

Transcription automatique de la musique de piano

Tarak Arbi

1. Objectif du projet

Ce projet porte sur la transcription de la musique de piano. Il consiste à développer un programme permettant d'analyser le son d'un fichier audio pour estimer les notes jouées avec les paramètres associés et en particulier, les instants d'attaque et les durées de ces notes de musique. L'application doit ensuite créer une représentation visuelle intelligible des notes jouées.

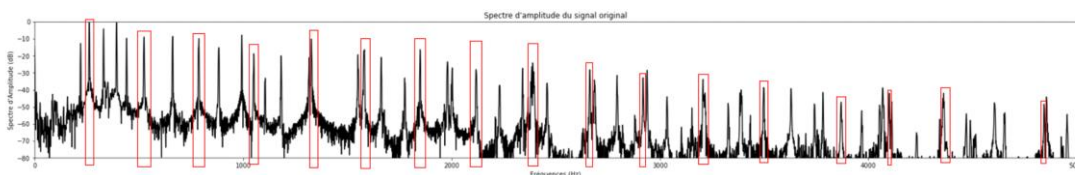
Le piano est un des instruments les plus difficiles à transcrire, du fait notamment de la forte polyphonie, et la transcription automatique de la musique de piano reste un sujet actif de recherche [1]. Pour parvenir à cette transcription, vous pourrez considérer plusieurs hypothèses simplificatrices afin de restreindre ce vaste cadre. Pour commencer, vous pourrez ne traiter que des fichiers audio ne comprenant qu'une seule note et chercher par la suite à supprimer cette hypothèse en considérant par exemple : des fichiers audio avec plusieurs notes successives et le cas de plusieurs notes jouées en même temps ...

2. L'estimation de hauteur d'une note de musique au piano

L'estimation de hauteur d'une note consiste à estimer la fréquence fondamentale d'un signal périodique ou quasi-périodique. Pour cela, plusieurs algorithmes ont été proposés dans la littérature. Parmi ces méthodes, on trouve la méthode de produit spectral ; elle consiste à estimer la fréquence fondamentale telle que :

$$f_0 = \underset{f}{\operatorname{argmax}} \log \left(\prod_i^H |X(if)|^2 \right),$$

où $X(f)$ est la transformée de Fourier du signal et H est le nombre total d'harmoniques considérées. En effet, cette méthode considère les harmoniques (voir Figure ci-dessous [2]) multiples de la fondamentale pour estimer la fréquence fondamentale car une approche qui consiste à chercher le pic principal du spectre est certes simple, mais ses performances en revanche restent insatisfaisantes, notamment à cause du bruit et du nombre d'échantillons.



De surcroît, les algorithmes de hauteur d'une note qui utilisent la transformée de Fourier discrète souffrent de fuites spectrales dans le cas où la période d'observation est insuffisante. Pour une meilleure estimation de l'estimation de la fréquence fondamentale, il est souvent utile d'utiliser du fenêtrage (comme la fenêtre de Hamming).

Cet algorithme peut être utilisé pour l'estimation des fréquences fondamentales multiples (plusieurs notes jouées simultanément) [1]:

ENTRÉES: signal audio polyphonique x
 $r \leftarrow x$ {Initialisation}
Tant que r contient une note
 Sélectionner la fréquence fondamentale prédominante f_0 dans r
 Estimer le signal s correspondant à f_0
 $r \leftarrow r - s$
Fin Tant que
SORTIES: liste des f_0 successivement estimées

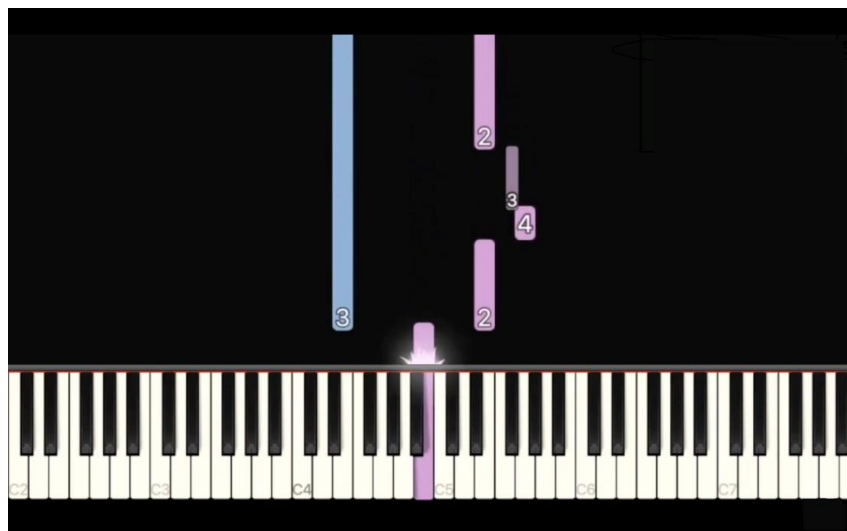
Toutefois, ses performances ne sont pas toujours satisfaisantes, en particulier, dans le cas où des notes en rapport harmonique sont jouées simultanément.

Enfin, d'autres algorithmes existent pour l'estimation de la hauteur d'une note et notamment des techniques dans le domaine temporel [3, 4] qui permettent d'estimer la période fondamentale ; ainsi l'algorithme de YIN [4] permet grâce à la fonction d'autocorrélation, de chercher le plus petit décalage temporel non nul pour lequel la forme d'onde et sa version décalée coïncident. Vous pouvez aussi si vous le souhaitez utiliser des filtres passe-bandes étroits centrés autour des 88 fréquences fondamentales des notes du piano [5]. D'autres algorithmes récents utilisant des méthodes de l'intelligence artificielle, comme les réseaux de neurones, ont aussi été proposés pour la transcription de la musique de piano [6]. Vous êtes libre d'implémenter votre méthode de prédilection.

NB. La bibliothèque GSL [7] peut vous être profitable ; elle définit plusieurs fonctions mathématiques comme la transformée de Fourier rapide.

3. Interface graphique

La bibliothèque SDL [8] comprend plusieurs fonctions qui permettent de créer des interfaces graphiques. Vous pourrez l'utiliser pour illustrer les notes de piano que votre algorithme d'estimation aura détectées, en mettant en surbrillance les touches de piano associées à ces notes, comme illustré sur la figure suivante :



4. Références

- [1] Valentin Emiya. Transcription automatique de la musique de piano. Télécom ParisTech, 2008. English. ffpastel-00004867f
- [2] <https://zestedesavoir.com/tutoriels/3013/reconnaissance-de-notes-de-musique/>
- [3] McLeod, Philip and Geoff Wyvill. “A Smarter Way to Find pitch.” *ICMC* (2005).
- [4] de Cheveigné, Alain and Hideki Kawahara. “YIN, a fundamental frequency estimator for speech and music.” *The Journal of the Acoustical Society of America* 111 4 (2002): 1917-30.
- [5] Müller, Meinard, Frank Kurth and Tido Röder. “Towards an Efficient Algorithm for Automatic Score-to-Audio Synchronization.” International Society for Music Information Retrieval Conference (*ISMIR*), (2004).
- [6] Böck, Sebastian and Markus Schedl. “Polyphonic piano note transcription with recurrent neural networks.” *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2012): 121-124.
- [7] <https://www.gnu.org/software/gsl/doc/html/fft.html>
- [8] <https://wiki.libsdl.org/>