

# 行政院國家科學委員會專題研究計畫 成果報告

## 古箏音樂的音色分析與演奏表現資料探勘 研究成果報告(精簡版)

計畫類別：個別型  
計畫編號：NSC 98-2221-E-216-030-  
執行期間：98年08月01日至99年07月31日  
執行單位：中華大學資訊工程學系

計畫主持人：劉志俊  
共同主持人：張儷瓊  
計畫參與人員：碩士班研究生-兼任助理人員：廖昶盛  
碩士班研究生-兼任助理人員：邱繼正  
碩士班研究生-兼任助理人員：張柏峰

處理方式：本計畫可公開查詢

中華民國 99年10月29日



# 行政院國家科學委員會專題研究計畫成果報告

## 古箏音樂的音色分析與演奏表現資料探勘

Automatic Analysis of Music Performance and  
Style Data Mining of Digital Music

計畫編號：NSC 98-2221-E-216-030

執行期限：98年08月01日至99年07月31日

主持人：劉志俊 中華大學資訊工程學系

共同主持人：張儷瓊 國立臺灣藝術大學中國音樂學系

計畫參與人員：江俊賢、沈瑞祥、廖昶盛、邱繼正、張柏峰

鄭中皓、余俊賢、張誼賢、陳奕安、鄭聿哲、黃志銘

中華大學資訊工程學系

### 一、中文摘要

多媒體資料的內涵式分析(content-based analysis)為近年來多媒體領域重要的研究主題之一。在音樂內涵分析方面，以往的研究大都以西方樂器為研究對象，而古箏是我國固有之民族樂器，其演奏美學與曲譜特性與西方樂器差異極大，值得深入進行探討其內涵分析技術與演奏技法的特色。本計畫以古箏作韻基本技法中的顫音為研究對象，討論其起音點偵測、基頻追蹤、音色分析、泛音組成等內涵分析技術，以及適用於古箏的顫音模型。並對空顫、按顫、游搖等演奏技法在音高、音量、音色等方面的聲學性質進行探討，微觀地分析個別顫音的動態變化，檢視各種作韻的基本形態與手法。

**關鍵詞：**古箏,顫音,起音點偵測,基頻追蹤,音色分析

### 二、緣由與目的

音樂資料的內涵式查詢與分析是目前資訊檢索、資料庫與多媒體訊號處理領域的焦點研究之一。目前此主題的研究對象，以西方18世紀以來的主要樂器，如鋼琴、小提琴、鼓等樂器為主。而古箏是我國固有之民族樂器，其演奏美學與曲譜特性與西方樂器差異極大，非常值得進行內涵式查詢與分析。

古箏是一項聲韻兼備的樂器，聲韻相合的特徵是古箏音樂重要的風格標記。「聲」指的是古箏的物理聲響，主要表現

了樂音的形質--音量、音高、音長及音色的總合。「韻」則指在樂音的基礎上所營造的音腔轉折，亦即在特定音長的時域範圍內，透過指端的操作賦予樂音音量、音高、及音色等方面的變化。樂人的做韻手法，基本上是以右手指端施力彈弦，激發琴弦的振動以製造古箏的基本物理音響；而後運用左手潤飾樂音，透過壓顫琴弦製造音頻的動態變化，刻畫出箏樂聲韻兼備的表現特質。

「韻」是箏樂特色的主要核心所在，然而中外的研究鮮少對古箏音韻進行科學性的分析。在顫音的分析技術方面，首先必須能精確地追蹤顫音的音高變化。以往樂音基頻追蹤的相關研究大都使用單一基頻偵測技術，但由於顫音的抖動性質、古箏撥奏的搭弦雜訊、前一個樂音的殘響等因素，造成單一基頻偵測法很難得到理想的結果。在本論文中，我們參考古箏的特型，先計算樂音的音澤向量來估算顫音的大致頻率範圍，運用 Tolonen 等人所提出的多重基頻偵測方法，考慮音頻連續性來選取正確的基頻值。我們針對取自代表性箏曲曹桂芬演奏《陳杏元和番》中的顫音與游搖音樣本，統計其各種指法的顫音參數，以期建構能自動偵測與分析古箏顫音的模式。

### 三、研究成果

#### ● 古箏顫音分析系統

本文所提出之古箏顫音分析系統整體架構

如圖1所示，說明如下。

- 起音點偵測：一首箏樂是由一連串的樂音所組成。要分析古箏顫音的內涵，首先必須將顫音的起音點(onset)與終音點(offset)標記出來。本文中，針對每個顫音的起音點，我們分別以人工手動編輯與系統自動偵測兩種方式標記其起音點，而以下一個樂音的起音點作為顫音的終音點。
- 基頻分析：分析顫音的根本技術是基頻分析。由於顫音的抖動性質、古箏撥奏的搭弦雜訊、前一個樂音的殘響等因素，造成一般基頻分析方法在古箏顫音的基頻追蹤產生很大的誤差。為解決此問題，我們提出一種限制式多重基頻追蹤法來量測古箏顫音的基頻。
- 顫音分析：對於每個選取的古箏顫音，先由古箏演奏專家仔細聆聽並與曲譜核對，對其是否有使用顫音技巧與可能彈奏技法做一專業上的預估。在自動顫音分析方面，我們利用限制式多重基頻追蹤法的基頻量測結果繪圖，再與人工專業預估結果比對主觀分析與客觀量測值是否一致。
- 顫音統計：根據顫音模型，對每個顫音自動計算其波峰位置、波谷位置、顫動範圍、顫動頻率平均值與變異度、音準、顫動模式等顫音統計值。
- 泛音分析：分析各個顫音的組成泛音，其振幅隨著頻率振動的變化方式，以及在整個顫動過程的泛音能量變化模式。
- 力度分析：在力度分析方面，我們計算其每個音框中n個音訊取樣之振幅xi的根均方值xrms來表示此音框的能量。

$$x_{rms} = \sqrt{\frac{\sum_{i=1}^n x_i^2}{n}} \quad (1)$$

- 音色分析：在音色分析，我們先利用快速傅立葉轉換計算每個音框的頻譜Xj：

$$X_j = \sum_{k=0}^{N-1} x_k e^{-\frac{2\pi j}{N}kj} \quad (2)$$

其中，N值為4096；傅立葉轉換窗函數為Hamming。計算頻譜之後，我們進一步計算其MFCC與MPEG-7頻譜輪廓特徵，來表示古箏滑顫音的音色在頻譜方面的表現，其計算方式詳述於5.3節。。

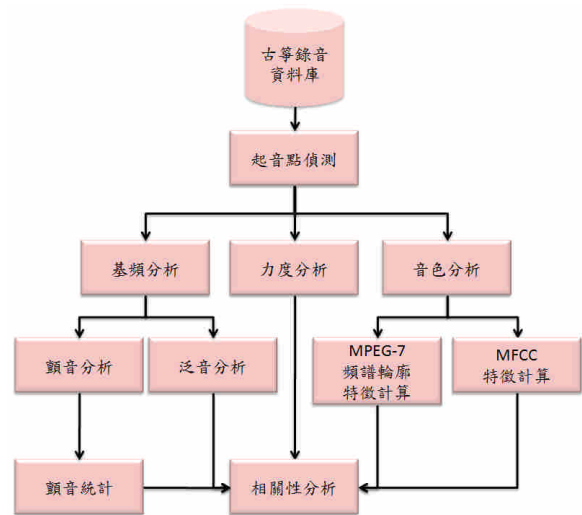


圖 1 古箏顫音分析系統架構圖

### ● 顫音基頻偵測技術

古箏顫音技法的主要變化在音高。所以能在餘音繚繞的古箏演奏環境中，精確地量測古箏顫音的音高變化，是分析顫音的基本議題。音高偵測(pitch detection)傳統上就是語音分析最重要的核心技術之一。常見的音高偵測技術大致上可以分為時域的自相關分析法(autocorrelational analysis)[1]，頻域的傅立葉分析法，以及頻譜係數法等。根據主要音高偵測比較結果，各種方法各有其優缺點，如頻域方法對泛音容易產生誤判、時域方法對低頻訊號有解析度上的問題等等，採用何種音高偵測方法必須視待分析的音訊特性而定。

由於古箏樂曲的演奏基本上是連續的樂音事件，所以一個顫音的起音往往是發生在前一個樂音的殘響環境下。所以以往使用單一基頻偵測的方法往往會偵測錯誤或有基頻不連續的問題。我們認為古箏顫音基頻偵測應該採用多重音高偵測的方法，以確保顫音的音高能正確被記錄，再配合使用音澤向量(chroma vector)估算顫音的大致彈奏音準範圍，經由選取適當的基頻值，可產生連續的顫音基頻量測值。

在多重音高偵測方法的選取方面，Tolonen與Karjalainen提出一種基於總和自相關函數(summary autocorrelation functions)來進行多重音高偵測的方法。有別於Meddis等人所提出的傳統時域音高偵測法所使用的多重頻帶聽覺濾波器(multichannel auditory filterbank)，Tolonen與Karjalainen認為這種架構需要耗用大量

計算，因此他們提出雙頻帶的架構來改善多重音高偵測的執行效率。實驗結果顯示其方法可以有效率地完成多重音高偵測，且偵測的準確率與傳統的多重頻帶音高偵測法相當。我們認為此架構很適合在餘音繚繞環境下準確偵測古箏顫音的音高。

本文使用的基頻偵測方法敘述如下：

- 步驟一：音框切割。古箏顫音音訊樣本為44.1k Hz取樣之錄音波形檔。在考量時間軸解析度與基頻量測值穩定度的平衡條件下，每一個音框(frame)大小為100ms，量測步進時值(hopping size)為10ms，亦即每秒量測基頻值100次。
- 步驟二：計算每一個音框的音澤向量(chroma vector)。
- 步驟三：雙頻道濾波。使用一個高通濾波器(highpass filter)與一個低通濾波器(lowpass filter)來進行雙頻道濾波，模擬人耳耳蝸的分頻處理方式，頻帶分割點設為1k Hz。分頻後的訊號進行半波整流(half-wave rectifier)。
- 步驟四：強化型加總式自相關分析(ESACF)。對半波整流後的高頻訊號與低頻訊號，各自使用自相關分析函數(autocorrelation function)來模擬聽覺神經系統的週期偵測功能。高頻自相關分析值與低頻自相關分析值相加後得到加總式自相關分析值。為降低泛音可能造成的誤判，加總式自相關分析值分別進行時間軸乘上2倍、3倍、4倍等，再與原自相關分析值相減，以降低2倍、3倍、4倍基頻上的峰值，得到較穩定清晰的加總式自相關分析基頻值，由此挑選最大三個峰值作為多重基頻的偵測結果。
- 步驟五：基頻追蹤。為提高偵測基頻值的穩定度，我們在加總式自相關分析的三個基頻峰值中選取頻率值落在預期頻率 $\pm 200$ 音分範圍之內，(預期頻率值由音澤向量與起音點八度音程能量最大頻帶產生)，且基頻值與前一音框的基頻值差距在100音分內的結果，以得到平滑化的古箏滑顫音基頻值。

#### ● 音樂表現的自動比較分析

在音樂領域中，經常有學者對於同一首曲子的不同演奏版本的音樂，進行詮釋比較分析。但因為傳統上以人工的方式比

較演奏表現差異是相當耗時的工作，難以進行大量而全面性的分析比較。因此，我們希望以電腦自動化的方式來協助演奏表現比較的進行。但在能夠進行演奏表現元素差異比較分析之前，音樂對齊(music alignment)是一項必要的核心技術。所謂音樂對齊，根據Orio的說法：“音樂對齊是將樂譜的音樂事件(即音符或樂音)與音樂訊號的時間軸中的每個時間點結合在一起。”由此可知，若要進行音樂表現差異分析比較，就要先對不同演奏版本的音樂之每個樂音將其對齊。但在傳統訊號處理領域中解決音樂對齊的問題並不容易，因此，我們將訊號轉換成符號式的音樂表示方式來進行音樂對齊。首先，我們將一段音樂訊號轉換成對應的音樂字串(music string)，把多重音樂對齊(multiple music alignment)的問題轉換成多重字串排比(multiple string matching)的問題。接著我們利用生物領域中為進行蛋白質字串排比，所發展出的漸進式多重序列排比(progressive multiple sequence alignment)技術，按照音樂的特性來修改比對分數計算所參考的相似矩陣(similarity matrix)及其計分方式，使其可以自動找到不同演奏版本的音樂的最佳對齊方式。

使用音樂自動多重排比方式進行音樂詮釋比較分析，跳脫了傳統音樂訊號處理方面的複雜問題，能夠自動排比多首不同演奏版本的音樂，解決人工比較分析所需大量反覆聆聽比對的耗時耗力狀況，藉由此方式我們得以進行大量而全面性的音樂比較分析。

本計畫提出之音樂演奏表現多重自動排比技術有以下方面之應用：

- 音樂作品比較分析：藉由音樂作品比較分析，我們可以從不同的演奏版本中，藉由速度、力度等的變化資訊，來比較不同詮釋版本之間的差異性之所在。
- 演奏表現特性具體呈現：我們可藉由演奏表現特性的定量化具體呈現，來分析每一位演奏家在處理每一段音樂時，其速度變化、力度變化以及音樂詮釋的張力等資訊，與其他演奏家演奏版本比較，找出一位演奏家有哪些獨特的表現性質。

音樂教育：藉由分析老師與學生的演奏版本之間的差異性，除了基本的音準與彈奏手法要求外，老師所演奏音樂的速度、力度等表現變化，都是具有音樂詮釋方面的深刻意義之珍貴指導。因此本論文提出的技術，可以提供音樂教育方面的指引，讓學生了解應該如何改進其演奏方式，以達到老師的音樂詮釋要求。

### ● MP3 音訊指紋之設計

由於MP3高音質與高壓縮比的特性，人們很容易利用它在網路上散播數位音樂。雖說這種行為往往是非合法的，但是卻沒有很好的方法來嚇阻。我們希望能使用自動偵測MP3數位音樂傳輸的方法，可以自動根據一小段MP3音樂的指紋特徵來辨識其身份，以達到自動監控MP3數位音樂非法傳輸之目的。在本計畫中，我們提出能有效辨識MP3音樂身份的MP3音訊指紋的設計及其搜尋比對方法。MP3音訊指紋所使用的音訊特徵包含梅爾頻率倒頻譜係數、MPEG-7音訊描述、音澤向量、修正式離散餘弦轉換係數等四大類特徵。為提高音訊特徵值計算的效率，這些特徵值是在壓縮領域中進行計算。之後對四種音訊指紋特徵進行主要成份分析，並依照分析的結果，使用QUC-Tree索引結構進行最近鄰居搜尋，找到包含此未知MP3音樂片段的反向索引。然後再根據此反向索引做查詢，來快速找出歌曲的身份。

## 四、計畫成果自評

本計劃目前的研究成果至已發表美國專利一篇[1]，期刊論文兩篇[2][3]，研討會論文五篇 [4][5][6][7][8]。

## 五、參考文獻

- [1] “Method of Online Detecting MP3 Music Transmissions and an Architecture Thereof,” 美國專利 US7765598 B2, July 2010.
- [2] 王鴻文、劉志俊，“MP3 音樂的聆賞情緒自動分類，”（期刊論文投稿已接受）。
- [3] 沈瑞祥、劉志俊、張儷瓊，“箏樂滑音之自動識別，”（期刊論文投稿已接受）。
- [4] 張儷瓊、劉志俊、林雅琇，“河南箏樂「滑顫音」之聲學分析內涵式分析技術在古

箏樂音上的應用研究，” 亞太民族音樂學會第十五屆國際學術研討會，新疆烏魯木齊，中國，2010.

- [5] 鄭中皓、余俊賢、鄭聿哲、陳奕安、劉志俊，“基於 IEEE1599 標準的 MP3 音樂自動斷句系統，” 2010 民生電子研討會，2010.
- [6] 張柏峰、劉志俊，“高效能 MP3 音訊指紋之設計，” 2010 民生電子研討會，2010.
- [7] 邱繼正、劉志俊，“使用多重序列排比技術來進行音樂表現的自動比較分析，” 2010 民生電子研討會，2010.
- [8] 劉志俊、張儷瓊，“箏樂顫音的內涵式分析，” 6th International WOCMAT and New Media Conference, Taoyuan, Taiwan, November 12-13, 2010.

# 國科會補助計畫衍生研發成果推廣資料表

日期 2010年10月29日

<p>國科會補助計畫</p>	<p>計畫名稱: 古箏音樂的音色分析與演奏表現資料探勘                  計畫主持人: 劉志俊                  計畫編號: 98 -2221-E -216 -030 - 學門領域: 資料庫系統及資料工程</p>		
<p>研發成果名稱</p>	<p>(中文) Method of Online Detecting MP3 Music Transmissions and an Architecture Thereof, 美國專利US7765598 B2, July 2010                  (英文) Method of Online Detecting MP3 Music Transmissions and an Architecture Thereof</p>		
<p>成果歸屬機構</p>	<p>中華大學</p>	<p>發明人 (創作人)</p>	<p>劉志俊</p>
<p>技術說明</p>	<p>(中文) 本創作提出一種MP3音樂傳輸偵測與身份識別之方法與系統。有別於一般使用浮水印(watermarks)的著作權保護(copyright protection)機制，本系統使用封包捕捉(packet capturing)的技術，攔截網路上的可疑檔案傳輸(file transfer)，利用MP3標準檔案格式的框架表頭(frame header)的資訊，初步判讀此檔案是否為MP3音樂檔案。若是MP3音樂檔案，則繼續用封包捕捉技術攔截一定長度的MP3音樂檔案，接著對此攔截到的部分MP3音樂檔案開始片段，以MP3特徵值擷取程式(MP3 feature extractor)取出其MP3特徵值(MP3 features)。最後與MP3音樂資料庫(MP3 databases)中的所有版權音樂的MP3特徵值作相似性比對(similarity measurement)。若相似度小於一定門限值，則將檔案傳輸的時間、傳輸雙方的IP位址、檔案名稱等資訊記錄在非法傳輸日誌中，作為取締非法下載MP3音樂的根據。或是送出終止連線的封包(reset packet)，及時中斷此項MP3音樂非法傳輸。</p> <p>(英文) We propose an approach to retrieve MP3 music objects based on their content. In our approach, the coefficients extracting from the output of the polyphase filters are used to compute the MP3 features for indexing the MP3 objects. We also propose an MP3 similarity measuring function to provide users the ability to approximately retrieve the desired MP3 objects.</p>		
<p>產業別</p>	<p>資訊服務業；休閒、娛樂服務業</p>		
<p>技術/產品應用範圍</p>	<p>可能應用之產業別(Applicable industries):                  音樂產業                  數位音樂網站                  網路硬體製造商                  可能應用之產品(Applicable Products)</p>		
<p>技術移轉可行性及預期效益</p>	<p>音樂產業                  數位音樂網站                  網路硬體製造商</p>		

註：本項研發成果若尚未申請專利，請勿揭露可申請專利之主要內容。





98 年度專題研究計畫研究成果彙整表

計畫主持人：劉志俊		計畫編號：98-2221-E-216-030-					
計畫名稱：古箏音樂的音色分析與演奏表現資料探勘							
成果項目		量化			單位	備註（質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等）	
		實際已達成數（被接受或已發表）	預期總達成數（含實際已達成數）	本計畫實際貢獻百分比			
國內	論文著作	期刊論文	2	2	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	5	5	100%		
		專書	0	0	100%		
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（本國籍）	碩士生	3	3	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
國外	論文著作	期刊論文	0	0	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	0	0	100%		
		專書	0	0	100%		章/本
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	1	1	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（外國籍）	碩士生	0	0	100%	人次	
		博士生	0	0	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		

<p>其他成果 (無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)</p>	<p>無</p>
--	----------

	成果項目	量化	名稱或內容性質簡述
科 教 處 計 畫 加 填 項 目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與(閱聽)人數	0	



# 國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

## 1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

達成目標

未達成目標（請說明，以 100 字為限）

實驗失敗

因故實驗中斷

其他原因

說明：

## 2. 研究成果在學術期刊發表或申請專利等情形：

論文： 已發表  未發表之文稿  撰寫中  無

專利： 已獲得  申請中  無

技轉： 已技轉  洽談中  無

其他：（以 100 字為限）

主要研究成果本年度獲得美國專利一項，期刊論文兩篇，研討會論文四篇

## 3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以 500 字為限）

由於 MP3 高音質與高壓縮比的特性，人們很容易利用它在網路上散播數位音樂。雖說這種行為往往是非法的，但是卻沒有很好的方法來嚇阻。我們使用自動偵測 MP3 數位音樂傳輸的方法，可以自動根據一小段 MP3 音樂的指紋特徵來辨識其身份，以達到自動監控 MP3 數位音樂非法傳輸之目的。

我們對不同演奏版本的音樂錄音進行起音點偵測，然後將兩個起音點間的音樂訊號以一個音澤向量來表示其主要內容，接著將一段音樂轉換為對應的音澤字串。為了自動進行音樂表現比較分析，我們使用多重序列排比來對齊不同演奏版本的音澤字串。完成音澤字串對齊後，便可進一步在單一樂音、樂句、或是全曲尺度上，進行力度與速度等音樂表現方面的自動比較分析。

古箏是我國固有之民族樂器，其演奏美學與曲譜特性與西方樂器差異極大，值得深入進行探討其內涵分析技術與演奏技法的特色。本文以古箏作韻基本技法中的顫音為研究對象，討論其起音點偵測、基頻追蹤、音色分析、泛音組成等內涵分析技術，以及適用於古箏的顫音模型。並對空顫、按顫、游搖等演奏技法在音高、音量、音色等方面的聲學性質進行探討，微觀地分析個別顫音的動態變化，檢視各種作韻的基本形態與手法。