Traitement du signal

Auteurs :
Ainhoa OLEAGA
Lucas BOUILLON
Sonia GARROUCH
Thi Minh Ha HO

Encadrant :
M. Charles DOSSAL

Table des matières

1 Introduction

Nous savons que l'anayse de Fourier et l'algorithme de Fast Fourier Transform (FFT) sont très puissants pour décomposer un signal comme une somme de sinusoïdes continues ou discrètes. Cette décomposition est très utilisée dans les domaines du traitement du signal pour faire du débruitage par exemple. Seulement l'analyse de Fourier est limitée à l'analyse de signaux stationnaires. En effet, procéder à une transformée de Fourier sur un signal non-stationnaire donnera un mauvais spectre en fréquence et on ne pourra pas vraiment savoir quelles fréquences sont présentes à quel moment.
La plupart des signaux (comme la voix par exemple) sont des signaux non-stationnaires. Il faut donc trouver un moyen de faire l'analyse de Fourier sur ce type de signaux.

2 Transformée de Fourier Court Terme

La Transformée de Fourier Court Terme (TFCT) aussi appelée Analyse de Fourier à fenêtre répond tout à fait à ce besoin.

2.1 Théorie

L'idée principale est de faire la transformée de Fourier du signal sur des petits intervalles. Ainsi le signal peut être considéré comme stationnaire sur ces intervalles. Dans le cas continu, l'idée est de multiplier le signal par une fonction fenêtre (fonction non nulle sur un intervalle de temps court) et de prendre la transformée de Fourier de la résultante. La fonction fenêtre est ensuite translatée sur l'axe des temps afin de prendre la transformée de Fourier de la prochaine partie du signal et ainsi de suite.
Mathématiquement, on peut l'écrire comme suit :

$$\mathbf{TFCT}\{x(t)\}(\tau,\omega) \equiv X(\tau,\omega) = \int_{-\infty}^{\infty} x(t)w(t-\tau)e^{-i\omega t}dt$$

avec w(t) une fonction fenêtre et x(t) le signal.

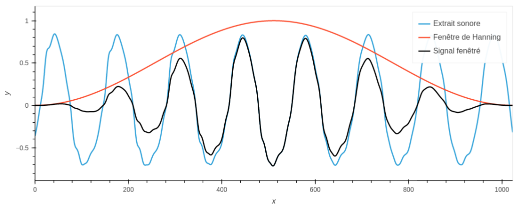

FIGURE 1 – Graphique illustrant un extrait sonore (en bleu), une fonction fenêtre (en rouge) et le signal fenêtré résultant (en noir)

Dans le cas discret on l'écrit :

$$\text{TFCT}\{x[n]\}(m,\omega) \equiv X(m,\omega) = \sum_{n=-\infty}^{\infty} x[n]w[n-m]e^{-j\omega n} \tag{1}$$

Dans le cas continu comme dans le cas discret, il est important de laisser du recouvrement entre les découpages du signal original. En effet, si les intervalles choisis sont disjoints le contenu fréquentiel du signal se retrouve altéré et l'analyse de Fourier n'est donc pas très fiable sur les bords de chaque intervalle. La résultante de la TFCT est une matrice où chaque colonne correspond à la transformée de Fourier du signal sur un intervalle de temps.
On obtient le spectrogramme du signal en prenant la norme au carré de la résultante.

2.2 Reconstruction du signal original

On pourrait penser que pour reconstruire le signal original il suffirait de calculer les FFT inverses des colonnes de le TFCT et de les mettre bout à bout mais à cause des fenêtre de recouvrement et des fonction fenêtre cela n'est pas possible. Il faut en fait d'abord multiplier les colonnes par la même fonction fenêtre qu'utilisée précedemment puis on les ajoute en utilisant le même décalage que lors du calcul de la TFCT :

$$x(t)w(t-\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\tau,\omega)e^{+j\omega t}d\omega \tag{2}$$

3 Vocodeur de phase

3.1 Théorie

Le but d'un vocodeur de phase est de modifier la durée d'un signal sans modifier sa tonalité. Les étapes de base d'un vocoder sont :

1. calculer la relation fréquence/amplitude instantanée en utilisant une TFCT
2. étape de processing sur les résultantes de la TFCT (comme leurs rééchantillonage par exemple).
3. effectuer les TFCTs inverses afin de reconstruire un signal (qui sera donc différent de l'orignal grâce à l'étape intermédiaire de processing).

Dans le TP nous avons choisi de stocker dans deux tableaux séparés les modules et les phases des FFT afin de pouvoir, d'une part, interpoler le spectrogramme de la TFCT initiale et d'autres parts, trouver la différence de phase entre le spectrogramme original et le nouveau spectrogramme. Ainsi, on reconstruit la phase fenêtre par fenêtre à partir de la phase de la fenêtre précédente du spectrogramme reconstruit en ajoutant la différence de phases entre les deux fenêtres correspondantes du spectrogramme original.

4 Séparation de sources

L'objectif de cette partie est de prendre un son stéréophonique qui est un mélange de plusieurs sources sonores (dans notre présentation différents instruments de musiques) et de séparer les sources à partir de deux voix d'enregistrement d'un même son.

4.1 Théorie

Dans le cas de notre étude, l'enregistrement des deux voix se déroule comme suit : deux instruments de musiques jouent à une certaine distance, la première voix enregistre la musique en étant plus proche du premier instrument et la deuxième voix enregistre en étant plus proche du deuxième instrument.

Nos seules hypothèses sont la connaissance du nombre d'instruments qui jouent et le fait que pour un instant donné, chaque instrument joue une note différente (donc une fréquence différente). Un spectrogramme dresse une carte temps-fréquence du signal sonore avec une variation de couleur qui représente l'intensité. En définissant le rapport d'intensité des deux instruments, on peut ainsi associer chaque point temps-fréquence du spectrogramme à un instrument. On peut reconstruire ainsi la piste de l'instrument 1, par exemple, en créant un spectrogramme uniquement constitué des points temps-fréquence associés à cet instrument.

Dans le projet nous essayons de séparer un son contenant deux sources et un son contenant trois sources en connaissant les coefficients de mélange et en utilisant comme seuil la moyenne géométrique des coefficients de mélanges $T = \sqrt{R_1 R_2}$ $avec$ R$_i$ le i^{me} coefficient de mélange.

4.2 Estimation des coefficients de mélange

Dans cette partie, nous supposons que nous ne connaissons plus les coefficients de mélanges mais que nous voulons quand même séparer les sources. On parle de séparation à l'aveugle. Nous cherchons donc à estimer le rapport d'intensité pour chaque instrument dans les deux voix.
L'idée est de faire un histogramme de $U = \frac{R}{1+R}$ avec R le rapport d'intensité des modules des TFCTs (l'histogramme de U est plus marqué que l'histogramme de R directement) et sélectionner les x valeurs les plus représentées dans l'histogramme avec x le nombre de rapports d'intensités que l'on souhaite trouver.

4.3 Clustering

Nous souhaitons trouver une autre manière de séparer des sources qui serait basée sur le clustering cette fois. Notre idée serait notamment de faire un K-Means sur les spectrogrammes issus des TFCTs des sons mélangés. Nous aimerions également mettre une note de confiance quant à la classification de chaque point.
Nous avons construit le dataset en utilisant la TFCT. Chaque individu, ie chaque point du spectogramme, dépend de 3 variables : le temps, sa fréquence et son intensité. L'intensité est le carré de la valeur absolue du coefficient correspondant dans la matrice de TFCT du signal.