

3MD4020 - Traitement du son (parole et musique)

Responsables : **Stephane ROSSIGNOL**

Langues d'enseignement : **FRANCAIS**

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

Nombre d'heures d'études élèves (HEE) : **35**

Nombre d'heures présentielles d'enseignement (HPE) : **18**

Année académique : **2024-2025**

Niveau avancé : **non**

Présentation, objectifs généraux du cours :

Le but de cet enseignement est d'exposer le corpus des méthodes d'analyse des signaux sonores musicaux et de parole. L'analyse, la modélisation, la synthèse, le codage et la reconnaissance de la parole sont traités. Des mesures, annexes mais nécessaires pour les méthodes décrites ci-dessus, sont elles aussi étudiées : le pitch, le voisement, le calcul des MFCC, le calcul de la DTW, la détection d'activité vocale... Quelques méthodes d'analyse spectrale doivent être étudiées plus en profondeur.

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

SM10

Prérequis :

Traitement du signal (1CC4000)

Statistique et apprentissage (1CC5000)

Système d'information et programmation (1CC1000)

Plan détaillé du cours (contenu) :

Cours Magistral, 9H

1. introduction au traitement des sons (perception et production des sons...)
2. analyse spectrale pour l'analyse des sons (non-paramétrique/modèles AR(MA)/projections sur l'espace-bruit)
3. modélisation et analyse de la parole
4. reconnaissance de la parole
5. reconnaissance de la parole (Deep learning)
6. suppléments : synthèse et codage de la parole et de la musique, détection d'activité vocale ; reconnaissance du locuteur ; dialogue homme machine ; pitch tracking parole/musique ; voisement ; ...
7. conclusion

TD/TL, 9H

1. BE (3h00) : mise en oeuvre de PSOLA (Pitch Synchronous Overlap and Add), avec ce qui accompagne l'algorithme : détection activité vocale et/ou pitch tracking et/ou mesure de voisement
2. BE (3h00) : mise en oeuvre d'un système complet : MFCC+DTW
3. BE (3h00) : mise en oeuvre d'un système complet : Deep learning (RNN et LSTM)

Examen : les BEs

Déroulement, organisation du cours :

9h Cours magistral

9h Travaux dirigés/Travaux de laboratoire : 3 BE de 3 heures

Organisation de l'évaluation :

Contrôle continu (50% de la note ; 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
- 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.

Description des compétences acquises à l'issue du cours

:

- C1.1 : Examine a problem in full breadth and depth, within and beyond its immediate parameters, thus understanding it as a whole. This whole weaves the scientific, economic and social dimensions of the problem.
- C2.3 : Rapidly identify and acquire the new knowledge and skills necessary in applicable domains, be they technical, economic or others.
- C3.6 : Evaluate the efficiency, feasibility and strength of the solutions offered.
- C9.4 : Demonstrate rigour and critical thinking in approaching problems from all angles, be they scientific, social or economic.

Bibliographie :

Transparents

