptions of the assignments due and their relation to the assessment:
Learning of topics is assessed with bi-weekly questionnaires

Short tests at the end of each topic	Hours: 1h Theory classes: 1h

Course project	Hours: 60h Theory classes: 60h

Description:
Team project realization, which includes audio processing experimental work, and is presented both orally and in writing.

Oral presentations	Hours: 0h 40m Theory classes: 0h 40m

Description:
Oral presentation of the course project in three times: proposal, review, and final presentation.

230622 - DSAP - Digital Speech and Audio Processing

Qualification system

Attendance/participation in class (10%)
Short tests every two weeks (30%)
Course project (60%)

Bibliography

Basic:

Quatieri, T.F. Discrete-time speech signal processing: principles and practice. Upper Saddle River, NJ: Prentice Hall, 2002. ISBN 013242942X.

Gold, B.; Morgan, N.; Ellis, D. Speech and audio signal processing: processing and perception of speech and music. 2nd rev. ed. Wiley-Blackwell, 2011. ISBN 978-0-470-19536-9.

Dutoit, T.; Marqués, F.; Rabiner, L.R. Applied signal processing: a MATLAB-based proof of concept. New York ; London: Springer, 2009. ISBN 978-0-38774534-3.

Complementary:

Rabiner, L.R.; Schafer, R.W. Theory and applications of digital speech processing. Prentice Hall, 2010. ISBN 9780136034285.

Huang, Y.A.; Benesty, J. (eds.). Audio signal processing for next-generation multimedia communication systems [on line]. New York: Kluwer Academic Publishing, 2004 [Consultation: 23/07/2013]. Available on: <http://link.springer.com/book/10.1007/b117685/page/1>. ISBN 1402077688.

Others resources:

Lecture slides
Practical work statements and programs

Audiovisual material

Slides

Slides used in lectures

Computer material

Codi programes

Software codes in Matlab or similar