

Le responsable scientifique devra développer au moins les points suivants :

- *Contexte, positionnement du projet présenté par rapport aux thématiques / projets du laboratoire*
 - *Avancées potentielles*
 - *L'équipe ressource mobilisée*
 - *Argumentaire de la demande*
-

- *Contexte, positionnement du projet présenté par rapport aux thématiques / projets du laboratoire*

Titre détaillé : « Descripteurs haut niveau issus de l'analyse spectrale pour une classification de la musique basée sur la signature sonore ».

Ce projet est directement lié aux projets du laboratoire au sein de l'axe « E-Formes : Arts et pratiques du numérique » et notamment du 2^e sous-axe (sur 3) : « écriture, analyse, perception... ».

Traditionnellement, le timbre dans la musique occidentale a souvent été considéré comme un élément secondaire de la musique, les paramètres principaux étant la hauteur des notes (régissant les mélodies et l'harmonie), leur amplitude (régissant les nuances) et leurs positions et durées dans le temps (définissant le rythme). Le timbre n'est déterminé que par le choix des instruments employés, ce qui constitue l'orchestration.

Mais la notion de timbre, ce qui doit décrire les caractéristiques du son, a progressivement changé de sens, d'une part par l'introduction de nouveaux sons dans la musique au XX^e siècle, qu'il s'agisse de sons concrets (sons enregistrés de toutes natures : sons instrumentaux, vocaux, industriels, de la nature, ...) ou de sons électroniques (sons de synthèse) et d'autre part par le travail de production des musiques enregistrées en studio, réalisé avec des appareils de traitement et de mixage électroniques et numériques qui permettent de modifier et de sculpter le son de façons très variées.

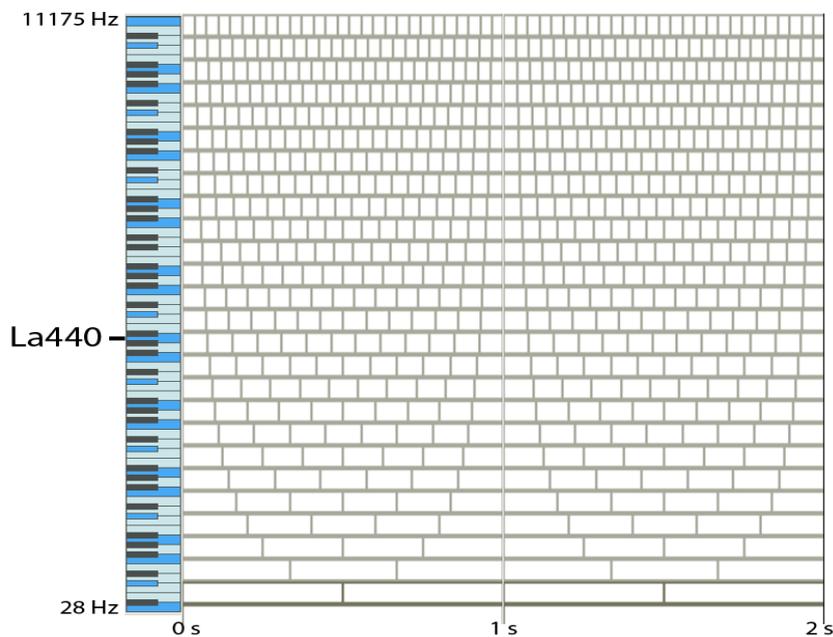
Ce que l'on perçoit du son est maintenant devenu un phénomène très complexe qui va bien au-delà de la liste des instruments de musique employés dans une composition. Quand on parle d'un son de guitare, la qualité sonore à laquelle on fait référence varie du tout au tout selon qu'il s'agisse d'une guitare acoustique classique, d'une guitare électrique de jazz ou du son saturé du rock (pour lequel existent en outre une infinité de variations de timbres). Qu'il s'agisse de musique contemporaine, de musique électronique ou de musique populaire, les compositeurs, les musiciens, les ingénieurs du son ont réalisé des expérimentations très variées et très sophistiquées vers la recherche de nouveaux sons et de nouveaux contrastes sonores.

Il nous semble donc fondamental de pouvoir décrire précisément les caractéristiques acoustiques de la musique, cette qualité qu'ont les sons qui est perçue par l'auditeur et nous ne disposons pas, aujourd'hui, d'outils très performants pour cette description.

Méthode d'analyse

Nous voulons développer une nouvelle méthode d'analyse qui permette de représenter qualitativement et quantitativement la nature du son présent dans un enregistrement.

La technique proposée consiste à prélever un échantillon du son d'une durée de quelques secondes (correspondant, pour des musiques pulsées, à une ou deux mesures) pour en effectuer des analyses spectrales (répartition de l'énergie dans différentes bandes de fréquences audibles). L'objectif est de découper le spectre en 27 bandes de fréquences, sur des plages correspondant à la perception (par tiers d'octaves, proche des Mels), pour ensuite mesurer l'énergie contenue dans chaque bande, sur des fenêtres glissantes de plus en plus fines en montant dans les aigus.



A partir de ce tableau de données, il est possible de représenter :

- l'enveloppe spectrale globale (sur la durée totale),
- l'enveloppe d'amplitude globale (sur l'énergie totale),
- les variations des enveloppes d'amplitudes par bandes de fréquences,
- les variations des enveloppes spectrales par fenêtres temporelles.

Le tableau peut être considéré tel quel, ou en version normalisée (amplitude maximum remplacée sur 0 dB).

D'autres données de plus haut niveau peuvent être extraites de ce tableau (centroïde spectral, flux...), mais c'est le tableau complet qui représente la signature sonore que nous voulons mettre en évidence et qui pourra nous permettre de calculer des distances entre divers échantillons. Il offre l'avantage d'être relativement réduit (756 valeurs) par rapport à l'analyse FFT complète du même échantillon sonore (20 000 valeurs pour une taille de fenêtre de 4096 échantillons sans recouvrement). Il prend en compte à la fois la richesse spectrale du son, en liaison avec les capacités de perception de l'oreille humaine, et les évolutions dynamiques du timbre (avec une précision inférieure au 1/10e de seconde). Cette signature, ne permet pas par contre détecter le rythme, le tempo, les notes jouées, l'harmonie. Elle n'est destinée qu'à la caractérisation de la qualité sonore.

Notre objectif vise à constituer un corpus d'œuvres qui nous paraissent représentatives, dans les différents domaines cités (musique contemporaine, musique électronique, musique populaire) d'un ensemble de qualités sonores contrastées. Pour chaque œuvre, nous déterminerons les extraits à caractériser, ce qui constituera la base de données de départ du projet, et nous élaborerons une sémantique de description.

Fouille de données

A partir de techniques de fouilles de données, il sera dans un premier temps nécessaire de réaliser des calculs de distances entre les différents échantillons, sur des critères à préciser (pouvant porter sur des descripteurs audio, sur des caractéristiques de timbres comme par exemple son de guitare acoustique, son de guitare saturée, chœurs, orchestre à cordes, ou des catégories perceptives : son brillant, sourd, granulé, ...).

Dans une œuvre donnée, un passage automatique de l'analyse sur l'ensemble du morceau devra permettre de mettre en évidence les similitudes rencontrées localement avec les échantillons de la base de donnée issus de ce même morceau. Cela permettra également de calculer un coefficient lié aux variations sonores du morceau.

Dans un second temps, des calculs de distances pourront être réalisés pour établir les similitudes et éloignements entre les différentes œuvres du corpus.

Toutes ces opérations permettront de calibrer progressivement des programmes automatiques de calcul de distances et de choisir ainsi les poids à apporter aux différents critères d'analyse, grâce à des tests de perception comparés réalisés entre les résultats de la machine et les résultats obtenus par des êtres humains entraînés à la pratique musicale.

Enfin, une fois les outils finalisés, ils pourront être déployés à plus grande échelle, à la demande, sur des bases de données musicales très vastes, afin de faire des rapprochements, de faire des recommandations à des amateurs de musique pour leur proposer d'autres œuvres possédant soit des signatures similaires, soit des progressions, des contrastes de signatures comparables.

- *Avancées potentielles*

Les débouchés pourraient donc être tout-à-fait novateurs et intéresser le plus grand nombre, la caractéristique des amateurs de musique étant toujours de chercher à découvrir de nouveaux talents, mais pas forcément par les biais que proposent actuellement la plupart des moteurs de recherches sur le WEB. La question du timbre, de la qualité, des caractéristiques du son des musiques nous paraît extrêmement actuelle et surtout d'avenir.

Une autre application concerne l'analyse musicologique des musiques du XXI^e siècle. De nombreuses publications parues ces quinze dernières années montrent l'utilisation croissante des descripteurs audio pour analyser la musique. Si ces descripteurs proposent des résultats intéressants pour montrer des variations des qualités audio d'une œuvre, pour réaliser sa segmentation, elles restent encore très rudimentaires pour décrire réellement le contenu sonore. La signature que nous proposons doit permettre de décrire le sonore dans sa globalité à un instant donné. Il faudra toutefois trouver des moyens de représentation graphique des résultats pour les rendre plus facilement perceptibles dans une analyse.

Par ailleurs, ces données (la signature sonore) pourront être utilisées dans des recherches sur la synthèse sonore, comme paramètres pour créer de nouveaux sons respectant une signature, ou pour en effectuer des transformations et créer des familles de sons nouveaux, par synthèse. C'est donc là un domaine de recherche et d'application qui peut également intéresser la création musicale.

Enfin, ces recherches doivent pouvoir trouver des applications dans des domaines non artistiques, comme l'étude des sons industriels. Le laboratoire Vibrations Acoustique de l'INSA mène des études pour savoir jusqu'où il est possible de simplifier des analyses de sons sans dégrader le spectre. L'approche que nous proposons pourrait servir pour évaluer la proximité entre un son calculé (avec un modèle plus ou moins grossier) et le son réel.

- *L'équipe ressource mobilisée*

Dans le projet, nous avons engagé des collaborations entre le laboratoire Hubert Curien et les chercheurs en informatique sous la direction de Fabrice Muhlenbach, le CIEREC et les chercheurs en informatique musicale sous la direction de Laurent Pottier, le laboratoire Vibrations Acoustique de l'INSA de Lyon, sous la direction de son directeur Etienne Parizet, l'IRCAM, à travers l'équipe Analyse et synthèse des sons, sous la direction de Geoffroy Peeters et de ses projets sur l'indexation des matériels audio et musicaux, et enfin la société stéphanoise 1Dlab-1Dtouch. 1D Lab est une Société coopérative d'intérêt collectif (Scic), qui s'inscrit dans les dynamiques et les valeurs de l'économie sociale et solidaire (ESS) et de l'innovation sociale. Un de ses projets majeurs est 1DTouch, la première plateforme mondiale de « streaming équitable » entièrement dédiée à la création indépendante.

Promotion de l'action culturelle territoriale

L'application de ce travail, avec les possibilités de proposer des recommandations musicales à travers des critères inédits, nous permettrait d'enrichir des collaborations existant déjà entre le laboratoire Hubert Curien et 1D Lab (Fabrice Muhlenbach co-encadre la thèse de Pierre-René Lhérisson qui travaille à 1D Lab suivant une convention industrielle de formation par la recherche).

- *Calendrier des recherches*

1) Etape 1 (M1) : mise au point de la méthode d'analyse. Premières missions à l'IRCAM pour discuter avec les chercheurs de l'équipe « Analyse-synthèse » des outils de développements qui nous pourrions employer pour mettre au point notre méthode d'analyse.

2) Etape 2 (M2-M3) : établissement du corpus d'œuvres représentatives, dans trois différents domaines : musique contemporaine, musique électronique, musique populaire) présentant un ensemble de qualités sonores contrastées. Pour chaque œuvre, découpage des différents extraits à caractériser ce qui constituera la base de données de départ du projet.

3) Etape 3 (M4-M6) : analyse du corpus et mise en place des outils informatiques permettant des calculs de distances et de similarités. Recrutement d'un stagiaire en traitement du signal (Masters ATIAM, INSA, RIM...) pour le développement d'une application optimisée et automatisée. Application de ces outils pour cartographier la base de données sur différents critères perceptifs.

4) Etape 4 : (M7-M8) : affinage des outils d'analyse et de fouille de données, en parallèle avec la réalisation de tests perceptifs pour valider la pertinence des résultats obtenus.

5) Etape 5 : (M9-M12) : phase de valorisation. Application et intégration sur la plateforme de streaming de l'entreprises 1D Touch, séances de travail avec le laboratoire Vibrations acoustique pour tester les outils sur des applications industrielles, publications des résultats, participation à des congrès sur les « Music Informations Retrieval ».

Le budget global comprend l'acquisition de machines rapides pour le calcul, de dispositifs de stockage des données, l'organisation de séminaires de recherches avec invités, déplacement et missions de travail avec les partenaires, le recrutement d'un stagiaire en traitement du signal et enfin l'intégration des outils de mesure et de recommandations élaborés par notre projet sur la plateforme 1DTouch. Ce dernier point nécessite la commande d'une API spécialisée à la société Mnémotix, actuellement responsable des développements de la plateforme 1DTouch.

A l'issue de cette première année de travaux, l'objectif est d'avoir des partenaires pour monter un projet de plus grande envergure par la suite (ANR, ou autre).