Teaching Unit title : Advanced speech and audio signal processing

Number of ECTS credits	3 ECTS (in addition to 3 ECTS from image processing)
Semester	M2 S1

a) Educational objectives

This teaching unit aims at introducing advanced methods and techniques in audio signal processing, with a focus on applications to speech signals used in many human-to-machine interaction systems. The main objective is to make the students able to combine skills in acoustics, signal processing and machine learning to solve complex tasks in audio processing, and more precisely in automatic speech processing and acoustic/auditory scene analysis.

b) Skills acquired by the student

At the end of this teaching unit, the students will be capable of:

- Analyze and characterize speech signal;
- Implement and calibrate some automatic speech processing techniques, or exploit audio signals to solve some tasks of localization, separation, recognition, synthesis or authentication.

c) Program

Acoustics/signal block

- C1 (2h) + C2 (1h) Reminders on acoustics and sound propagation: derivation of the wave equation in three space dimensions, solution in 1 space dimension without source: plane wave, and solution in 3 space dimensions in the presence of any localised source, using a Green's function.
- C2 (1h) + C3 (2h) Sound source localization through a microphone array. From acoustics to array modelling; beamforming; TDOA estimation through generalized cross-correlation (GCC).

Signal/data block

- C4 (2h) Sound source separation: monaural, binaural et multi-channel algorithms for sound source separation (NMF).
- C5 (2h) Automatic speech processing: introduction to speech communication and general linguistics, speech-to-text and text-to-speech (history, algorithms and trends).
- C6 (2h) Music information retrieval. Introduction to music representations. Application to pitch estimation and cover detection using triplet loss. Source separation for music using convolutional neural networks.

- **Student project (16h)**. The students have to exploit their knowledge and competencies in some projects focused on specific tasks, like sound localization, separation, etc. The evaluation of the project is composed of two marks : one on the report and the other on an oral presentation.

d) Prerequisites

Mathematical tools for analog signal processing: Fourier series, Fourier transform, Laplace transform, differential equations, computer programming in Python. Knowledge of acoustics is a plus.

e) Terms and conditions of assessment

2 written tests + 1 project (report + oral presentation).

Educational organisation

Classroom sessions	Total timetable	Student Groups
Lectures	6 2-hour classes	Total
Tutorials		
Project	4 4-hour classes	
Other	1 quizz + 1 written test + project presentation by groups of 4 students (20 mn)	

Cadencement temporel de l'UE :

N°	Lectures	Tutorials	Project	Assessment
S1	C1 (HB)			
S2	C2 (HB & SA)			
S3	C3 (SA)			
S4				CC1 (quizz)
S5	C4 (NO)			
S6	C5 (NO)			
S7	C6 (ACH)		Project1	
S8			Project 2	
S9			Project 3	
S10			Project 4	
X				CC2 (written test)
X				