Course guides
230622 - DSAP - Digital Speech and Audio Processing

Unit in charge: Barcelona School of Telecommunications Engineering
Teaching unit: 739 - TSC - Department of Signal Theory and Communications.

Degree: MASTER'S DEGREE IN TELECOMMUNICATIONS ENGINEERING (Syllabus 2013). (Optional subject).
MASTER'S DEGREE IN ADVANCED TELECOMMUNICATION TECHNOLOGIES (Syllabus 2019). (Optional subject).

Academic year: 2021 ECTS Credits: 5.0 Languages: English

LECTURER
Coordinating lecturer: Climent Nadeu

Others:

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
At least 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.
TEACHING METHODOLOGY

- Lectures (50%)
- Application classes (with Matlab or similar) (50%)
- Team work: course project, application classes
- Individual work: homework assignments, related to the applications

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

<table>
<thead>
<tr>
<th>Type</th>
<th>Hours</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self study</td>
<td>86,0</td>
<td>68.80</td>
</tr>
<tr>
<td>Hours large group</td>
<td>39,0</td>
<td>31.20</td>
</tr>
</tbody>
</table>

Total learning time: 125 h

CONTENTS

Introduction

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

Full-or-part-time: 12h
Theory classes: 6h
Self study: 6h

Modeling and representation of audio signals

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

Full-or-part-time: 8h
Theory classes: 4h
Self study: 4h
<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
<th>Full-or-part-time: 16h</th>
<th>Theory classes: 8h</th>
<th>Self study : 8h</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Enhancement of voice and audio signals</strong></td>
<td>Description: Denoising: spectral subtraction, Wiener-based filtering, neural nets (deep learning) Blind source separation: NMF Cancellation: echo, interference</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Recognition and detection of audio and speech</strong></td>
<td>Description: Pattern-matching dynamic approaches. Statistical and deep learning approaches Approaches based in dynamic pattern matching Statistical and deep learning approaches Audio activity detection Application to speech and sound recognition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Multi-microphone audio processing</strong></td>
<td>Description: Room acoustics Array beamforming Acoustic source localization and tracking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Analysis and synthesis of audio signals</strong></td>
<td>Description: Short-term analysis-synthesis of (cuasi)periodic signals. Time-scale and pitch modification Spatial audio synthesis with HRTF functions</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Project realization and presentation

**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

**Full-or-part-time:** 48h
Theory classes: 4h
Self study: 44h

### ACTIVITIES

#### Assignments and short tests

**Full-or-part-time:** 25h
Self study: 25h

#### Course project

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

**Full-or-part-time:** 60h
Self study: 60h

#### Oral presentations

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

**Full-or-part-time:** 1h
Self study: 1h

### GRADING SYSTEM

Attendance/participation in class (10%)
Assignments and tests (30%)
Course project (60%)

### BIBLIOGRAPHY

**Basic:**
Complementary:

RESOURCES

Audiovisual material:
- Slides. Slides used in lectures

Computer material:
- Codi programes. Software codes in Matlab or similar

Other resources:
Lecture slides
Practical work statements and programs