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 ADVANCED TELECOMMUNICATION TECHNOLOGIES (Syllabus 2019).
(Teaching unit Optional)
MASTER'S DEGREE IN TELECOMMUNICATIONS ENGINEERING (Syllabus 2013). (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

<table>
<thead>
<tr>
<th>Total learning time: 125h</th>
<th>Hours large group: 39h</th>
<th>31.20%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hours medium group: 0h</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>Hours small group: 0h</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>Guided activities: 0h</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>Self study: 86h</td>
<td>68.80%</td>
</tr>
</tbody>
</table>
# Content

## 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

**Learning 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

**Learning time:** 12h
- Theory classes: 6h
- Self study: 6h

## Enhancement of voice and audio signals

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

**Learning time:** 12h
- Theory classes: 6h
- Self study: 6h
<table>
<thead>
<tr>
<th><strong>Recognition and detection of audio and speech</strong></th>
<th><strong>Learning time:</strong> 12h</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description:</strong></td>
<td>Theory classes: 6h</td>
</tr>
<tr>
<td>Pattern-matching dynamic approaches.</td>
<td>Self study : 6h</td>
</tr>
<tr>
<td>Statistical and deep learning approaches</td>
<td></td>
</tr>
<tr>
<td>Audio activity detection</td>
<td></td>
</tr>
<tr>
<td>Application to speech and sound recognition</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Multi-microphone audio processing</strong></th>
<th><strong>Learning time:</strong> 12h</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description:</strong></td>
<td>Theory classes: 6h</td>
</tr>
<tr>
<td>Room acoustics</td>
<td>Self study : 6h</td>
</tr>
<tr>
<td>Array beamforming</td>
<td></td>
</tr>
<tr>
<td>Acoustic source localization and tracking</td>
<td></td>
</tr>
<tr>
<td><strong>Specific objectives:</strong></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Analysis and synthesis of audio signals</strong></th>
<th><strong>Learning time:</strong> 12h</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description:</strong></td>
<td>Theory classes: 6h</td>
</tr>
<tr>
<td>Short-term analysis-synthesis of (quasi)periodic signals. Time-scale and pitch modification</td>
<td>Self study : 6h</td>
</tr>
<tr>
<td>Spatial audio synthesis with HRTF functions</td>
<td></td>
</tr>
</tbody>
</table>
## Project realization and presentation

**Learning time:** 55h  
Theory classes: 4h  
Self study: 51h

**Description:**  
Design, implementation and test of an 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 |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description:</strong></td>
<td>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.</td>
</tr>
<tr>
<td><strong>Support materials:</strong></td>
<td>Theory slides, Matlab code, data (signals, etc)</td>
</tr>
<tr>
<td><strong>Descriptions of the assignments due and their relation to the assessment:</strong></td>
<td>Learning of topics is assessed with bi-weekly questionnaires</td>
</tr>
</tbody>
</table>

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

| Course project | Hours: 60h  
Theory classes: 60h |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description:</strong></td>
<td>Team project realization, which includes audio processing experimental work, and is presented both orally and in writing.</td>
</tr>
</tbody>
</table>

| Oral presentations | Hours: 0h 40m  
Theory classes: 0h 40m |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description:</strong></td>
<td>Oral presentation of the course project in three times: proposal, review, and final presentation.</td>
</tr>
</tbody>
</table>
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:


Complementary:


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