

樂器聲音重建之自動化分析模組開發與應用

Development and Application of Automatic Analysis Module for Musical Sound Reconstruction

王栢村¹、蘇秉翔²、溫儒琳^{2*}、吳盈輝³

¹ 國立屏東科技大學機械工程系教授

² 國立屏東科技大學機械工程系研究生

³ 國立內埔高級農工職業學校機械科技士

摘要

以往對於樂器的探討，為了瞭解樂器聲音特徵，須人工進行特徵分析，包含聲音擷取、頻譜分析、抓取峰值頻率等等，分析過程十分繁瑣。為了加快樂器聲音探討的效率，本文建立一套聲音重建之自動化分析模組，透過模組可以快速的進行特徵分析，並將樂器聲音特徵，利用理論方程式來進行聲音重建。首先將聲音特徵區分為音色參數與時間參數，音色參數包含峰值頻率、振幅、阻尼比與相位角，音色參數可分別透過自身功率頻譜與傅立葉頻譜兩種方式進行計算，兩者差異為透過自身功率頻譜方式無法取得相位角，傅立葉頻譜方式可以取得相位角；時間參數則分為衰減率與包絡線，衰減率是依照時間域訊號衰減程度，計算出對應阻尼比，包絡線為原始時間域訊號波形輪廓趨勢，將聲音重建時間域訊號與包絡線進行比例調整。最後應用於中提琴、響鈴板與銅鐘的聲音重建，中提琴有持續拉奏的聲音特徵，適合使用包絡線進行聲音重建，打擊樂器則適合使用衰減率進行聲音重建。未來可透過此聲音重建之自動化分析模組，可以快速分析不同聲音，取得對應的聲音特徵，並能應用於不同樂器聲音的組合。

關鍵字：聲音重建、自動化分析、音色參數、時間參數

Abstract

For investigating musical instruments, musical sound characteristics are of interest. Sound measurement, spectral analysis and capturing peak frequencies are required and mostly relied by manual operation. To expedite the analysis process, this work aim to develop the automatic analysis module for musical sound reconstruction. The module can extract key parameters of sound and rebuild the sound. The sound parameters contains both tonal parameters and time parameters. Tonal parameters, including peak frequencies, amplitudes, damping ratios and phase angles, can be derived from sound auto spectrum or Fourier spectrum. The difference is that only Fourier spectrum can provide the phase angle data while auto spectrum cannot. Time parameters are defined with two aspects, i.e. the decay rate and enveloping methods. The decay rate refers to the decay effect of sound signal, in which the damping ratio can be obtained. The

enveloping method is to mask the time waveform of sound. The masked enveloping profiles are used to adjust the reconstructed sound. The module is then applied to examine musical sounds for viola, ring plate and copper bell. The viola sound with the continuous playing sound due to the bow and string in frictional contact is suitable for using the enveloping method to rebuild the sound. The percussion instrument sound having the decay effect is better to use the decay rate method. The developed sound reconstruction analysis module can easily and quickly analyze different types of musical sound or others to obtain the key parameters of sound. The rebuilt sound can be integrated to generate the combination of different musical instrument sounds as well as other applications.

Keywords: sound reconstruction, automatic analysis, tonal parameters, time parameters

一、前言

不同的樂器都會發出獨自特有的聲音，大部分的樂器是透過演奏方式，使樂器結構振動，結構振動會帶動周圍空氣，進而發出聲音，過去在探討樂器時會將聲音做為研究的重點。Park et al.[01]利用機器手臂拉奏小提琴，模擬人按壓琴弦並拉奏的動作，收集多筆人拉奏效提琴的聲音特性，並與機器手臂拉奏的聲音特性進行關聯性探討，結果顯示，人類的皮膚會緩減弓與弦的振動，可以更準確的演奏正確的音階。Debut et al.[02]研究古老歐洲鈴鐺，因現有的鈴鐺實體以損壞，透過有限元素分析，建構出無損壞鈴鐺，從材料參數、幾何尺寸繪製、結構模態分析與聲音模擬等步驟，重現出古老歐洲鈴鐺的聲音。王與蘇[03]透過理論與實驗探討吉他弦之振動模態特性，量測吉他弦聲音取得聲音頻譜特性，發現吉他聲音有產生倍數簡諧頻率之特徵，此為弦振動較高模態之自然頻率，可以確認與吉他弦振動之關聯性，證明吉他結構有產生擴音的效果。

王等人[04]探討聲音模擬方程式的可行性，發展簡易的時間域訊號之模擬方式，透過輸入系統參數，包含：時間點數、取樣開始時間長度、平均次數等等參數，即可產生聲音檔，並透過實際打擊樂器銅鑼之聲音，進行驗證，結果顯示，可模擬出近似之聲音聽覺得效果。Ilmoniemi et al[05]運用心理聲學理念，探討不同聲音音色的差異，將音色區概分為四種不同特性：奇數偶數簡諧倍頻、亮度、衝擊時間、噪音量，透過主觀者聆聽，評估音色上不同的特性差異，而產生一組主觀之聲音評估指標。

本文目的旨，為開發聲音重建之自動化分析模組，針對樂器聲音進行自動化分析，取得聲音特徵，包含音色參數與時間參數，再透過理論方程式進行聲音重建，最後應用模組於中提琴、響鈴板與銅鐘，成功取得對應的聲音特徵，及重建聲音。

二、聲音重建之理論解析

本節介紹聲音重建理論解析，聲音特徵主要包含音色參數與時間參數，首先說明音色參數，音色參數包括了頻率、振幅、阻尼比與相位角，取得音色參數的方式，分為功率頻譜密度函數(PSD)方法及傅立葉頻譜(Fourier spectrum)方法。

圖 1 為音色參數檔案流程圖，功率頻譜密度函數 $G_{pp}(f)$ 方法，是將原始時間域訊號 $p(t)$ ，透過 SM 軟體[08]進行頻譜分析，過程中為了降低雜訊而進行平均處理，可使 $G_{pp}(f)$ 有較平滑的曲線，信噪比較佳，但 $G_{pp}(f)$ 為實數，因此無法計算相位角數值，所以相位角數值以 0 取代。

傅立葉頻譜 $P(f)$ 方法，是將原始聲音時間域訊號 $p(t)$ ，進行快速傅立葉轉換 FFT，沒有透過平均處理，因此信噪比較差，但是可以保留較原始的訊號。且傅立葉頻譜 $P(f)$ 為複數，因此可以使用實數部與虛數部，計算對應的相位角數值。

因要取得音色參數：峰值頻率、振幅，都需要手動抓取，並透過半能量點法計算阻尼比與相位角，透過人工抓取並計算，是沒有效率的，因此本文開發自動化分析模組，透過模組，可自動抓取頻率、振幅，並自動計算阻尼比與相位角，加快音色參數提取的效率。

時間參數分為衰減率及包絡線兩種，圖 2 為衰減率計算阻尼比流程圖，衰減率是透過原始時間域訊號 $p(t)$ ，使用 SM 軟體進行抓取，接著透使用衰減率計算阻尼比，其方程式如下：

$$\xi_r = \frac{\sigma}{\omega_r} \quad (1)$$

其中， ξ_r 為阻尼比， r 為第幾個峰值， σ 為衰減率、 ω_r 為峰值頻率(rad/sec)，透過衰減率及可計算出阻尼比。

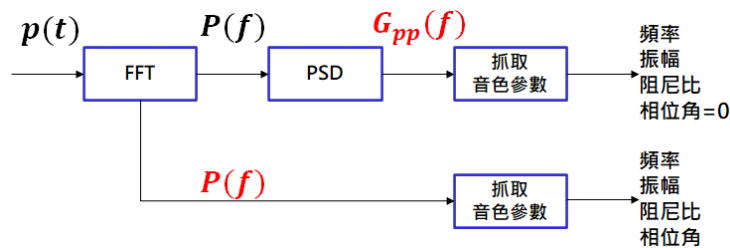


圖 1 音色參數檔案流程圖

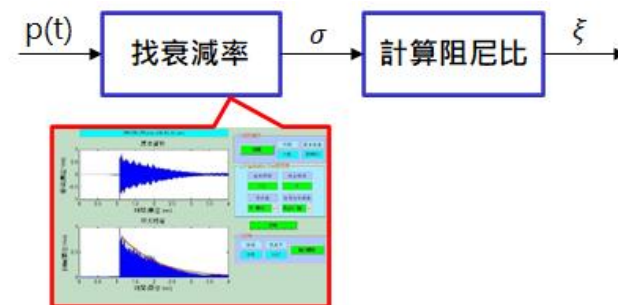


圖 2 衰減率計算阻尼比流程圖

另一種時間參數為包絡線方法，圖 3 為包絡線示意圖，第 1 張為原始聲音時間域訊號 $p(t)$ ，第 2 張為抓取 $p(t)$ 的包絡線 $E(t)$ ，第 3 張為重建聲音時間域訊號 $p_n(t)$ ，為了使 $p_n(t)$ 與 $p(t)$ 趨勢相同，因此取時間調整比例參數 $s(t)$ ，其方程式如下：

$$s(t) = E(t)/p_n(t) \quad (2)$$

接著將 $p_n(t)$ 與 $s(t)$ 進行振幅的調整，第 4 張圖為調整時間參數聲音重建時間域訊號 $p_{ns}(t)$ ，其方程式如下：

$$p_{ns}(t) = p_n(t) * s(t) \quad (3)$$

取得聲音音色參數與時間參數後，及可透過理論解析方程式進行聲音重建，使用具多頻率對數衰減正弦波[4]，其方程式如下：

$$p_n(t) = \sum_{r=1}^n P_r e^{-\xi_r \omega_r t} (\omega_r t - \phi_r) \quad (4)$$

$$\omega_r = 2\pi f_r \quad (5)$$

其中， $p_n(t)$ 為聲音重建之時間域訊號、 P_r 為聲音之振幅、 ξ_r 為聲音阻尼比、 ω_r 為聲音之頻率(rad/sec)、 ϕ_r 為聲音之相位角、 f_r 為聲音之頻率(Hz)。因透過 $G_{pp}(f)$ 方法與 $P(f)$ 方法所取得振幅單位為 dB，而據多頻率對數衰減正弦波，所使用的振幅單位為 Pa，因此需透過單位換算，其方程式如下：

$$P_r(P_a) = 10^{\frac{P_r(dB)}{20}} * P_{ref} \quad (6)$$

其中， P_{ref} 為聲音壓力參考值。

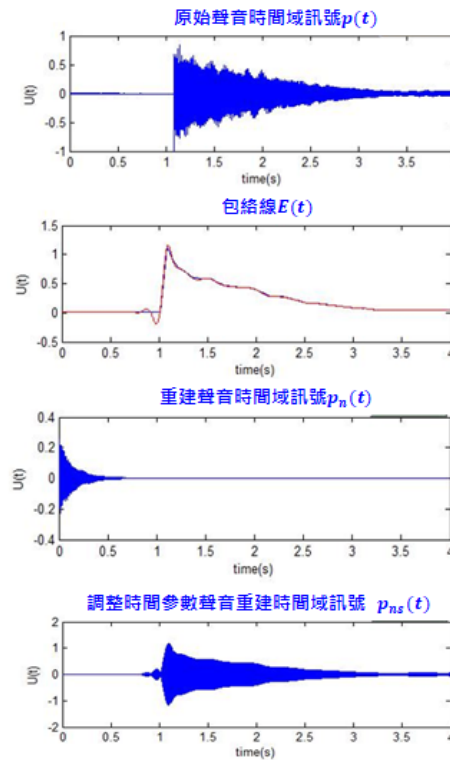


圖 3、包絡線示意圖

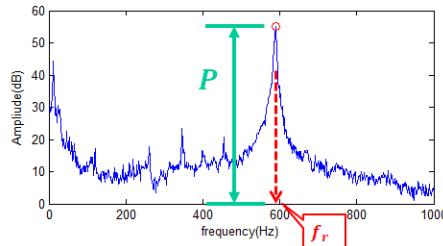


圖 4.聲音頻譜峰值振幅示意圖

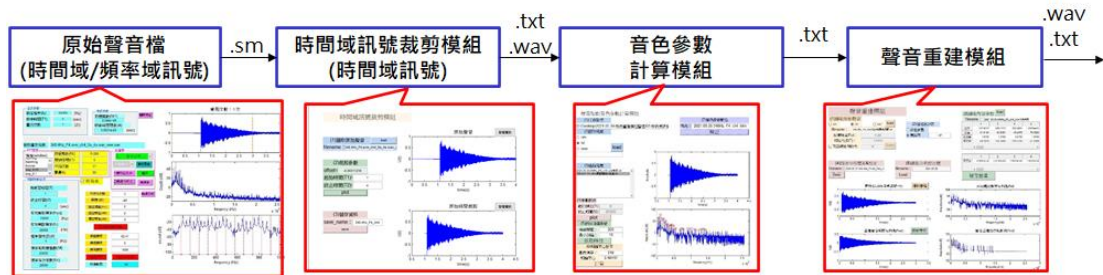


圖 5、聲音重建之檔案流程圖

圖 4 為聲音頻譜峰值振幅示意圖，峰值振幅為 P ，因每一個峰值頻率與阻尼比不同，因此帶入方程式(3)需進行換算，以單自由度假設推導，其方程式如下：

$$P_r = P * (4\pi\xi f_r) \quad (7)$$

三、聲音重建之程式發展

上節說明透過理論方程式進行聲音重建，需要有音色參數與時間參數，聲音頻譜特徵，包含峰值頻率、振幅、阻尼比與相位角，如果人工提取，當分析檔案一多，分析效率很低，因此本節介紹透過自動化分析模組，自動提取聲音頻譜特徵，並進行聲音重建。

圖 5 為聲音重建之檔案流程圖，首先是原始聲音檔，包含時間域訊號與頻率域訊號.sm 檔；因原始音檔可能雜訊很多，或者聲音檔案頭尾是沒有量測到聲音，此時可透過時間域訊號裁剪模組，進行時間裁減，可以輸出裁剪後時間域訊號.txt，或者聲音音檔.wav；接著使用音色參數計算模組，自動取得音色參數，包含峰值頻率、振幅、阻尼比與相位角，並輸出成音色參數.txt；最後透過聲音重建模組，進行聲音重建，輸出聲音重建的聲音時間域訊號.txt 及聲音音檔.wav。

3.1 時間域訊號裁剪模組

本小節介紹時間域訊號裁剪模組 ISOC 架構及使用介面，因原始聲音時間域訊號 $p(t)$ 可能有雜訊，或者沒有量測到聲音的時間區段，甚至音 DC 訊號 offset 造成時間域訊號有偏差狀況發生，因此可透過時間域訊號裁剪模組，進行時間域訊號裁剪及偏移，圖 6 為時間域訊號裁剪模組 ISOC 流程圖，其討論如下：

1. Input 輸入為實驗檔案.sm，是原始聲音時間域訊號 $p(t)$ 。
2. System 系統為時間域訊號裁剪模組，進行時間域訊號的裁剪，以及進行 offset 偏移調整。

- Output 輸出為時間域訊號裁剪並偏移完成後，原是聲音裁剪後時間域訊號 $p_t(t)$ ，並輸出對應的.txt 與聲音檔.wav。
- Control 控制變數，包含 offsert 設定，需設定時間域訊號偏移的數值；時間參數，起始時間至終止時間，設定時間裁剪範圍。

在實驗量測中，實驗儀器加速規貨麥克風，所量測到的震動訊號為電壓，再經由所設定的靈敏度進行換算，將電壓值轉換成物理量，在轉換過程中會產生 DC offset 狀況，產生 DC offset 會造成進行時間參數控制，衰減率與包絡線時，會有偏移現象，聲音特徵會變形。圖 7 為 DC offset 示意圖，原始聲音時間域訊號 $p(t)$ 距離 0 點，有偏移 0.5 的 p_{DC} ， $p(t)$ 的表示如下：

$$p(t) = p_{DC}(t) + p_{AC}(t) \quad (8)$$

因此透過 offset 調整 $p_{DC}(t)$ ，後續的原始聲音時間域訊號 $p(t)$ 的表示如下：

$$p(t) = p_{AC}(t) \quad (9)$$

圖 8 為時間域訊號裁剪模組，首先要選擇要進行時間域訊號裁剪的實驗檔案.sm，裁剪參數會出現預設的 offset 數值，可依照需求進行設定，輸入起始時間與終止時間，即可進行裁剪。也規劃音檔撥放功能，可以比對原始音檔與裁剪後音檔有無不同，確認裁剪後可以設定存檔名稱，儲存為時間域訊號 $p_t(t)$ 及聲音檔.wav，供後續進行音色參數計算模組使用。

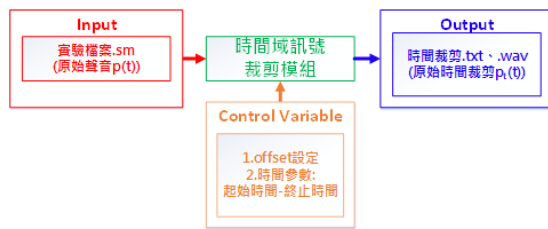


圖 6、時間域訊號裁剪模組 ISOC 流程圖

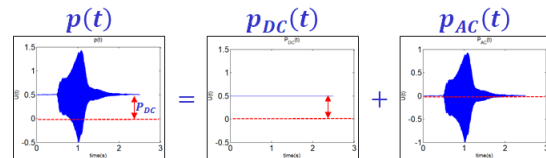


圖 7、DC offset 示意圖

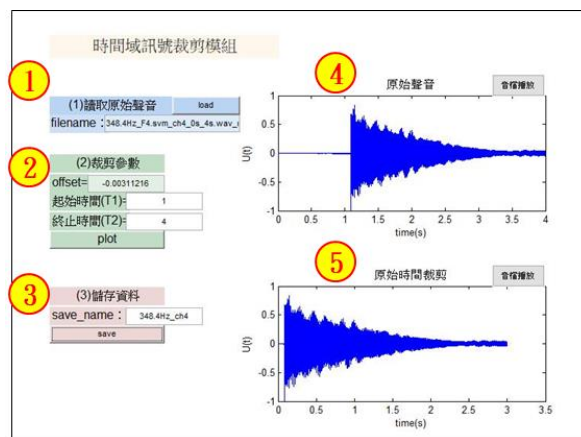


圖 8、時間域訊號裁剪模組

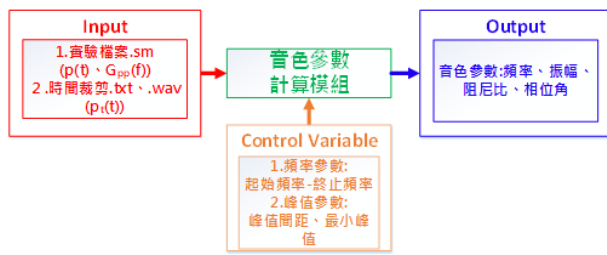


圖 9、音色參數計算模組 ISOC 流程圖

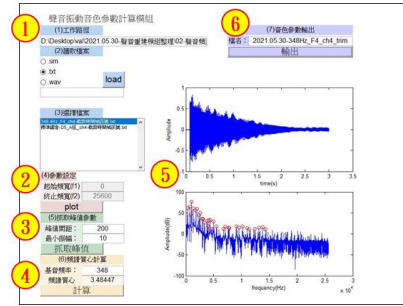


圖 10、音色參數計算模組

3.2 音色參數計算模組

音色參數模組，會自動抓取峰值參數，並計算出音色參數，包含峰值頻率、振幅、阻尼比與相位角。圖 9 為音色參數計算模組 ISOC 流程圖，其說明如下：

1. Input 輸入分為兩種，第一種為實驗檔案.sm，包含原始聲音時間域訊號 $p(t)$ 與頻率域訊號 $G_{pp}(f)$ ；第二種為時間裁剪後.txt 或.wav，為原始聲音裁剪後時間域訊號 $p_t(t)$ 。
2. System 系統為音色參數計算模組，以照輸入的檔案，自動判別使用功率頻譜密度函數(PSD)方法，或者傅立葉頻譜(Fourier spectrum)方法取得音色參數。
3. Output 輸出為音色參數，包含峰值頻率、振幅、阻尼比、相位角。
4. Control 控制變數為頻率參數或者峰值參數，頻率參數為控制抓取峰值的頻率範圍；峰值參數為自動抓取峰值時，可調整的參數，包含峰值間距與最小峰值振幅。

圖 10 音色參數計算模組，可選擇實驗檔案.sm，會自動讀取頻率域訊號 $G_{pp}(f)$ 資料，進行峰值抓取，因 $G_{pp}(f)$ 為實數，因此無法取得相位角，相位角以 0 取代。分析檔案選擇.txt 或.wav，為時間裁剪時間域訊號 $p_t(t)$ ，會先進行快速傅立葉轉換(FFT)，取得傅立葉頻譜 $P_t(f)$ ，進行 FFT 過程沒有進行重疊平均處理，所以訊號較雜亂，但是所抓取的峰值數量較多，且 $P_t(f)$ 為複數，可取得相位角數值。

3.3 聲音重建模組

本小節介紹聲音重建模組規劃及使用架構，圖 11 為聲音重建模組 ISOC 流程圖，其討論如下：

1. Input 輸入有三種，第一種為原始聲音.sm，包含 $p(f)$ 與 $G_{pp}(f)$ ；第二種時間裁剪.txt、.wav，為 $p_t(t)$ ，兩者檔案主要用於取得原始聲音包絡線，可用於控制時間參數；第三種是音色參數.txt，用於聲音重建參數輸入。
2. System 系統為聲音重建，透過控制時間參數與讀取音色參數，進行聲音重建。
3. Output 輸出為聲音重建.wav，可重複播放的音檔，以及聲音重建得時間域訊號 $p_{tn}(t)$ 與頻率域訊號 $P_{tn}(f)$ 。

4. Control 控制變數分為音色參數與時間參數。音色參數分為 $G_{pp}(f)$ 方法或者 $P(f)$ 方法取得，時間參數分為衰減率與包絡線。

接下來介紹聲音重建模組主要架構與使用程序說明，圖 12 為聲音重建模組，首先選擇要讀取的聲音資料，.sm、.txt、.wav，接著控制時間參數，選擇是否使用包絡線方法，接著為阻尼比的選擇，可透過半能量點法或著衰減率所計算得阻尼比。

音色參數選擇讀取 $p(f)$ 方法或者 $G_{pp}(f)$ 方法，模組會自動套用理論方程式進行聲音重建，並計算每一峰值所對應的衰減率或者阻尼比。在結果顯示上，上排會顯示原始聲音時間域與頻率域訊號，下排會顯示重建聲音之時間域訊號與頻率域訊號，可方便比對原始聲音與重建聲音的頻譜特徵。最後規劃儲存分析設定功能，方便後續重新讀取設定檔，使用介面圖片檔也會自動輸出。

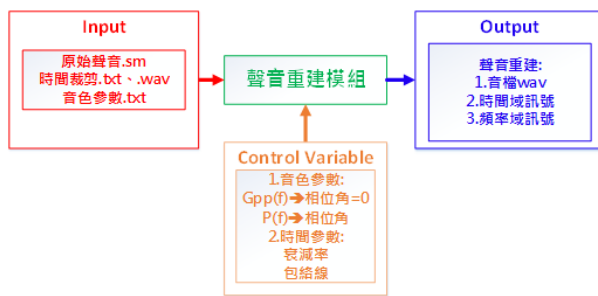


圖 11、聲音重建模組 ISOC 流程圖

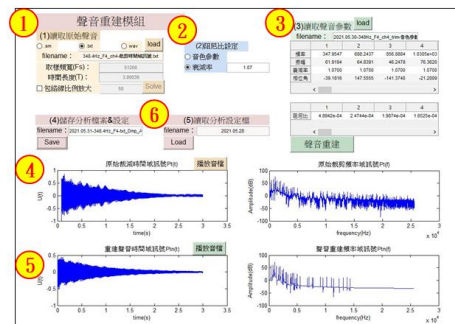


圖 12、聲音重建模組



(a) 中提琴



(b) 響鈴板



(c) 銅鐘

圖 13、聲音重建應用樂器

四、案例應用分析

本節透過實際樂器聲音，進行聲音重建，圖 13 為聲音重建應用樂器，圖 13(a) 為中提琴，圖 13(b) 為響鈴板，圖 13(c) 為銅鐘。

表 1 為不同樂器聲音重建示意圖比較表，分別為中提琴、響鈴板與銅鐘，從原始聲音包含時間域訊號與頻率域訊號觀察，時間域訊號，中提琴有持續拉奏的特徵，響鈴板與銅鐘為打擊樂器，其訊號從量測到聲音後，就會緩慢衰減；頻率域訊號，中提琴有明顯的簡諧倍頻特徵。

- 模式 A 使用 $G_{pp}(f)$ 方式取得音色參數，因為 $G_{pp}(f)$ 有經過重疊平均處理，因此峰值振幅會略微降低，造成阻尼比計算較大，聲音重建聽感衰減很快。

- 模式 B 時間參數有透過包絡線進行調整，中提琴有持續拉奏特徵，透過包絡線可以調整時間域訊號，使得取原始聲音最相似；但打擊樂器經由包絡線調整後，重建聲音不相似，推測音打擊樂無持續拉奏特徵，透過包絡線會調整會使音色失真，因此重建聲音聽感很差。
- 模式 C 是透過衰減率進行聲音重建，但衰減率只取時間域後段進行計算，因此中提琴聲音重建聽感衰減很快；打擊樂器重建效果佳，各模態音衰減率都相同，因此聲音重建聽感會比較平穩。
- 模式 D 是由傅立葉頻譜 $P(f)$ 方式取得音色參數，音 $P(f)$ 沒有經過重疊平均處理，因此與原始聲音最相似，適合應用於打擊樂器的重建，而中提琴因為有持續拉奏聲音，因此聽感較不相似。
- 模式 E 時間參數是使用包絡線，可以模擬持續拉奏的時間域特徵；對於打擊樂器，因包絡線調整時間域訊號，音色改變，聽感最差。模式 F 時間參數是使用衰減率，與模式 C 相同，中提琴的持續拉奏特徵無法模擬，因此中提琴聲音重建聽感差；而打擊樂器的聲音重建聽感會較平穩。

表 2 為不同樂器聲音重建聽感比較表，評分標準以重建的聲音聽感為主，分為最佳(5)、佳(4)、普通(3)、差(2)、最差(1)，進行不同樂器重建讀聲音結果比較。對於中提琴，有持續拉奏特徵的樂器，適合使用模式 B、E，時間參數設定為包絡線，進行聲音重建聽感最佳；但對於打擊樂器，不適合使用包絡線，推測包絡線調整時間域訊號時，變更了音色參數，因此打擊樂器聽感最差。

而打擊樂器適合使用模式 D、C、F 方式進行聲音重建，模式 D 的聲音重建效果最佳，時間參數無調整，阻尼比為半能量點法所得，各模態與原始聲音相進，聽感最相近；模式 C、F 時間參數設定衰減率，透過衰減率計算阻尼比，但因各模態衰減率相同，因此聽感較為平穩。

表 1 不同樂器聲音重建示意圖比較表

樂器種類			中提琴			鑼鈴板			銅鐘		
模式	音色參數	時間參數	時間域訊號	頻率域訊號	包絡線 $E(t)$ 衰減率 σ	時間域訊號	頻率域訊號	包絡線 $E(t)$ 衰減率 σ	時間域訊號	頻率域訊號	包絡線 $E(t)$ 衰減率 σ
原始	-	-			-			-			-
重建聲音	A	$G_{PP}(f)$	無設定			-			-		
	B	$G_{PP}(f)$	$E(t)$								
	C	$G_{PP}(f)$	σ								
	D	$P(f)$	無設定			-			-		
	E	$P(f)$	$E(t)$								
	F	$P(f)$	σ								

表 2 不同樂器聲音重建聽感比較表

	模式	音色參數	時間參數	中提琴	響鈴板	銅鐘	適用建議
重建聲音	A	$G_{pp}(f)$	無設定	2	2	2	無
	B	$G_{pp}(f)$	$E(t)$	5	1	1	持續拉奏樂器
	C	$G_{pp}(f)$	σ	2	4	4	打擊樂器
	D	$P(f)$	無設定	2	5	5	打擊樂器
	E	$P(f)$	$E(t)$	4	1	1	持續拉奏樂器
	F	$P(f)$	σ	2	4	4	打擊樂器

五、結論

本文開發樂器聲音重建之自動化分析模組，包含了時間域訊號裁剪模組、音色參數計算模組與聲音重建模組，並應用於不同樂器進行聲音重建，其結論如下：

1. 建立一整套聲音重建自動化分析模組，可快速的從原始實驗數據.sm，進行數據分析、抓取峰值、計算音色參數、計算時間參數，並進行聲音重建，加快樂器聲音聲音探討。
2. 透過理論解析進行聲音重建，將聲音特徵區分為音色參數與時間參數，並提出自身功率頻譜(PSD)方法與傅立葉頻譜($P(f)$)方法進行音色參數計算；時間參數分為衰減率與包絡線。
3. 最後應用於中提琴、響鈴板與銅鐘的聲音重建，結果顯示，中提琴有持續拉奏特徵，適合透過包絡線進行聲音重建，打擊樂器則適合使用衰減率進行聲音重建。
4. 未來可應用此聲音重建自動化分析模組，進行不同樂器的分析及重建，取得對應的聲音特徵，並能應用於不同樂器聲音的組合。

六、參考文獻

1. Hyeonjun, P., Lee, B., and Kim, D., "Violin Musical Tone Analysis using Robot Finger", Procedia Science, Vol. 94, pp. 398-403, 2016.
2. Debut, V., Carvalhom, M., Figueiredo, E., Antunes, J., and Silva, R., "The Sound of Bronze : Virtual Resurrection of a Broken Medieval Bell", Journal of Heritage, Vol. 19, pp. 544-554, 2016.
3. 王栢村，蘇集銘，中華民國音響學會第十九屆學術研討會論文集，「吉他弦之振動與聲音特性探討」，台南，論文編號：A7，2006。
4. 王栢村，陳品豪，吳銘峰，董時沛，中華民國音響學會第二十三屆學術研討會論文集，「聲音模擬程式之發展與應用」，台中，第 205-212 頁，2010。
5. Ilmoniemi, M., Valimaki, V., and Huutilainen, M., "Subjective Evaluation of Musical Instrument Timbre Modifications", Joint Baktic-Nordic Acoustics Meeting, pp. 1-6, 2004.