

Reconnaissance et séparation de notes et d'accords dans un fichier audio musical par analyse spectrale

La reconnaissance de notes m'intéresse particulièrement du fait de son application directe dans les jeux de rythmes afin de générer des partitions automatiquement. C'est un genre de jeu vidéo que j'ai découvert récemment et qui me passionne par sa simplicité d'apprentissage, mais aussi par sa rigueur.

Jouer de la musique est en quelque sorte un jeu d'où l'expression. De plus, la détection de notes peut s'appliquer à la génération de partitions pour un jeu vidéo de rythme. C'est en cela que notre étude s'inscrit dans le thème "Jeux, sports".

Ce TIPE fait l'objet d'un travail de groupe.

Liste des membres du groupe :

Positionnement thématique (ÉTAPE 1) :

- *PHYSIQUE (Physique Ondulatoire)*
- *INFORMATIQUE (Informatique pratique)*

Mots-clés (ÉTAPE 1) :

Mots-clés (en français) Mots-clés (en anglais)

<i>Analyse spectrale</i>	<i>Spectral analysis</i>
<i>Détection de notes</i>	<i>Note detection</i>
<i>Transformée de Fourier</i>	<i>Fourier transform</i>
<i>Spectre fréquentiel</i>	<i>Frequential spectrum</i>
<i>Reconnaissance</i>	<i>Recognition</i>

Bibliographie commentée

Les fichiers sons en informatique permettent de condenser plusieurs signaux analogiques en une seule suite de données numériques grâce à l'échantillonnage et à la capture de son. Cependant, ce processus est irréversible et ne permet pas de récupérer les signaux originaux. L'enjeu du

projet est donc de savoir s'il est possible de récupérer les signaux originaux à partir du fichier initial ou bien de résoudre des problèmes comme si on possédait les signaux originaux à partir d'un seul fichier son.

De nombreux chercheurs se sont penchés sur le sujet, notamment grâce aux avancées récentes dans le domaine de l'intelligence artificielle. Le défi de ce projet est de ne pas utiliser l'intelligence artificielle dans nos recherches, puis de comprendre les éventuelles limitations d'une telle approche.

La difficulté du sujet se résume majoritairement à identifier les différentes fréquences qui se superposent : lesquelles sont des harmoniques, à quelles notes elles correspondent, de quel instrument elles proviennent etc.

Pour cela, nous allons recourir à l'analyse de Fourier, et plus particulièrement à la décomposition en séries de Fourier, passant par le concept de transformée de Fourier [1]. De plus, la compréhension des méthodes de prise de son et de conversion numérique-analogique [2] nous permettront une meilleure interprétation des résultats obtenus et l'identification d'éventuels problèmes rencontrés.

En raison de l'importance de la justesse d'un instrument, il existe de nombreux types d'accordeurs sur le marché. Ces accordeurs fonctionnent de différentes manières : Certains exploitent l'effet stroboscopique, d'autres font usage de la technique du passage à zéro, et d'autres encore se servent de la détection des pics et de pentes comme dans [3]. Un objectif de notre TIPE est de mettre au point un accordeur en utilisant uniquement la FFT [1] et de permettre au dispositif d'analyser la justesse de plusieurs notes en même temps, lors d'un accord. Cela permettrait aux musiciens d'avoir un retour immédiat en temps réel sur l'accordage de leur instrument.

Nous nous intéressons également à la séparation d'instruments à partir d'un fichier audio, ce qui consiste à « séparer » le fichier en deux autres contenant chacun la piste audio d'un seul instrument jouant sa partie. La difficulté est de parvenir à « supprimer » une note tout en laissant les autres intactes, mettant ainsi en jeu l'étude des harmoniques et la création de bases de données appropriées. Ce problème est appelé Séparation de Sources Audio (M.S.S.) et représente un sujet d'actualité dans le monde de la recherche en raison des avancées dans le domaine des réseaux neuronaux et de l'apprentissage profond [4] [5].

Une autre partie du problème consiste à générer des partitions de musique, et notamment des partitions spécialement conçues pour les jeux vidéo de rythme. En effet, concevoir ce type de partition peut se révéler long et complexe étant donné que le nombre de notes est limité. Des recherches utilisant l'intelligence artificielle ont déjà été menées [6], c'est pourquoi nous avons choisi de créer une partition à partir d'un fichier son sans l'aide de réseaux de neurones. Cette approche a déjà été étudiée par Karl O'Keefe [7] qui a étudié le cas des jeux vidéo de rythme n'ayant besoin que d'une seule partition. Il serait donc intéressant de déterminer s'il est possible de trouver un moyen de générer plusieurs partitions à partir d'un même fichier son. Par

exemple, on remarque que le sonogramme d'une musique peut se corréliser avec la position des notes dans une partition d'un jeu de rythme comme le montrent les travaux de Chris Donahue et Zachary C. Lipton Julian McAuley [6]. Nous allons donc nous en inspirer pour les adapter à un jeu dans lequel plusieurs joueurs peuvent jouer sur une partition différente. Par souci de temps, nous nous limiterons à la détection de notes, afin de générer la partition dans un second temps.

Problématique retenue

Comment isoler et reconnaître des notes extraites d'un morceau musical ?

Objectifs du TIPE du candidat

1. Conjecture des instants où une note est jouée à partir d'un sonogramme généré grâce à l'algorithme FFT d'une musique pour le jeu *Friday Night Funkin'* où seules les voix des personnages sont présentes dans un fichier son.
2. Mise en application de cette conjecture afin de détecter les occurrences de notes en utilisant un programme informatique Python.
3. Analyse des résultats en appliquant le programme sur des musiques sur lesquelles les partitions ont déjà été créées par la communauté, puis en comparant les instants trouvés par le programme avec la position des notes dans les partitions.

Références bibliographiques (ÉTAPE 1)

- [1] STÉPHANE BALAC : La transformée de Fourier vue sous l'angle du calcul numérique : 2011, *Cours de l'Université de Rennes 1*, cel-01862054
- [2] CARINE PASCAL ET NATHALIE TOMAS : Musique et Mathématiques : 2000, *Faculté des Sciences de Luminy, Université de la Méditerranée*
- [3] BHIMSEN BUDHATHOKI : Design of an Electric Guitar Tuner : 2018, *Thesis, Metropolia University of Applied Sciences*
- [4] JIALE QIAN, XINLU LIU, YI YU AND WEI LI : Stripe-Transformer: deep stripe feature learning for music source separation : 2023, *EURASIP Journal on Audio, Speech, and Music Processing*, DOI: 10.1186/s13636-022-00268-1
- [5] KILIAN SCHULZE-FORSTER : Informed Audio Source Separation with Deep Learning in Limited Data Settings : 2021, *Thèse de doctorat de l'Institut Polytechnique de Paris préparée à Télécom Paris*
- [6] CHRIS DONAHUE, ZACHARY C. AND LIPTON JULIAN MCAULEY : Dance Dance Convolution : 2017, *arXiv:1703.06891v3 [cs.LG]*

- [7] KARL O'KEEFE : Automated creation of step files for Dance Dance Revolution : 2003, MEng Individual Project Report, <https://monket.net/dancing-monkeys/DancingMonkeys.pdf>
- [8] MICHÈLE CASTELLENGO : Écoute musicale et acoustique, avec 420 sons : 2015, Edition Eyrolles, ISB13 : 978-2-212-13872-6

DOT

- [1] : Recherche documentaire sur les différentes représentations du son en septembre.
- [2] : Recherche d'une problématique commune au groupe et répartition de différents types de travaux entre nous en octobre.
- [3] : Création d'un sonogramme à partir d'un fichier sonore en novembre à l'aide du langage de programmation Python.
- [4] : Entre décembre et février, mise en place d'une stratégie afin de détecter les notes à l'aide de variations d'intensités sonores d'un fichier sonore en fonction de la fréquence et du temps.
- [5] : En mars, mises en oeuvre des stratégies sur des fichiers sonores dont les notes ont déjà été placées, puis analyse des résultats afin de voir si la stratégie est efficace.