

行政院國家科學委員會專題研究計畫 成果報告

古琴與古箏音樂資料庫的建構, 內涵分析與查詢 研究成果報告(精簡版)

計畫類別：個別型
計畫編號：NSC 99-2221-E-216-046-
執行期間：99年08月01日至100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：游智為
碩士班研究生-兼任助理人員：徐志豪
碩士班研究生-兼任助理人員：劉哲辰

公開資訊：本計畫可公開查詢

中華民國 100年10月26日

行政院國家科學委員會專題研究計畫成果報告

古琴與古箏音樂資料庫的建構,內涵分析與查詢

The Construction, Content-Based Analysis, and Retrieval of

Guqin and Zheng Music Databases

計畫編號：NSC 99-2221-E-216-046

執行期限：99 年 08 月 01 日 至 100 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：邱繼正、張柏峰、鄭中皓、黃志銘、游智為、劉哲辰

蔡君彥、徐志豪、程弘錡、張智凱、徐詩媛、蕭秀婉

中華大學資訊工程學系

一、中文摘要

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。而古琴與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。本計畫以古琴與古箏經典曲譜資料庫建構、減字譜數位化儲存、古琴/古箏的顫音(吟猱)分析、古琴與古箏內涵式指法查詢與樂句分析等方面來探討古琴與古箏樂曲文化的特色。

關鍵詞：古琴,古箏,顫音,吟猱,起音點偵測,基頻追蹤,減字譜

二、緣由與目的

西方樂譜是以五線譜的方式來記載音樂，而電腦化的樂譜檔案，長久以來不管在學術界或是產業界都呈現非常分歧的現象。在“Beyond MIDI”一書中描述了超過 20 種以上不同的樂譜編碼方式；而各種樂譜軟體，如 Finale 與 Sibelius 各有其專屬的檔案格式，彼此間缺乏共通的轉換標準。

隨著網際網路的興起，在網路上交易數位樂譜檔案的需求日益增加。目前數位樂譜的檔案主要以 pdf 為主。但 pdf 格式的樂譜只是記錄了樂譜的印刷資訊，欠缺記錄音樂方面的涵義。另一方面，MIDI 格式雖然成功地用於電腦合成音樂的領域，但 MIDI 檔所能記錄的資訊與樂譜比較顯得

十分貧乏。

由於 XML 所具備可擴充定義標準的特性，且文字檔的形式同時適合人類閱讀與機器剖析，所以用 XML 來表示樂譜資訊的幾種方法中，以 MusicXML 發展最為成功。其主要原因有二：第一，MusicXML 以 MuseData 以及 Humdrum 樂譜格式作為設計的參考對象。而 MuseData 與 Humdrum 皆已發展多年，各自建立了數量龐大的樂譜典藏，所以在樂譜表示的功能上設計十分成熟而完備。第二，MusicXML 與業界主要的五線譜編輯軟體合作，避免了 SMDL 之類標準，理論完備但缺乏業界支持的狀況。

然而 MusicXML 是以十七世紀以來西方音樂的樂譜作為設計表示之範疇，設計上並不支援中國傳統樂器對記譜的特別需求。對古琴與古箏曲譜而言，有表韻指法(左手)與表聲指法(右手)的演奏指示需求，尤其是古琴千年的文化傳承之下，所產生的龐大減字譜琴譜珍貴典藏，在曲譜的數位化表示模型方面需要特別的設計，才能完整的儲存與呈現古琴與古箏的曲譜。

顫音(vibrato)是 17 世紀西方音樂所發展出的裝飾音(ornament)技術，最早應用在中提琴、長笛與人聲演奏表現上，隨後成為重要的音樂表情的表現重要手段之一。在 19 世紀，顫音成為廣義音色的一部分，大部分的西方樂器都有可以讓演奏者控制顫音的方法。

在西方音樂的顫音分析，最早是由 Seashore 與其同儕在 Iowa 大學心理學實驗室開始了開創性的一系列顫音分析研究。Seashore 等人由顫音訊號的分析與顫動感知量測兩方面著手進行，建立了最早的顫音分析模型，明確地標示出顫音本質上包含音高(pitch)、音量(intensity)、與音色(timbre)三個層面的振動現象。而顫動範圍(vibrato extent)與顫動頻率(vibrato rate)等關於顫音模型的重要參數，亦是由 Seashore 建立明確定義。

《五知齋琴譜》說「五音活潑之趣，半在吟猱」，可見吟猱等顫音在古琴音樂中的重要性。古琴吟猱音的劃分方式，各家說法並不一致。根據《梅庵琴譜》，古琴顫音可以分為吟與猱兩大類。吟音的指法描述為「得音後將指急急搖動，約三四轉，猶吟哦之餘韻，上下不出四五分，雖動搖不離其位。要至勻至實，不可輕搖漫動。然必鬆活，得靈動之機為妙。」。猱音的指法描述為「得音後稍偏下緩緩搖動約一二轉，如猱狻之悲啼，較吟稍大而緩。」。吟可以再細分為游吟、飛吟、落指吟、長吟、細吟等五種指法，而猱可以再細分為蕩猱與小猱兩種指法，可見吟猱音演奏法的細膩程度。

三、研究成果

● 顫音基頻偵測技術

古琴顫音技法的主要變化在音高。所以能在餘音繚繞的古琴演奏環境中，精確地量測古琴顫音的音高變化，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)，頻域的傅立葉分析法，以及頻譜係數法等。根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。

由於古琴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。我們認為古琴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琴顫音的音高。

本文使用的基頻偵測方法敘述如下：

步驟一：音框切割。古琴顫音音訊樣本為 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 ± 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。

● Music-BLOSUM 矩陣的計算

我們參考生物資訊領域中的區塊置換矩陣 BLOSUM(BLOcks Substitution Matrix)的基本建構原理，考慮音樂的特性，提出稱為 Music-BLOSUM (Music BLOcks Substitution Matrix) 矩陣的建構方式。此矩陣的建立過程主要可分為四個步驟，各個步驟的計算方式說明如下。

步驟 1：計算每種音澤符號的出現機率

我們將每個樂句的音樂訊號轉換為對應的音澤字串後，將多個演奏版本的音澤字串根據起音點開始位置對齊後，產生對應之音樂區塊(*music blocks*)。圖 1 為一個由 4 個不同演奏版本 (EDEDB、ECEDC、EDED、以及 ECECB)，長度為 5 個音澤符號的音樂區塊。我們首先計算一個音樂區塊中每種音澤符號的出現機率。

E	D	E	D	B
E	C	E	D	C
E	C	E	D	C
E	C	E	C	B

圖 1. 一個包含 4 個不同演奏版本長度為 5 的音樂區塊範例

假設一個由 k 個演奏版本構成長度為 l 的音樂區塊為 $Block = \{S_1, S_2, \dots, S_k\}$ ，其中 S_i 為第 i 個演奏版本的音澤字串。令 $S_i = C_{i,1}, C_{i,2}, \dots, C_{i,l}$ ，其中 $C_{i,j}$ 為 S_i 中的第 j 個音澤符號。對每個音澤符號 $X \in Chroma = \{C, C\#, D, D\#E, F, F\#, G, G\#, A, A\#, B\}$ 而

言， X 在音樂區塊 B 中的出現機率 $P(X)$ 為

$$P(X) = \frac{\sum_{i=1}^k count(X, S_i)}{\sum_{Y \in Chroma} \sum_{i=1}^k count(Y, S_i)} \quad (1)$$

其中 $count(X, S_i)$ 表示 X 在 S_i 中出現的次數。對圖 1 的例子而言，音澤符號 B 、 C 、 D 、 E 出現的機率 $P(B)$ 、 $P(C)$ 、 $P(D)$ 、與 $P(E)$ 分別為

$$P(B) = \frac{1+0+0+1}{2+6+4+8} = \frac{2}{20} = 0.1 \quad (2)$$

$$P(C) = \frac{0+2+2+2}{2+6+4+8} = \frac{6}{20} = 0.3 \quad (3)$$

$$P(D) = \frac{2+1+1+0}{2+6+4+8} = \frac{4}{20} = 0.2 \quad (4)$$

$$P(E) = \frac{2+2+2+2}{2+6+4+8} = \frac{8}{20} = 0.4 \quad (5)$$

步驟 2：計算兩種音澤符號隨機配對發生變異的機率

在多重音樂表現字串排比過程中，我們需要計算兩兩音澤符號配對出現變異的機率，而此機率值必須與隨機發生的配對變異機率進行比較，才能顯現出此種配對變異的相對機率大小。因此，我們必須先計算在音樂區塊中任兩種音澤符號隨機配對發生變異的機率。

假設任意兩個音澤符號 $X, Y \in Chroma$ 。在音樂區塊 $Block$ 中 X 與 Y 隨機發生變異的機率 $P(X, Y | random)$ 為

$$P(X, Y | random) = \begin{cases} P(X)^2, & \text{if } X = Y \\ 2P(X)P(Y), & \text{if } X \neq Y \end{cases} \quad (6)$$

上式表示若 X 沒有發生變異時，音樂區塊中任兩個音澤字串中的每個音澤符號進行兩兩比對後，隨機出現 XX 配對的機率為 $P(X) * P(X) = P(X)^2$ ；而若 X 發生變異轉為 Y 時，有可能發生 XY 配對或是 YX 配對，故機率值為兩種可能的配對機率之和 $P(X) * P(Y) + P(Y) * P(X) = 2P(X)P(Y)$ 。

對圖 5 的例子而言， $P(B,B|random)=P(B)^2=0.01$ ， $P(B,C|random)=2*0.1*0.3=0.06$ 。

步驟 3: 計算兩種音澤符號在音樂區塊中實際出現配對變異的機率

兩條長度為 l 音澤字串進行對齊，會產生 l 個音澤符號配對。若音樂區塊 *Block* 由 k 個演奏版本構成，則一共有 $C_k^2 = \frac{k(k-1)}{2}$ 次音澤字串對齊，故產生 $M = \frac{lk(k-1)}{2}$ 次音澤符號配對。故兩個音澤符號 $X, Y \in Chroma$ 在音樂區塊 *Block* 中實際出現配對變異的機率 $P(X, Y | Block)$ 為

$$P(X, Y | Block) = \frac{\text{count}(X, Y | Block)}{M} \quad (6)$$

其中 $\text{count}(X, Y | Block)$ 表示 XY 或 YX 配對在音樂區塊 *Block* 中實際出現配對變異的次數。

對圖 1 的例子而言，音樂區塊由 4 個演奏版本構成，故一共需要進行 $C_4^2 = \frac{4(4-1)}{2} = 6$ 次音澤字串對齊，每次音澤字串對齊產生 5 個音澤符號配對，所以總共有 $M = 5*6 = 30$ 次音澤符號配對。

圖 1 的音樂區塊中，BB 配對僅出現一次，故

$$P(B, B | Block) = \frac{\text{count}(B, B | Block)}{M} = \frac{1}{30} \quad (7)$$

圖 1 的音樂區塊中，BC 或 CB 配對共出現 4 次，故

$$P(B, C | Block) = \frac{\text{count}(B, C | Block)}{M} = \frac{4}{30} \quad (8)$$

步驟 4: 計算 Music-BLOSUM 矩陣

由圖 1 的範例中我們可以發現，兩種音澤符號在音樂區塊中實際出現配對變異的機率，往往與兩種音澤符號隨機配對發生變異的機率出現極大的差異。例如 BC 實際出現配對變異的機率 $P(B,C|Block) = 0.1333$

高於 $P(B,C|random) = 0.06$ 。我們希望音樂演奏表現相似矩陣能夠反應出在真實的演奏比較中，某一個音的出現兩種不同音澤符號的機率比隨機出現的機率高出多少倍，故可取兩種變異機率的比值來衡量真實變異的機率相對高低。此外，由於兩種比值的範圍很大，我們取比值的對數值，放大兩倍，並四捨五入來產生音樂演奏表現相似矩陣。

因此，任意兩個音澤符號 $X, Y \in Chroma$ ，在音樂區塊 *Block* 中 X 與 Y 表現的相似性為

$$\text{Music-BLOSUM}(X, Y) = \text{round}(2\log_2 \frac{P(X, Y | Block)}{P(X, Y | random)}) \quad (9)$$

舉例而言，圖 1 中的 BC 音澤符號配對的 $\text{Music-BLOSUM}(B, C)$ 為

$$\begin{aligned} \text{Music-BLOSUM}(B, C) &= \text{round}(2\log_2 \frac{P(B, C | Block)}{P(B, C | random)}) \quad (10) \\ &= \text{round}(2\log_2 \frac{0.1333}{0.06}) \\ &= 2 \end{aligned}$$

表示 BC 音澤符號配對的真實出現機率約為隨機出現之機率兩倍。

● 減字譜字符拓樸樹的 XML 編碼

我們參考漢字拓樸結構分析的研究成果，歸納出減字譜字符的八種不同的拓樸結構，以此為編碼設計的基礎，利用 XML 的編碼技術，擴充 MusicXML 五線譜的編碼語法，提出一套稱為 QinXML 的古琴減字譜編碼方式，可以完整記錄減字譜字符所攜帶的演奏指示資訊，以及合成減字譜字符所需的組成結構空間關係。由於 XML 可以標記減字譜字符拓樸樹中各結點之語意的特性，使得我們的編碼方式較其他古琴編碼相關研究，在左右手指法的涵義標記方面更為明確清楚，而不只是記錄合成減字譜字符的資訊，此方法非常有利於減字譜樂譜查詢與分析等後續研究工作的進行。而以往古琴編碼相關研究所提出之語法樹編碼方式，雖然編碼較精簡，但各節點的彈奏涵義不夠明確，且其拓譜結構較為簡單，未能完整表達減出字譜字符的八種不

同的拓樸結構之差異。

四、計畫成果自評

本計劃目前的研究成果至已發表期刊論文兩篇[1][2]，EI 研討會論文兩篇 [3][4]，國內研討會論文七篇 [5][6][7][8][9][10][11]。

五、參考文獻

- [1] 王鴻文、劉志俊，“MP3 音樂的聆賞情緒自動分類,” *Journal of Information Technology and Applications*, Vol. 4, No. 4, pp. 160-171, Dec. 2010.
- [2] 張儷瓊、劉志俊、林雅琇，“河南箏樂「滑顫音」之聲學分析—以曹桂芬演奏的五首曹派箏曲為例,” *藝術學報*, No.88, pp.313-348, April 2011.
- [3] Chih-Chin Liu, “Automatic Phrase Segmentation of MP3 Songs Based on the Technique of Breath Sound Detection,” *7th IEEE International Conference on Information Technology and Application (ICITA 2011)* (EI), Sydney, Australia, 2011.
- [4] Chih-Chin Liu and Po-Feng Chang, “An Efficient Audio Fingerprint Design for MP3 Music,” *9th ACM International Conference on Advances in Mobile Computing and Multimedia (MoMM-2011)* (EI), Ho Chi Minh City, Vietnam, 2011.
- [5] 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*, 台南, 台灣, 2010. (獲研討會最佳論文獎)
- [6] 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*, 台南, 台灣, 2010. (獲研討會最佳論文獎)
- [7] 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.
- [8] 劉哲辰、邱繼正、劉志俊，“Music-BLOSUM 音樂演奏表現相似矩陣與多重音樂自動對齊,” *第六屆智慧生活科技研討會*, 台中, 台灣, June 2011.
- [9] 黃志銘、劉志俊，“呼吸聲偵測技術在

MP3 音樂自動斷句的應用,” *數位生活科技研討會*, 雲林, 台灣, July 2011. (獲研討會最佳論文獎)

- [10] 程弘錡、張智凱、徐詩媛、蕭秀婉、劉志俊，“MP3 國語流行音樂自動斷句與詞曲同步撥放系統,” *2011 民生電子研討會*, 台中, 台灣, 2011.
- [11] 游智為、劉志俊，“使用 XML 來建立古琴減字譜之數位編碼,” *NCS 2011 全國計算機會議*, 2011.

國科會補助計畫衍生研發成果推廣資料表

日期:2011/10/26

國科會補助計畫	計畫名稱: 古琴與古箏音樂資料庫的建構, 內涵分析與查詢
	計畫主持人: 劉志俊
	計畫編號: 99-2221-E-216-046- 學門領域: 資料庫系統及資料工程
無研發成果推廣資料	

99 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：99-2221-E-216-046-				計畫名稱：古琴與古箏音樂資料庫的建構, 內涵分析與查詢	
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	三篇論文獲得國內研討會最佳論文獎
		研究報告/技術報告	0	0	100%		
		研討會論文	7	7	100%		
		專書	0	0	100%		
	專利	申請中件數	0	0	100%	件	
		已獲得件數	1	1	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	2	2	100%		
		專書	0	0	100%	章/本	
	專利	申請中件數	0	0	100%	件	
		已獲得件數	1	1	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 (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>	三篇論文獲得國內研討會最佳論文獎
--	------------------

	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）

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）

古琴與古箏的內涵式分析對珍貴文化的保存意義重大

MP3 指紋相關研究獲得中華民國與美國專利