2EL5060 – Analysis and processing of audio data (speech and music)

Instructors: Stephane Rossignol
Department: CAMPUS DE METZ
Language of instruction: ANGLAIS
Campus: CAMPUS DE METZ
Workload (HEE): 60
On-site hours (HPE): 35,00
Elective Category : Fundamental Sciences
Advanced level : Yes

Description

The aim of this course is to present the corpus of non-parametric and parametric spectral analysis methods, as part of the analysis of musical and speech sound signals.

Spectral analysis is one of the elements of signal processing chains; therefore, it is not just the visualization of spectra. The main objective is to decide and/or estimate. Some examples: What was the original score? Or what instruments are present in the orchestra? What is the fundamental frequency of this or that sound? What does this person say? Where is this person or this other one? Etc. The choice of this or that spectral analysis method is crucial, depending on the current problem.

The focus is put on the different concepts underlying each method, and the performance of these methods are compared. This last approach also highlights the concept of modeling (physical modeling/signal modeling/...), inherent in an efficient engineering approach.

Moreover, the various tools are studied in the context of the observation of the human being, who communicates with his fellow beings and his environment through his senses. Of these, sight and hearing are the best known, and only they allow a remote approach to the environment. Communication systems (human-to-human, human-to-machine or machine-to-human) are designed to acquire and reproduce these perceptions as faithfully as possible. It is therefore useful to know and be able to model in detail on the one hand the system of human perception, that is to say the receiver (the ear, here), but also the system of production of the signal concerned by the perception (human speech, music, various sounds), that is, the transmitter.

The sound platforms of the Metz campus (holophony room and anechoic room) are used for the practical aspects of this course.

This course has many areas of application: in-depth analyzes of time series; sound analysis for recognition and coding; phonetic coders for telephony; tools for the arts.

Quarter number
SG6

Prerequisites (in terms of CS courses)
Signal processing (1CC400).
Statistics and machine learning (1CC5000)
Programming experience (1CC1000)

Syllabus

- 1. Non-parametric spectral analysis methods (mainly based on the Fourier transform)
    - 1.1. Further information on these methods, already known
    - 1.2. Review of their limits
    - 1.3. Statistical means used to make the most of these methods
- 2. Parametric spectral analysis methods
    a. 2.1. Introduction to some of them
    b. 2.2. Contributions, compared to non-parametric methods
    c. 2.3. Cost of parametric methods
- 3. The sounds
    - 3.1. Models of perception (ear) and production (voice, mainly)
    - 3.2. Sound localization
    - 3.3. Virtualization of sound sources (holophony)

Class components (lecture, labs, etc.)

18h Lecture

9h Tutorials. (3 hours for each part of the lecture; sequencing Lecture/Tutorials: 6h L; 3h T; 6h L; 3h T; 6h L; 3h T)

8h Labs. A single topic.

Grading

Continuous monitoring (50%, 2/3 MCT at the beginning of the tutorials; individual score) and oral presentation at the very end of the labs (50%). Labs : grading by pair; differentiated in the event of an anomaly in a pair.

Course support, bibliography
Slides.

Resources

- Teacher : Stéphane Rossignol
- Room size for tutorials : 34
- Max room size for labs : 34
- Software : Matlab (34 licences)/Octave (Python)
- Rooms for labs : rooms on Metz campus

Learning outcomes covered on the course

- Design a complete signal processing chain.
- Compare the performances of the various tools at our disposal for the analysis of complicated time series, in order to choose the one which will be best suited for this or that signal to be analyzed.
- Program in an interpreted computer language (matlab/octave/python/...).
- Mastering the basic and advanced principles of analog signal processing and digital signal processing.
- Mastering the basic principles of sound perception (cognitive perception).

Description of the skills acquired at the end of the course

- C1 : Analyze, design, and build complex systems with scientific, technological, human, and economic components
- C2 : Develop in-depth skills in an engineering field and a family of professions
- C6 : Be operational, responsible, and innovative in the digital world