COURSE DATA

Data Subject	
Code	34907
Name	Digital speech and audio processing
Cycle	Grade
ECTS Credits	6.0
Academic year	2023 - 2024

Study (s)

Degree	Center	Acad. year	Period
1403 - Degree in Telematics Engineering	School of Engineering	4	Second term

Subject-matter

Degree	Subject-matter	Character
1403 - Degree in Telematics Engineering	19 - Optional subjects	Optional

Coordination

Name	Department
COBOS SERRANO, MAXIMO	240 - Computer Science

SUMMARY

The course "Digital Audio and Speech Processing" is a fourth year course that is part of the elective offer of the Degree in Telematics Engineering. The course complements the contents seen in other subjects of the Degree such as "Digital Signal Processing", "Linear Systems and Signals " and "Mathematical Fundamentals of Communications", offering an applied vision of the concepts studied throughout these subjects. Thus, the topics covered by this course are oriented to the application of digital signal processing and machine learning in the field of audio and speech processing.

The course justifies the importance of digital audio signal processing in current multimedia systems, briefly reviewing some basic concepts studied in previous courses. The structuring of the contents follows an approach that begins by introducing the more theoretical concepts in the first part of the course related to digital signal processing, pacing the way to other topics more focused on machine learning, which have become very important nowadays. Thus, we start with a review of basic concepts of signal processing, introducing several representations of audio signals, paying special attention to time-frequency representations. The theoretical part is complemented with the study of the human auditory system and its impact on the design of practical lossy audio coding systems, the fundamentals of the speech

production system and the source/filter models used in speech coding.

The second part of the course introduces the fundamentals of machine learning based audio systems, introducing the use of features widely used in audio problems and their combination with classical binary classifiers such as support vector machines or logistic regression. Finally, the student is introduced to current deep learning systems.

PREVIOUS KNOWLEDGE

Relationship to other subjects of the same degree

There are no specified enrollment restrictions with other subjects of the curriculum.

Other requirements

OUTCOMES

LEARNING OUTCOMES

This course allows for the following key learning outcomes. Note that since this course is optional, there are not specific competencies linked to these learning outcomes.

-Acquire new knowledge and techniques suitable for the design, development and implementation of telecommunication systems, especially those related to multimedia communications.

-Follow the design process of systems based on auditory perception, relating the aspects involved in it with the associated mathematical theory.

In addition to the above results, this subject also allows to acquire the following technical and social skills:

-Apply the knowledge on auditory perception to the design of signal processing systems.

-Identify the limitations in the analog / digital conversion of audio signals.

-Select properly the parameters needed for the encoding and storage of audio and speech signals.

-Identify the models that guide the design of existing telecommunication systems.

-Implement basic signal processing systems oriented to audio processing.

-Adequately describe the principles ofaudio signal compression, relating them to the corresponding physical and mathematical theories.

-Promote teamwork and organization into tasks and subtasks.

DESCRIPTION OF CONTENTS

0. Introduction

Digital audio, image and video processing overview. Signal processing in multimedia.

1. Signal processing review

Descripción de contenidos (English):
Introduction. Review of convolution and filtering. Sampling theorem. Discrete-time signal processing. DFT and DTFT. Random processess.

2. Audio and psychoacoustic principles

Descripción de contenidos (English):
Introduction. Fundamentals of Acoustics. Hearing and perception. Loudness and critical bands.

3. Introduction to machine learning in audio

Signal processing and artificial intelligence in audio. Evolution and history of machine learning based audio systems. Challenges related to the use of machine learning techniques in audio. Current applications.

4. Audio representations for machine learning

Signal processing and artificial intelligence in audio. Evolution and history of machine learning based audio systems. Challenges related to the use of machine learning techniques in audio. Current applications.

5. Machine learning techniques for audio and speech applications

Introduction to supervised and unsupervised learning. Linear binary classifiers: SVM, Logistic regression. Nonlinear classifiers: k-NN, Kernel models. Gaussian Mixture Models (GMM), k-Means. Introduction to neural networks and deep learning.

WORKLOAD

ACTIVITY	Hours	% To be attended
Theory classes	30,00	100
Laboratory practices	20,00	100
Classroom practices	10,00	100
Development of group work	20,00	0
Development of individual work	10,00	0
Preparation of evaluation activities	20,00	0
Preparing lectures	10,00	0
Preparation of practical classes and problem	15,00	0
Resolution of case studies	15,00	0
TOTAL	**150,00**	

TEACHING METHODOLOGY

1.- Classroom work consisting of:

a) Theory classes, which will consist of the presentation and basic explanation of the corresponding subject. Short activities will be proposed, which will require the intervention of the students with the objective of confirming the understanding of the theory exposed.

b) Exercise and seminar classes, designed to solve problems of greater temporal or conceptual scope.

c) Laboratory classes, designed to experimentally test some of the most relevant issues seen in the theory classes.

2.- Non-attendance work consisting of:

(a) Preparation of subject questionnaires to be done after finishing each didactic unit.

b) Preparation of the laboratory practices, for which the student must have read and assimilated the content of the practice bulletin, as well as having reviewed the relevant theory.

Consulting sessions: A certain number of hours are established each week, which the students can attend in order to solve doubts

EVALUATION

The evaluation mechanism is based on a continuous evaluation of the work developed by the student throughout the course. The following items and evaluations are taken into account:

a) - Result of topic questionnaires carried out after the completion of each thematic unit (50% of the final grade)

b) - Preparation of practical reports (25% of the final grade)

c) - Completion of the global final questionnaire (25% of the final grade)

To pass the subject, an average grade higher than 5 is required. The minimum required to pass the course will be the equivalent of 5 out of 10 both in the average grade obtained in the topic questionnaires (a), and in the final questionnaire (c).

The score obtained in the labs (b) will not be recoverable on second call.

REFERENCES

Basic

- Virtanen, Plumbley, Ellis, (eds): Computational Analysis of Sound Scenes and Events, Springer, 2018.

- Giannakopoulos, T. And Pikrakis, A. Introduction to Audio Analysis, Elsevier, 2014.
- Zölzer, Udo., Digital Audio Signal Processing, 2nd edition, Wiley, 2008. ISBN: 0470997850
- Pulkki, V., Karjalainen, M. Communication Acoustics: An Introduction to Speech, Audio and Psychoacoustics. Wiley (2015). ISBN: 978-1-118-86654-2

Additional

- Aurélien Géron, Hands-on Machine Learning with Scikit-Learn, Keras & Tensorflow, OReilly, 2020.
- -Smith III, Julius O., Spectral Audio Signal Processing, W3K Publishing, 2011. 978-0974560731
- -Bosi, M. and Goldberg, Richard E., Introduction to Digital Audio Coding and Standards, Kluwer Academic Publishers, 2003. ISBN: 978-1402073571