



Advanced Speech and Audio Processing (MU5RBI20)

M2 Engineering of Intelligent Systems
& Advanced Systems and Robotics

2023-2024,

H. Boutin, S. Argentieri, N. Obin, A. Cohen-Hadria

Education team



Henri Boutin
C1(2h) + C2(1h)
*Basics on acoustics
and sound
propagation*



Sylvain Argentieri
C2 (1h) + C3 (2h)
*Sound source
localization
through a
microphone array*



Nicolas Obin
C4 (2h) – *Source
separation*
C5 (2h) -
*Automatic speech
processing.*



**Alice Cohen-
Hadria**
C6 (2h) – *Music
information retrieval
(resp.)*

Objectives

- The aim of this course are:
 - To provide students with advanced methods and technics in audio signal processing, with a strong focus on speech processing;
 - To address complex issues in the field of automatic speech processing, scene analysis and soundscapes, using basic skills in acoustics, signal processing and machine learning.
- Skills to be acquired:
 - Speech analysis;
 - Implementation and calibration of systems of Automatic Speech Processing for speech localization, separation, recognition, analysis, synthesis and authentication;
 - Audio and speech recordings.

Prerequisites

- This course is:
 - part of the ISI M.Sc. **Engineering in Smart Systems** from Sorbonne Université, Paris;
 - a logical follow-up to the introductory course of 1st year M.Sc.: **Introduction to Audio Signal Processing**.
- Prerequisites:
 - **Maths for analog signal processing:** Fourier series, Fourier transforms (DSFT, DFT, STFT), Laplace transform, differential equations;
 - Programming in **Python**;

Course syllabus (1)

Lectures (12h = 6×2h):

- **Acoustics / Signal**
 - 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**
 - 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.

Evaluation: 2 assessment tests : 1 quizz + 1 written exam ⇒ 2 marks

Course syllabus (2)

Practical works (16h = 4×4h):

- Purpose: The students have to exploit their knowledge and competencies in some projects focused on specific tasks, like sound source localization, separation, etc.
- Organization: 3 groups of ~12 students, 1 expert teacher per group
 - TP1 (H. Boutin) : Basics of acoustics and sound propagation
 - TP2 (S. Argentieri) : Sound source localization and beamforming
 - TP3&4 (A. Cohen-Hadria & N. Obin) : Sound source separation, Music information retrieval, Automatic speech processing, Text-to-speech.
- Evaluation: 1 report per session

Course syllabus (3)

Content = **28h**: lectures = **12h** (**6 × 2h**) + practical works= **16h** (**4 × 4h**)
+ 2 assessment tests

Week No	Tests	Lectures	Practical works
1		C1: 18/9: 1:45-3:45 pm	
2		C2: 25/9: 1:45-3:45 pm	
		C3: 26/9: 1:45-3:45 pm	
4	ER1:? 9/10 10:45am	C4: 9/10: 1:45-3:45 pm	
5		C5: 16/10: 1:45-3:45 pm	
6		C6: 23/10: 8:30-10:30 am	
9			P1: <u>G1:</u> 13/11: 1:45-6pm, <u>G2:</u> 14/11: 1:45-6pm
10			P2: <u>G1:</u> 20/11: 1:45-6pm, <u>G2:</u> 22/11: 1:45:6pm
11			P3: <u>G1:</u> 27/11: 1:45-6pm, <u>G2:</u> 29/11: 8:30-12:45pm
13	ER2: 13/12: 1:45pm		P4: <u>G1:</u> 13/12: 8:30-12:45pm, <u>G2:</u> 14/12: 8:30-12:45pm

Resources

Resources and information on Moodle:

<https://moodle-sciences-23.sorbonne-universite.fr/course/view.php?id=2135>

+ *at the end of presentations...*

Provisional planning

<https://sciences.sorbonne-universite.fr/masters/master-automatique-robotique/parcours-ingenieure-des-systemes-intelligents-isi>



Advanced Speech and Audio Processing (MU5RBI20)

M2 Engineering of Intelligent Systems
& Advanced Systems and Robotics

2023-2024,

H. Boutin, S. Argentieri, N. Obin, A. Cohen-Hadria