DSP Final Project Team 3
樂譜產生器

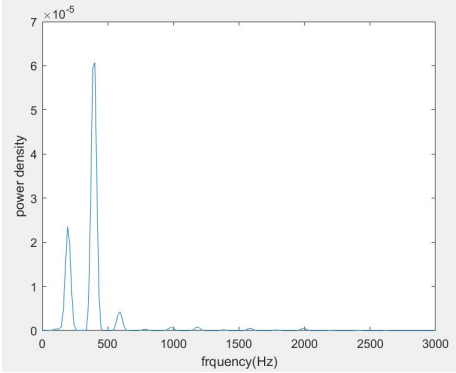
Members: 高智遠、劉宇瑩

I. 問題定義/應用場景

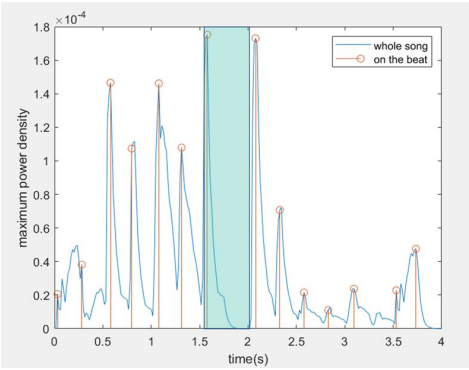
我們想提供一個將音樂轉成五線譜的軟體，輸入為音樂檔案及曲速，音樂檔案限定為單音樂器錄音，曲速則以bpm為單位，輸出則是五線譜的pdf檔。使用這個軟體，只要將創作者的即興演出錄製成音檔(我們的測資使用WAV檔)，就能用軟體產生五線譜將創作記錄下來，如此一來創作音樂的過程變得更簡單，也更快速。

II. 問題分析

1. 雖然曲速是已知的，但通常在樂曲開始之前就會開始錄音，首先我們要找到音樂開始的時間點。
2. 程式中需要判斷音的長度，例如我們的測資中有四分音符、八分音符，要從錄音的資訊辨別這個音是長音，或是相同的音重複彈奏。
3. 樂器會產生泛音，從頻譜判斷音高會產生困難。我們的測資使用線上產生音樂的網站Song-Maker的鋼琴單音，有些頻譜會出現泛音比主音能量更強的狀況。例如小星星中的第三個音(g)，頻率應該是195.998Hz，但其泛音(g’391.995Hz) 的能量更高。



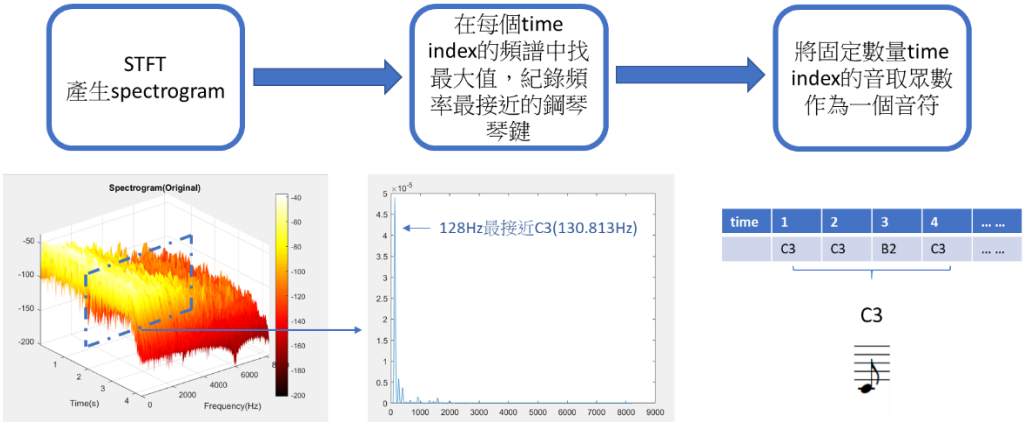
4. 無法判斷音樂演奏過程中的停頓為長音或是休止符。



III. 定義方法

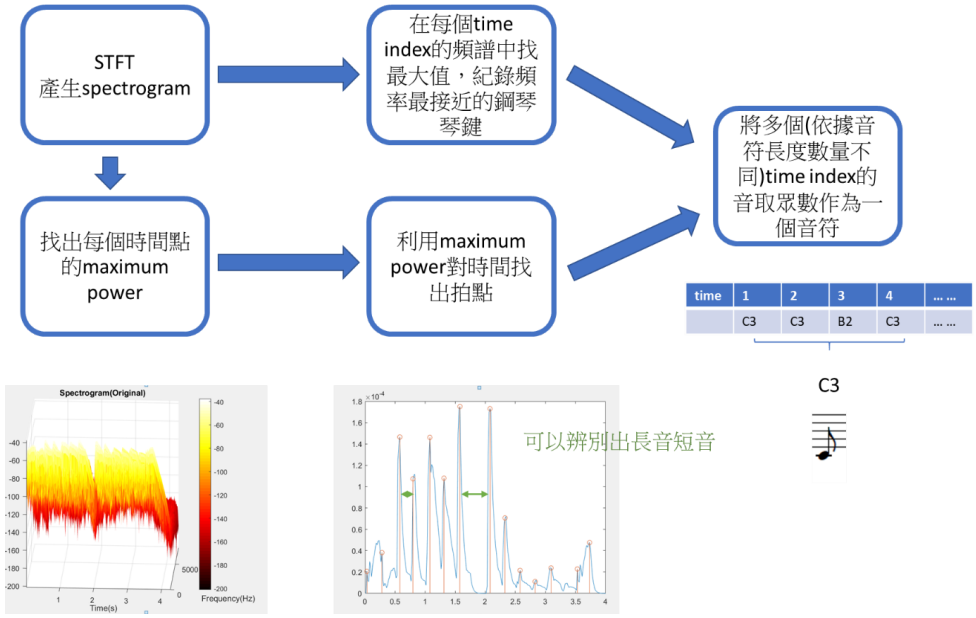
method 1

我們會將輸入的錄音檔做短時距傅立葉變換(short-time Fourier transform)，並建立鋼琴琴鍵鍵數對頻率的表，從spectrogram中找到每個時間點最接近的鋼琴音，由short-time Fourier transform的window長度、訊號的取樣頻率、及曲速可以算出一個音符包含幾個frame，再取眾數為音符。（例如曲速120bpm，window長度512個點，重疊256個點，取樣頻率16000Hz即可算出一個八分音符約是16個frame，所以把16個琴鍵音取眾數得到一個八分音符)

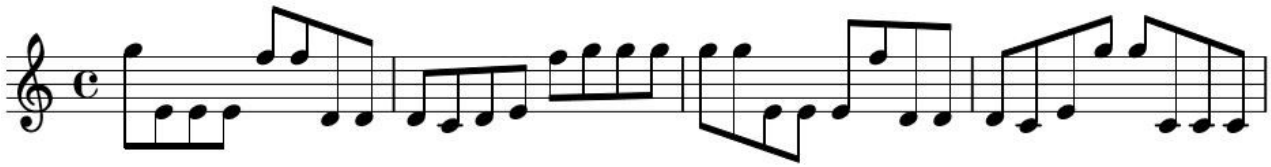


method 2

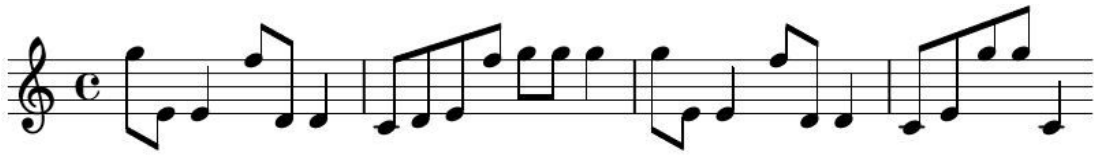
由於上述的方法只能產生固定長度的音符(例如全部為八分音符或全部為四分音符)，所以我們用能量密度對時間的關係當作判斷音符長度的依據，和前述方法相似，先利用曲速和STFT的window長度可以算出一個八分音符為幾個frame(我們此次用的最小單位音符，也可以設定的更短)，測資中約為16個，所以我們以16及其倍數間隔找能量密度對時間的peak當成音符的拍點，並以peak值的時間差為音符的長度。



IV. 結果分析



上圖為第一種方法產生的小蜜蜂樂譜(樂譜所產生的音樂於./TestPattern/star_v1.wav)，可以發現大部分的音符是有抓到，已能辨識出歌曲，然而固定數量的time index使得產生的音符僅能為固定長度，以小蜜蜂為例，其中有些四分音符在這個方法下會變成由兩個八分音符取代。



在第二種方法中，將音符出現的時間點也考慮進去，便能解決第一個方法中僅能產生固定長度的音符的限制，由上圖可觀察到有四分音符的出現，更符合原本音檔的節拍。

由兩種方法所產生的樂譜可以觀察到音高的分布十分極端，有些連續兩個音符的音距將近一個八度，而這是因為泛音的能量大於基頻，在我們的方法下取到的是高一個八度的音，然而音名是正確的，故聽起來仍算和諧。

V. 結論

在第一種方法上我們可以找到每個時間點下的音符，而第二個方法則是透過在能量密度對應時間關係上找peak值的方法有效解決音樂的起始點以及音符的長度的問題，然而長音以及休止符在目前方法下仍難以辨別。在頻率的選擇上，找能量最強的頻率上會找到泛音的問題亦有待解決。

VI. 貢獻

A. 方法貢獻

除了short-time Fourier transform的概念主要是從「數位聲訊分析與合成」課程所學的概念外，其餘方法皆為原創。

B. 實作貢獻

short-time Fourier transform的部分在Python使用到librosa這個聲音訊號分析的library，而在音符轉成可視化樂譜上，所使用的是Python的mingus這套音樂創作的library。

VII. 組內分工

method1方法及程式實現: 高智遠
method2方法: 劉宇瑩
method2程式實現:共同完成

VIII. 參考

[1] 測資來源: song maker https://musiclab.chromeexperiments.com/Song-Maker/

[2] Python package: mingus https://bspaans.github.io/python-mingus/#

[3] Python package: librosa https://librosa.github.io/librosa/