Représentation d'un signal audio par chromagramme

1 Introduction

Aujourd'hui la musique est principalement diffusée par des enregistrements sonores. Pour la plupart d'entre eux, la partition n'est pas disponible ou n'existe pas. Afin de pouvoir rejouer ces morceaux ou de faciliter leur analyse musicale, une partition est nécessaire. L'un des enjeux principaux de la recherche actuelle est de développer des algorithmes permettant d'obtenir une transcription automatique des morceaux de musique. Actuellement, les algorithmes existants sont encore peu fiables et l'on ne sait pas encore extraire la partition à partir d'un signal audio, surtout lorsque celui-ci contient des sons vocaux ou percussifs. Cependant, certaines représentations du signal permettent d'en extraire de manière satisfaisante des informations sur l'harmonie, la tonalité ou le rythme. L'objectif de ce TP est d'étudier un signal audio en utilisant l'une de ces représentations, le *chromagramme*.

2 Partie signal

Cette première partie est consacrée à la construction du *chromagramme*. Le chromagramme représente l'intensité de chaque note présente dans le signal audio à un instant donné. Il est constitué de *vecteurs de chroma* qui sont des vecteurs à 12 dimensions représentant les 12 demi-tons de la gamme chromatique.

La figure 1 montre l'exemple du chromagramme des quelques premières secondes du morceau *Misery* des Beatles. On peut voir que les notes DO, FA et LA sont présentes dans le signal.

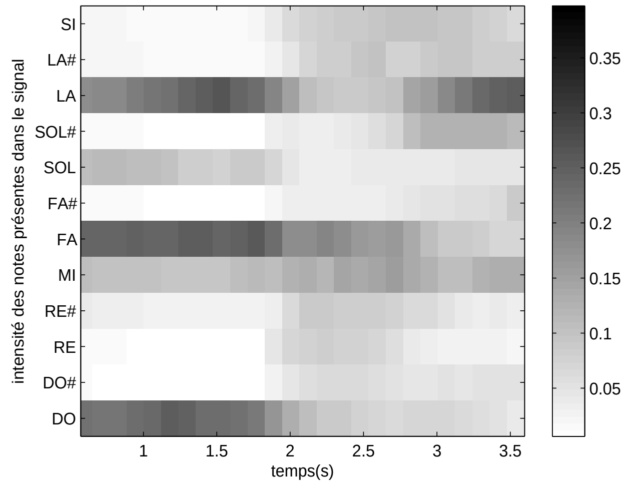

FIG. 1 – chromagramme du début de *Misery* des Beatles

2.1 Transformation du signal temporel dans le domaine fréquentiel

Dans le travail qui va suivre, la fonction matlab principale sera nommée **Chromagram**.

Question 1 : Dans **Chromagram**, convertir le signal audio stéréophonique échantillonné à $11025Hz$ en signal monophonique en prenant la moyenne sur les deux canaux.

Afin de pouvoir analyser le signal au cours du temps, le découper en "trames" se recouvrant partiellement (fenêtres de taille 0.5 s).

2.2 Conversion des fréquences f_k du spectre en notes midi

Le chromagramme ne tient pas compte de la position des notes par rapport aux différentes octaves. Les hauteurs séparées par un nombre entier d'octaves sont perçues comme "sonnant" de manière équivalente et ont le même nom. On dit qu'elles partagent le même *chroma*. L'ensemble des notes partageant un même chroma est appelé *classe de hauteur* ou *pitch class*.

Dans l'espace des classes de hauteur, il n'y a pas de distinction entre les notes qui sont séparées par un nombre entier d'octaves. Par exemple LA3 (440 Hz), LA4 (880 Hz) et LA5 (1760 Hz) appartiennent à la même classe de hauteur.

Les théoriciens de la musique se réfèrent en général aux différentes classes de hauteur en utilisant des nombres réels (on appelle cela échelle midi).

Question 2 : Ecrire une fonction matlab **freq2midi** qui permet de transformer la fréquence fondamentale f d'un son en un nombre réel p selon l'équation suivante (conversion en échelle midi) :

$$p = 69 + 12 log_2 \frac{f}{440} \tag{1}$$

Le le DO du milieu d'un piano correspond à la note midi 60.

Question 3 : Ecrire une fonction matlab **midi2freq** qui permet d'effectuer la transformation inverse.

Ces fonctions seront utilisées dans la question 4.

2.3 Construction d'un spectre de demi-tons

Dans cette partie, il s'agit de construire un spectre de demi-tons (de notes) en associant chaque fréquence du spectre à la note midi la plus proche. On ne considèrera que l'intervalle de fréquences $[61.7Hz; 1046Hz]$ (notes midi $[35; 84]$).

Question 4 : Construire un ensemble de filtres centrés sur les notes midi n' correspondant aux demi-tons appartenant à l'intervalle de fréquences considéré. Pour une résolution précise, chaque filtre représente un tiers de demi-ton. Le filtre $H_{n'}$ centré sur $n' \in [35 + \frac{1}{3}, 35 + \frac{2}{3}, \ldots, 84]$ est défini par :

$$H_{n'} = \frac{1}{2} \tanh(\pi(1 - 2x)) + \frac{1}{2} \tag{2}$$

où x est la distance relative entre le centre du filtre n' et les fréquences de la transformée de Fourier : $x = R|n' - n(f_k)|$.

Question 5 : Construire le spectre de demi-tons en appliquant l'ensemble des filtres à la transformée de Fourier du signal. Les valeurs N(n') du spectre de demi-tons sont obtenues par :

$$N(n') = \sum_{f_k} H_{n'}(f_k) A(f_k) \tag{3}$$

où les $A(f_k)$ sont les valeurs de la transformée de Fourier.

2.4 Construction du chromagramme

Question 6 : Construire le chromagramme (succession de vecteurs à 12 dimensions) en cumulant les valeurs du spectre de demi-tons correspondant à un même chroma :

$$C(l) = \sum_{n' \ tel \ que \ c(n') = \ l} N(n') \quad avec \quad l \in [0, 12[\tag{4}$$

Question 7 : Normaliser le chromagramme de manière à ce que la somme selon les lignes fasse 1.

3 Partie Audio

Question 8 : Utiliser la fonction **Chromagramme** pour représenter le chromagramme de différentes notes jouées par des instruments de musique. Reconnaît-on les notes jouées ? Entend-t-on d'autres notes que celle correspondant à la fréquence fondamentale ? En regardant le chromagramme, peut-on voir d'autres notes présentes dans le signal ? Comparer avec la transformée de Fourier.

Question 9 : Construire une gamme chromatique pour un instrument à cordes et un instrument à vent. Jouer la gamme pour vérifier puis calculer le chromagramme de cette succession de notes. Retrouve-t-on dans le chromagramme la gamme construite ? Y a-t-il une différence entre les deux instruments ?

Question 10 : En utilisant la fonction jouerAccord(note1, note2) du TP sur les Gammes,

- construire des accords d'octaves. Y a-t-il une différence avec les notes seules ?
- construire d'autres accords de 2 ou 3 sons. Peut-on facilement reconnaître les accords ? D'autres notes sont présentes dans le signal. En regardant quelles sont les harmoniques des notes émises, essayer d'expliquer ce qui est observé.

Question 11 : Choisir un court extrait d'un morceau de musique annoté en accords parmi ceux proposés. Construire un ensemble $C_i, i \in [1 : 24]$ de 24 vecteurs à 12 dimensions correspondant aux 24 accords majeurs et mineurs de trois sons. Pour chaque vecteur C_i, on aura $C_i(k) = 1$ si la note k appartient à l'accord, $C_i(k) = 0$ sinon. Par exemple, le vecteur correspondant à l'accord Do Majeur sera : 100010010000.

Pour retrouver les accords du morceau, calculer le produit scalaire des vecteur C_i avec chaque vecteur du chromagramme. Retenir à chaque instant la valeur la plus élevée et le vecteur C_i correspondant. Commenter.