

行政院國家科學委員會專題研究計畫 成果報告

數位音樂的演奏表現力自動分析與演奏風格之資料探勘研究

研究成果報告(精簡版)

計畫類別：個別型

計畫編號：NSC 97-2221-E-216-032-

執行期間：97年08月01日至98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

計畫參與人員：此計畫無其他參與人員

處理方式：本計畫可公開查詢

中華民國 98 年 12 月 11 日

行政院國家科學委員會專題研究計畫成果報告

數位音樂的演奏表現力自動分析與演奏風格之資料探勘研究

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

計畫編號：NSC 97-2221-E-216 -032

執行期限：97 年 08 月 01 日 至 98 年 07 月 31 日

主持人：劉志俊 中華大學資訊工程學系

計畫參與人員：王鴻文、江俊賢、蔡咏昇、沈瑞祥、曾柏嘉、張俊堂、

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

本計畫中，我們針對 MP3 與 AAC 數位音樂，試著探討音樂給人的感知的高階情緒特性，據此提出一個音樂情緒模型。我們計畫先從壓縮領域中擷取出其 MDCT 特徵值係數，再進一步計算其 MPEG-7 與 MFCC 等音效低階特徵值作為分析其音樂內涵的來源。接著對重要的音樂特性，如起音點、基頻、泛音、和弦、調性、拍子、力度、節奏、速度、音色等，我們經由樣本分類技術與各音樂特性的樂理法則為基礎來自動進行計算。我們針對力度、節奏、音色等重要影響音樂演奏表現力因素提出自動分析的技術。

關鍵詞：human perception, music emotion model,

二、緣由與目的

隨著網路的快速發展以及多媒體壓縮技術的進步，目前有大量的多媒體資料在網際網路上快速的傳播，所以對於多媒體資料的分類與查詢，顯得日益重要。因此，有關多媒體資料內涵式分析的相關研究，越來越受到學術界的重視。

音樂是最主要的多媒體資料之一。早期的音樂內涵分析，主要是由音樂理論學者與心理學家，透過人工的方式來進行音樂作品的分析探討。近年來以訊號處理的方式來對音樂內涵進行分析，成為多媒體訊號處理方面的主要研究焦點之一。目前的主要研究方式是透過擷取音樂低階特徵值，對主要的音樂特性如旋律、節奏、調性、

速度等，在偵測、比對、分析、分類、索引等自動技術進行探討。近年來，隨著研究的快速進展，低階特徵已逐漸無法滿足對音樂內涵分析的高階需求，因此分析人類感知、情緒等高階特徵，乃至於對音樂家演奏風格詮釋的量化分析與探勘，已成為目前音樂內涵分析領域的研究挑戰之一。

三、研究成果

(一) 音樂感知情緒模型

Hevner 設計了一系列音樂引發情緒的實驗，請聽者寫下當演奏不同類型音樂時，情緒有何反應**錯誤！找不到參照來源**。透過此實驗來瞭解音樂的聲音和聽者的情緒反應間的關係，並提出八組情緒相關的形容詞組(adjective group)，每一組代表性的形容詞分別為高貴的(dignified)、傷心的(sad)、悅耳的(dreamy)、平靜的(serene)、優美的(graceful)、快樂的(happy)、使人興奮的(exciting)、強而有力的(vigorous)，此模型為針對音樂引發情緒，最早提出的模型。

我們以 Hevner 模型為基礎，並加入 Thayer 的二維模型以及 Juslin 的模型中區分四種情緒的概念，依照音樂的調性、音樂的速度、音樂的力度三種音樂表情因素的組合，歸納出三個音樂表情影響因素與 Hevner 的八個情緒的對應關係，如圖 1 所示。據此，我們提出之音樂感知情緒自動分類系統整體架構如圖 2 所示。首先我們從 MP3 音樂檔案中取出每一小節的音樂片段，接著透過兩種方法分析影響音樂表情

之因子：速度、力度、調性等特徵。取得音樂表情影響因子特徵後，我們透過音樂感知情緒分類器，判定此小節音樂之音樂片段在八種音樂情緒的各自比重。

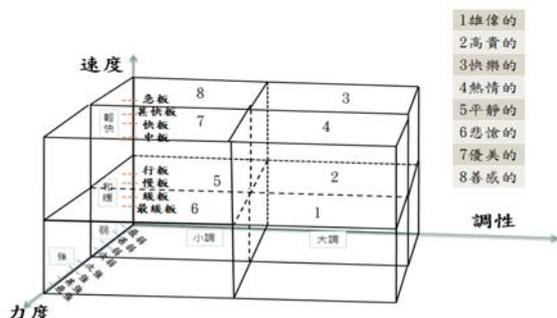


圖 1 音樂感知情緒模型

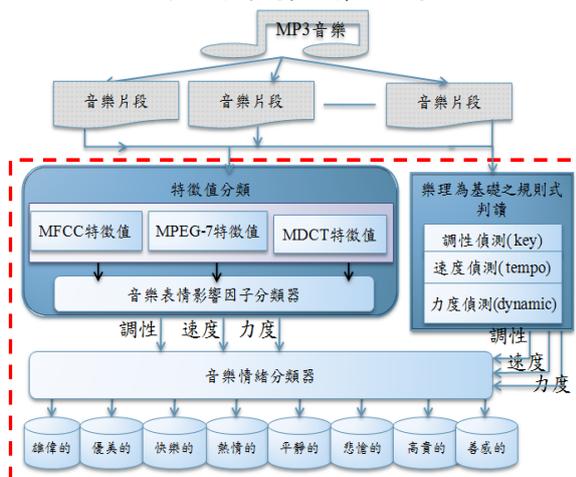


圖 2 音樂感知情緒自動分類系統架構

(二) 滑音之自動識別

中國民族樂器彈奏技法的特色是對於「韻」的重視，在常見的民族樂器中，如古箏、古琴、二胡、笛，尤其是對古箏與古琴類的樂器而言，特別追求「以韻補聲」的境界。但自古以來，有關韻的演奏分析與傳承都只能夠以師徒間親自彈奏教授的方式進行；在本研究中，我們希望可以藉由對古箏滑音技法進行自動偵測，能將古箏演奏樂曲轉換為樂譜，並標示出其中的滑音及其特性。滑音識別技術是對傳統樂器內涵分析的利器，此外，相較於傳統的師徒面授方式，此技術更有利於後人自我學習時所用。

我們提出之滑音的識別研究，初步以上滑音與下滑音為主。在一個完整的上滑音與下滑音彈奏過程中，可依時間劃分為起滑區段(A)、滑音變化區段(B)與終滑區段(C)等三個區段。如圖 3 所示，各區段涵義說明如下。

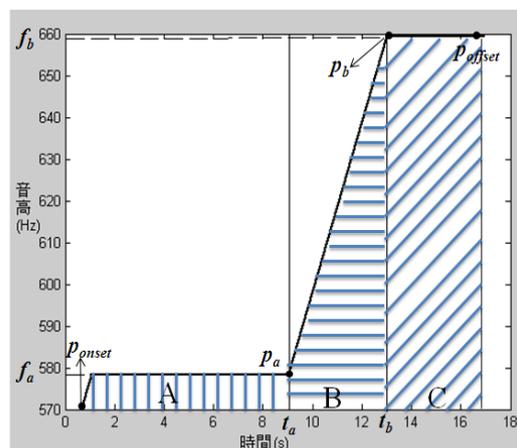


圖 3 滑音模型

- 起滑點區段(A)**：箏樂訊號中，一個樂音的開始彈奏之處稱之為起音點 (p_{onset})；起音點之頻率值，即稱之起音點頻率 (f_a)。在此區段中，頻率大致上維持一定值。
- 滑音變化區段(B)**：為滑音技法偵測最重要的區段，包括起滑點 (p_a)、終滑點 (p_b)、頻率變化量 (Δf)、時值變化量 (Δt) 等滑音主要參數，其中頻率變化量與時值變化量是識別滑音的重要特徵。滑音技法中，開始樂音與終止樂音之音程，即頻率變化量 Δf ，大約相差大二度、小二度或是小三度。因此，起滑點與終滑點之音程，會產生約 100 至 300 音分的頻率變化量。整個滑音變化區段時值 Δt 約佔滑音總時值範圍之四分之一，而此區段大致位於 2/4 至 3/4 滑音總時值之間。
- 終滑點頻率區段(C)**：這個區段中，包括了終滑點 (p_b) 與終音點 (p_{offset})。當頻率滑動至要求音準之時稱為終音點；滑音在終滑點之後會有一段頻率穩態的區塊，在這個區塊中的頻率值，即稱為終滑點頻率值 (f_b)。最後是整個樂音的結束點，我們稱之終音點 (p_{offset})。我們提出之滑音識別方法
 - 步驟一：偵測起音點 (onset) 與終音點 (offset)，找出各樂音的起始點 p_{onset} 與終音點 p_{offset} 。
 - 步驟二：求前 f_a 、 f_b (前 1/8T 與後 1/8T) 頻率值。計算滑音總時值的前 1/8 穩定頻率區段之平均頻率值作為起滑點

頻率 f_a ；滑音總時值的後 $1/8$ 穩定頻率區段之平均頻率值作為終滑點頻率 f_b 。

- 步驟三： f_a 、 f_b 頻率穩定判斷。由於 f_a 、 f_b 對於滑音點偵判別影響相當大，所以必須要判斷 f_a 、 f_b 是否為穩定值。如頻率變化太大，即判斷為非滑音，並將其去除。
- 步驟四：追蹤起滑點頻率 f_a 、終滑點頻率 f_b ，以作為計算 Δf 所用。接著偵測起滑點 p_a 與終滑點 p_b 發生之時間點 t_a 與 t_b ，以作為計算 Δt 所用。
- 步驟五：求滑音變化區段之時間間隔。取出 t_b 與 t_a 進行相減，可算出滑音變化區段中的時間長度 Δt 。
- 步驟六：求滑音變化區段之音程。
- 步驟七：檢驗滑音之音程與時間間隔是否合於滑音限制。

(三) 韻律結構分析

以往在音樂的內涵式分析著重於旋律比對以及拍子追蹤為主，對於音樂的韻律結構(metrical structures)分析方面較少定量研究。拍子是分析音樂的重要內涵，人們聽到一段音樂，往往會不由自主的配合音樂節奏打拍子，甚至身體也跟著韻律律動。一首音樂的樂譜上也必定會標記其拍子資訊，樂曲的節奏感覺是音樂詮釋的重要面向。我們提出一套自動拍子感測系統，目的在從音樂中取出不同結構之音長特徵，作為音樂分析的依據。樂曲中取出起音點資訊，到泰坦及拍的感測單元，統計起音點長度，利用赤池資訊量準則逼近並取出高斯混合模型之資訊，進行泰坦檢測及拍點檢測，泰坦檢測本文設計了三個實驗，依據不同出發點找尋最小時間刻度泰坦(利用最短起音點長度找泰坦、最長起音點長度找泰坦及最頻繁起音點長度找泰坦)；拍點的檢測本文分離音樂高音資訊，找尋低音資訊中最頻繁的時間跨步(一個簡單的時間跨步通常發生在低頻，如，Bass)；然而要達到拍子自動判別還需要小節及拍子的資訊，小節及拍子在音樂中通常是伴隨再一起的，所以感知拍子需要透

過一些知覺因子，拍子並不是雜亂無章的，由強弱交替的循環造就了拍子的形成，所以加入了知覺因子”重音”對拍點找尋精準之小節判斷。

四、計畫成果自評

本計劃目前的研究成果至已發表研討會論文四篇[1][2][3][4]，期刊論文一篇[5]，另有兩篇期刊論文投稿中[6][7]。

五、參考文獻

- [1] 蔡咏昇、劉志俊，“MP3 音樂的力度自動偵測與表現分析，” 二〇〇八數位生活科技研討會，2008.
- [2] 沈瑞祥、劉志俊，“箏樂滑音之自動識別，” NCS 2009 全國計算機會議，2009.
- [3] 江俊賢、劉志俊，“數位音樂的自動韻律結構分析，” 2009 年資訊科技應用學術研討會，2009.
- [4] 劉志俊、鄭中皓、余俊賢、張誼賢、陳奕安、鄭聿哲、黃志銘，“基於 IEEE 1599 標準的 MP3 音樂自動斷句系統，” 2009 年資訊科技應用學術研討會，2009.
- [5] 張俊堂、劉志俊，“XML 資料整合與關聯法則分析及其在蛋白質資料庫之應用，” Journal of Information Technology and Applications, Vol. 2, No. 1, pp. 43-53, 2009.
- [6] 王鴻文、劉志俊，“MP3 音樂的聆賞情緒自動分類，”(期刊論文投稿審稿中).
- [7] 沈瑞祥、劉志俊，“箏樂滑音之自動識別，”(期刊論文投稿審稿中).