

2EL5060 - Analyse et traitement de données audio (parole et musique)

Responsables : **Stephane ROSSIGNOL**

Langues d'enseignement : **ANGLAIS**

Type de cours : **Electif 2A**

Campus où le cours est proposé : **CAMPUS DE METZ**

Nombre d'heures d'études élèves (HEE) : **60**

Nombre d'heures présentielles d'enseignement (HPE) : **30**

Année académique : **2024-2025**

Catégorie d'électif : **Sciences fondamentales**

Niveau avancé : **oui**

Présentation, objectifs généraux du cours :

Le but de cet enseignement est d'exposer le corpus des méthodes d'analyse spectrale non-paramétriques et paramétriques, ceci dans le cadre de l'analyse des signaux sonores musicaux et de parole.

L'analyse spectrale est l'un des éléments des chaînes de traitement du signal ; de ce fait elle n'est pas que la simple visualisation des spectres. L'objectif principal est de décider et/ou d'estimer. Quelques exemples : Quelle était la partition originale ? Ou encore : quels instruments sont présents dans l'orchestre ? Quelle est la fréquence fondamentale de tel ou tel son ? Que dit telle personne ? Où se trouve telle autre personne ? Etc. Le choix de telle ou telle méthode d'analyse spectrale est crucial, selon le problème traité.

L'accent est mis sur les différents concepts sous-jacents à chacune des méthodes et leurs performances sont comparées. Cette dernière approche met aussi en relief le concept de modélisation (physique/signal/...), inhérent à une démarche d'ingénierie efficace.

De plus, les outils étudiés le sont dans le cadre de l'observation de l'être humain, qui communique avec ses semblables et son environnement grâce à ses sens. Parmi ceux-ci, la vue et l'ouïe sont les mieux connus, et seuls eux autorisent une approche à distance de l'environnement. Les systèmes de communication (humain à humain, humain à machine ou machine à humain) sont conçus afin d'acquérir et de restituer le plus fidèlement possible ces perceptions. Il est donc utile de connaître et de pouvoir modéliser en détails d'une part le système de perception humain, c'est-à-dire le récepteur (l'oreille, ici), mais également le système de production du signal concerné par la perception (parole humaine, musique, sons divers), c'est-à-dire l'émetteur.

Les plateformes sonores du campus de Metz (salle d'holophonie et chambre sourde) sont mises à contribution pour les aspects pratiques de ce cours.

Ce cours compte de nombreux domaines d'application : analyses approfondies des séries temporelles ; analyse des sons pour la reconnaissance et le codage ; codeurs phonétiques pour la téléphonie ; outils pour les arts.

Période(s) du cours (n° de séquence ou hors séquence) :

SG6

Prérequis :

- Traitement du Signal : 1CC4000
- Statistique et Apprentissage : 1CC5000
- Expérience de programmation :
- Système d'Information et Programmation : 1CC1000

Plan détaillé du cours (contenu) :

- 1. Première partie : les méthodes d'analyse spectrale non-paramétriques (basées principalement sur la transformée de Fourier)
 - 1.1. Approfondissement concernant ces méthodes, déjà vues
 - 1.2. Revue de leurs limites
 - 1.3. Moyens statistiques utilisés pour tirer le maximum de ces méthodes
- 2. Deuxième partie : les méthodes d'analyse spectrale paramétriques
 - 2.1. Introduction à certaines d'entre elles
 - 2.2. Apports par rapport aux méthodes non-paramétriques
 - 2.3. Coût de ces méthodes
- 3. Troisième partie : les sons
 - 3.1. Modèles de perception (l'oreille) et de production (la voix, principalement)
 - 3.2. Localisation des sons
 - 3.3. Virtualisation des sources sonores (holophonie)

Déroulement, organisation du cours :

17.5h Cours magistral

9h Travaux dirigés (3 heures par partie du cours ; séquençement Cours/TD : 6h C; 3h TD; 6h C; 3h TD; 6h C; 3h TD)

8h Travaux pratiques (sur machines). Un sujet courant sur la séquence.

Organisation de l'évaluation :

Contrôle continu (50% de la note, sous forme de 2/3 QCM en début de TD ; note individuelle) et exposé oral à la toute fin du long TP (50% de la note). TP : note par binôme ; différenciée en cas d'anomalie dans un binôme.

Moyens :

- Equipe enseignante : Stéphane Rossignol
- Taille des TD : 34
- Taille max des TP : 34
- Outils logiciels et nombre de licence nécessaire : Matlab (34 licences)/Octave (Python)

- Outils logiciels et nombre de licence nécessaire : Matlab (34 licences)/Octave (Python)
- Salles de TP (département et capacité d'accueil) : salles sur le campus de Metz

Acquis d'apprentissage visés dans le cours :

- Concevoir une chaîne de traitement du signal complète.
- Comparer les performances des divers outils à notre disposition pour l'analyse des séries temporelles compliquées, afin de choisir celui qui conviendra le mieux pour tel ou tel signal à analyser.
- Programmer dans un langage interprété (matlab/octave/python/...).
- Connaître les principes de base et approfondis du traitement du signal analogique et du signal numérique.
- Connaître les principes de base de la perception des sons (perception cognitive).

Description des compétences acquises à l'issue du cours :

- 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

Bibliographie :

Transparents.