

資工系實務專題研究計畫成果報告

自動移調撥放器

專題編號：098-CSIE-S003

執行期限：97 年 1 學期至 98 年 1 學期

指導教授：楊士萱 教授

專題計劃參與人員：林文靖、林婉茹

一、中文摘要

本系統主要是開發一個音調比對系統，讓使用者哼唱一小段曲子，經由 FFT 快速傅立葉轉換，計算使用者哼唱結果之頻率陣列，再和原唱頻率陣列做音調比對，計算出和原唱相差幾個半音，最後藉由播放器呈現移調結果。

關鍵詞：STFT、FFTW、MFC、Phase Vocoder

二、緣由與目的

現今，音樂普及化造成日常生活中接觸之機會增加，聽音樂成為年輕人的最愛之一，同時，許多人對唱歌漸感興趣。唱歌對人們帶來之吸引力除能趕上潮流，另一方面也能藉由娛樂達到抒解壓力的目的。然而，唱歌時往往會有走音或破音等問題，在 KTV 唱歌時，這樣的問題常常發生，雖然遙控器能做升降調的動作，但始終有些人不懂樂理，很難選出適合自己的音調。有鑑於此，我們想到，若能有一系統讓使用者只哼唱一小段就能找出適合自己音調的配樂，使用者能更精確地唱出歌曲，因此以自動移調播放器做為專題主題。

三、系統介紹

步驟一.取得頻率陣列

將讀取進來的 wav 檔音波，以 4096 個 sample 為一個音框，也就是總共有 (wav_length / 4096)個音框，再對每個音框做頻率偵測，產生頻率陣列。

$$X[k] = \sum_{n=0}^{N-1} x[n]e^{-j2\pi kn/N}$$

此公式即 DFT 公式。其中 $X[k]$ 代表的值就是頻率為該音框時間內有 k 個全波的組成大小。頻率計算公式如下：

$$\begin{aligned}\text{frequency} &= k / \text{window_time} \\ &= k / (\text{frame_size} / \text{sample rate}) \\ &= k * \text{sample rate} / \text{frame_size}\end{aligned}$$

步驟二.找出需要移多少半音

得到原唱與使用者聲音陣列後，必須將兩個陣列對齊，去除錄音一開始的空白期(因為每個人按下錄音鍵的時間不同)。人最低的聲音大約在 65.4Hz，因此，將小於 65.4Hz 的頻率陣列忽略，當陣列頻率開始大於等於 65.4Hz，再把之後的頻率陣列存放至另一新的頻率陣列，即完成去除空白期之動作。

將兩個已對齊的陣列，做相除的動作，產生一個「頻率比值」陣列。依據 12 平均律公式，計算出相差幾個半音。12 平均律公式如下 (f :頻率比值, k :相差幾個半音)：

$$2^{k/12} = f$$

將以上公式轉換即可求出相差的半音數

$$k = 12 \times \log_2 f$$

步驟三.配樂移調

每 1024 samples 為一個音框，對音框做 FFT 轉換成 $X[k]$ 取得 phase 陣列與 Magnitude 陣列。音高調整方法 Phase Vocoder 分為以下兩部分：

【振幅調整方法】

原訊號第一音框的 Magnitude 陣列即為新訊號第一音框的 Magnitude 陣列。原訊號第二音框的 Magnitude 陣列減第一個音框的 Magnitude 陣列，除以 2，再加上第一音框的 Magnitude 陣列，即為新訊號第二音框的 Magnitude 陣列。以此類推計算，其公式為：

$$gLastMagn[k] + (magn - gLastMagn[k])/2$$

【頻率調整方法】

原訊號第一個音框的 phase 陣列即為新訊號的第一個音框 phase 陣列。原訊號第二音框的 phase 陣列減第一個音框的 phase 陣列，再減去其對應的期望值 $expect = 2\pi k / osamp$ 即為新訊號第二個音框的 phase 陣列。以此類推計算。

最後對 Phase Vocoder 產生的新訊號做 Resample，即可產生移調效果。

步驟四.介面呈現

【播放器】 利用 CActive Movie 達成播放器功能,具備 play, stop, pause, volume up, volume down 等功能,利用 C File Dialog 呼叫開檔 Dialog 並且將路徑存放在 CEdit。

【錄音機】 使用 MCI_RECORD 建立一個簡單的錄音裝置。功能:開始錄音、停止錄音、播放、試聽。

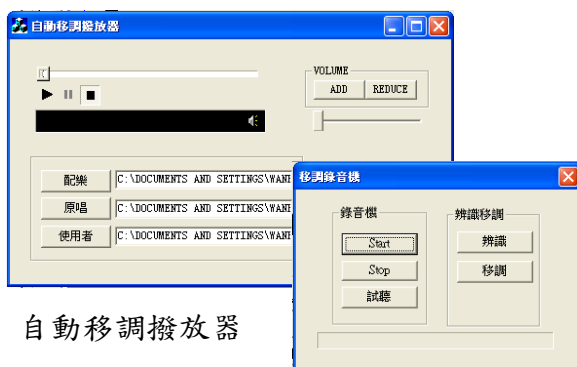
【自動移調系統與介面連結處】

「辨識」按鈕:

告知使用者與原唱相差多少半音。

「移調」按鈕:

讀取半音差對配樂進行移調功能。



自動移調播放器

四、結論

曾在學校課程中學習到 Signals and Systems 及 DSP，發現 DFT(傅立葉轉換)是個很好的處理頻率偵測方法，以此延伸出 FFT(快速傅立葉轉換)，使效率上更加快速。我們先從偵測單音做起，偵測電腦做出來的純音，再偵測人聲哼唱的單音，發現倍頻問題(找基頻)，並解決之。頻率偵測是整個系統最困難的地方，因為造成錯誤的因素很多，如程式、偵測方法、麥克風、錄音環境等，皆會影響到系統的精確性。此外，移調程式原先是從參考資料[3]的網址取得，但其支援聲音檔規格與本系統不符，且原本程式碼的移調效果不好，有許多雜音，後來我們修改了部分的程式碼解決這些問題，成功做出移調效果後，將使用者哼唱的聲音與原唱者的聲音進行音高辨識，然後做出升降調的效果，顯示出和原唱者差多少半音，最後產生出移調後的配樂，並使用 MFC 視窗介面來實現此移調系統。

五、未來期望

未來希望加入節奏偵測功能，將原唱者頻率陣列與使用者的頻率陣列比較對齊，找出原唱者與使用者聲音的節奏相似性，可讓偵測半音差的結果更加準確，將本系統的準確度趨於完美。

六、參考資料

- [1] 張元治，中華民國九十七年二月，台灣大學資訊工程學研究所碩士論文「用於伴唱機的伴奏音樂自動移調系統」
- [2] FFTW <http://www.fftw.org/>
- [3] The DSP Dimension <http://www.dsppdimension.com/admin/pitch-shifting-using-the-ft/>

