

# 原住民音樂之樂句分析\*

## Sentence Analysis of Aboriginal Music

陳良弼<sup>+</sup>      吳宜鴻      陳宏鎮      李珍焯<sup>++</sup>      簡敏紘  
Arbee L.P. Chen    Yi-Hung Wu    Hung-Chen Chen    Jennifer Lee    Ming-Hong Jian  
國立政治大學資訊科學系<sup>+</sup>  
國立清華大學資訊工程系  
國立東華大學資訊工程系<sup>++</sup>

### 摘要

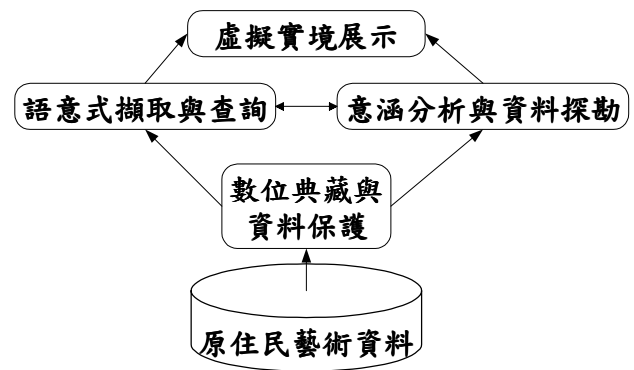
隨著電腦科技的進步，資料擺脫了以往文字的形式，以更多采多姿的方式展現；如何保存與分析多媒體形式的資料，並從中獲取知識，已成為當前重要的研究課題之一。原住民音樂極具特色，本研究除了蒐集並數位化許多珍貴的原住民音樂資料外，更以此資料為思考藍本，研發有助於分析音樂意涵的資料分析技術。

由於音樂格式的低階特徵有所不同，對於 MIDI 格式的音樂，我們基於音樂理論發展了一套有效的樂句分析方法；而對於音訊格式的音樂，我們則以人類聽覺為出發點，設計了一套方法來找出音樂內涵的轉折處（稱為樂句分界點）。我們所研發之樂句分析技術，除了可供探索原住民音樂的特色，協助專家分析音樂意涵之外，更可應用於自動化音樂分析及特徵擷取上，以便建立音樂內涵（content）及意涵（semantic）檢索系統。

**關鍵詞：**原住民音樂、樂句分析、數位典藏

### 1. 簡介

原住民文化的保存與傳承日趨重要，過去研究多著眼於文史資料的數位化，未深入研究不同族群間文化的關聯性，也缺乏對於原住民藝術意涵的分析。我們乃結合資訊工程與民族文化兩個不同領域的專家學者，一同致力於原住民藝術之數位典藏及其保護、分析、擷取、與展示系統之研究及實作，其四項研究主題如圖一所示。



圖一、整合架構

在「數位典藏與資料保護」主題下，我們除了實際執行原住民藝術資料的數位化工程外，也研究相關的保護機制如數位浮水印等技術，以保障其藝術創作之智慧財產權。以此為基礎，我們分別從兩個方向探討數位化資料的擷取與分析：「語意式擷取與查詢」研究視訊影像之低階特徵擷取與分析，藉以開發較高階之語意式視像內容檢索分類系

\*本論文為國科會補助之成果，計畫編號：  
NSC-91-2213-E-259-011

<sup>+</sup>Contact: alpchen@cs.nccu.edu.tw

統。「意涵分析與資料探勘」則從藝術意涵的觀點出發，分別研究不同族群在視覺圖式及音樂舞蹈上的異同；本論文研究扮演其資訊技術的提供者，開發多媒體資料的探勘技術，既加速了原住民藝術創作的分析過程，也可藉此探索更深一層的文化奧秘。最後，「虛擬實境展示」著眼於以虛擬實境技術立體呈現原住民藝術之美，建立互動式展示的介面。結合這些研究成果，將可完成一個跨族群藝術分析、鑑賞、及教學系統。

本論文將介紹「意涵分析與資料探勘」主題下，我們在音樂資料探勘技術上的研究成果。首先，我們將說明目前蒐集與數位化原住民資料的成果，接著介紹國內外相關研究的進展，然後提出我們針對兩種音樂格式所提出的樂句分析方法，並描述實驗結果與討論，最後則是本論文的結論與未來展望。

## 2. 原住民資料之數位典藏

以下列出我們目前所蒐集與數位化的原住民藝術資料，其中，部分資料已供本論文的觀察及實驗之用。

### 樂譜文件之數位化

為了便於進行自動化樂句分析的研究，我們先將蒐集到的樂譜文件掃描存成BMP圖片檔，再轉成MIDI音樂檔，目前已數位化資料筆數如表一所示：

表一、樂譜文件（單位：筆）

族名	圖片數	音樂數	族名	圖片數	音樂數
太魯閣	10	5	排灣	18	16
布農	6	6	鄒	4	4
阿美	160	110	魯凱	15	13
泰雅	16	13	卑南	13	10
小計	192	134	小計	50	43

### 曲目釋意文字之數位化

為了便於未來提供跨媒體查詢的系統，我們也陸續蒐集曲目釋意文字資料，每首歌曲以歌名存檔，其曲目釋意分為歌詞內容、歌詞大意、歌詞背景解說、以及單字解說等四部分，目前已數位化資料筆數如表二所示：

表二、曲目釋意文字（單位：筆）

族名	筆數	族名	筆數
魯凱	10	排灣	21
布農	4	鄒	5
阿美	96	魯凱	13
泰雅	17	卑南	14
小計	127	小計	53

### 原住民資料之蒐集與數位化

截至目前為止，我們已蒐集與數位化的原住民資料如表三所示：

表三、原住民資料總整理

資料類型	已蒐集資料	已數位化資料	資料類型	已蒐集資料	已數位化資料
音樂	242 筆	217 筆	舞譜	10 筆	10 筆
文字	180 筆	180 筆	圖式	90 筆	0 筆
影像	800 筆	150 筆	視訊	9 小時	9 小時

## 3. 相關研究

多媒體資料的特徵眾多，隨著媒體型態的不同，其重要性亦有所不同；因此，如何從眾多特徵中擷取適當的特徵，並賦予適當權重，一直是很重要的研究課題。相關研究如：[10]用模糊亂度（fuzzy entropy）來分析資料特徵值以選擇合適的特徵；[9]根據交互資訊（mutual information）擷取特徵；[6]則將所擷取之特徵點應用於影像檢索系統上。

我們研究重點在於因應不同資料格式，提供自動化的特徵擷取機制；其中，我們針對MIDI與音頻兩種音樂格式，提出相對應的

樂句分析方法。如同文章一般，音樂架構也是由長短不一的句子所組成的，這些句子統稱為樂句；樂句所形成的架構通常可以顯示其音樂特色，以原住民音樂為例，其樂句架構與「領唱—合唱」即有著密切的關係。

在 MIDI 格式方面，我們以音樂理論為根據來分析音樂架構[2]；而對於音頻格式，我們則藉著偵測人類聽覺的改變來劃分樂句[8]。相關研究如：[1]以休止符與長音符劃分音樂片段；[3]開始發展音樂架構的概念，然其所考慮之樂理都相當薄弱；[5]結合音頻特徵的各種量測方法，來分析由不同樂器所組成的音樂內涵；[4]則用音頻特徵計算兩兩音樂片段間的相似度，並提出一套量測某段音樂新奇程度（audio novelty）的方法。

當前原住民藝術研究者仍以文件蒐集、田野調查、與人工比對的方式，來分析藝術創作的內涵，因此，通常只能針對單一部落或少量資料加以研究。面對多個族群或大量資料時，我們研發的樂句分析技術，將可提供自動化且有效率的工具。

國內有關台灣原住民文化的研究，以台灣大學「台灣平埔族探源」和中央研究院「平埔文化資訊網」為代表；而研究原住民音樂的文化與發展則有[12][13][14][15][16]。

### 3. 樂句分析技術

#### 3.1 MIDI 格式

根據音樂理論，樂句（sentence）是音樂架構的基本單位，每一個樂句則是由一個以上的樂節所構成的，而樂節（phrase）即是在曲調和節奏上形成和諧流暢的音樂片段。

有鑑於過去研究鮮少利用音樂和諧與流暢的特性來擷取樂節，我們引入旋律形式（melodic shape）的概念來達成這個目標。[7]

統計了幾種旋律形式在 36,075 個樂節出現的比例，其結果如表四所示；其中，A 表示樂節第一個音符的音高，B 表示樂節最後一個音符的音高，而 C 表示樂節其他音符的平均音高。我們先利用這些統計資訊擷取出樂節，再將這些樂節分群，最後將常常重複出現的相似樂節組合成樂句。

表四、旋律形式的分類及統計資訊

形式種類	定義	圖示	比例
凸狀	$A < C \wedge B < C$		38.6%
下降	$A > C > B$		28.8%
上升	$B > C > A$		19.4%
凹狀	$A > C \wedge B > C$		9.7%
其他			3.5%

#### 樂節之擷取

首先，我們定義任何滿足下列兩條規則之一的音符都屬於終止音符：一、這個音符之前有一個休止符；二、相對於前後音符，這個音符的音長有顯著的差異。由這些終止音符，便可將整首音樂分成多個音樂分段，接著，對於每個音樂分段，我們再利用表四的分類及統計資訊來擷取其中的樂節。

首先，根據[7]的統計顯示，三分之四的樂節長度都在 6 到 12 之間，藉此可初步篩選可能存在多個樂節的音樂分段。在擷取樂節時，我們依序檢查每個音樂分段開頭長度 6 到 12 的部分，若其符合表四某類旋律形式的定義，該部分就會被標記成樂節並與整個音樂分段分離；然後，我們再對音樂片段的剩餘部分進行相同的檢查，以此類推。

#### 樂節之分群

一首音樂通常都包含許多樂節，且相似的樂節會分布在不同位置上，為了提升樂句擷取的準確性，我們依據樂節的起始音高與整體的音樂輪廓，計算兩兩樂節的相似度，

當相似度高於預設門檻值時，這兩個樂節就會被歸類到同一個分群中。

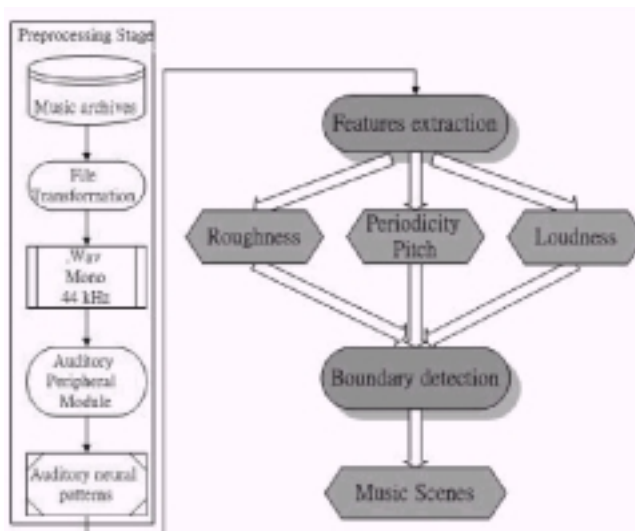
### 樂句之擷取

依據樂節分群的結果，我們就能將音樂表示成由分群標籤所組成的符號序列，同時我們也標記了可能是樂句開頭的樂節：包括整首音樂的第一個樂節與那些在休止符或長音符之後的樂節。根據這些樂節，我們即可從符號序列中找出常常重複出現的樂句。

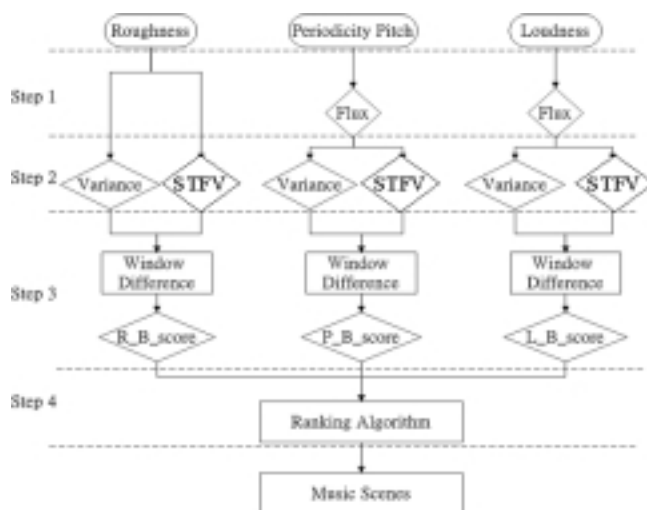
### 3.2 音頻格式

在音頻格式下，我們將樂句定義為音樂認知學上的和諧旋律 (harmonic melody)：即音長、音色、音高、音量都很相似的音樂片段。這四種特性的變化可以分別從粗糙度 (roughness)、週期音高 (periodicity pitch)、與響度 (loudness) 等三種聽覺特徵推導出，而且每個特徵的變化點無法由另外兩個特徵推出，我們稱為特徵互補性。

我們遂基於這個特性提出一套樂句分析方法，程序如圖二所示；首先對音樂資料的每個取樣點計算這三種聽覺特徵值，再套用圖三的分界點偵測演算法，選取符合人類聽覺的樂句分界點。



圖二、音頻格式之樂句分析程序



圖三、分界點偵測演算法

### 樂句分界點之偵測

首先，我們將一般常用的音訊處理技術如 Flux、Variance 與 STFV (Short-Term Feature Value) 先分別套用在每個取樣點的三種聽覺特徵上；其中，Variance 可顯示特徵值在局部區段內的變化程度，而 STFV 則可顯示特徵值於較大區段內的變化情況。

在第三個步驟中，我們利用兩兩滑動視窗 (sliding window) 在 Variance 與 STFV 上的差值，求出不同取樣點在單一特徵上的評分 (boundary score)，即  $w_1 * \text{Variance}$  差值 +  $w_2 * \text{STFV}$  差值，其中  $w_1$  與  $w_2$  為加權比重。在最後一個步驟中，我們先將單一特徵上擁有最高評分的取樣點列為分界點，並記錄其時間；接著，再從剩餘的取樣點中選取下一個擁有最高評分者，若其時間太接近某個已被列為分界點者則略過不計。重複這個選取程序，即可得到所有符合人類聽覺的樂句分界點。

## 5. 實驗結果與討論

### 5.1 MIDI 格式

我們從[11]中挑選 50 首原住民樂曲作為實驗資料，利用我們的樂句分析方法取出其

樂節與樂句；另一方面，我們也請專家標記這些樂曲的樂節與樂句，以便結果的驗證。表五顯示樂節擷取的正確性，其中準確度為吻合點數/吻合點數+誤判點數，而吻合比例為吻合點數/標示點數。

**表五、樂節擷取的正確性**

標示點數	吻合點數	誤判點數	準確度	吻合比例
596	465	215	68.38%	78.02%

表六顯示樂句分析的正確性，其中吻合比例為擷取樂句吻合數量/正確樂句數量。

**表六、樂句擷取的正確性**

正確樂句數量	擷取樂句吻合數量	吻合比例
126	79	62.70%

## 5.2 音頻格式

我們蒐集了 10 首不同類型的歌曲，分別套用我們的樂句分析方法後，再與受測者所認定的分界點做比較；表七是同時考慮三種特徵的結果，以下為效果評估的方法：

**表七、同時考慮三種特徵的結果**

風格	受測者分界點		我們方法的分界點				
			FB		BE		
	AG	%	FB	%	BE	M X	%
1. 流行	6/9	67	5/6	83	6/6	14	100
2. 流行	5/7	71	2/5	40	5/5	21	100
3. 流行	6/9	67	4/6	67	6/6	11	100
4. 電子	8/9	89	7/8	88	8/8	15	100
5. 電子	6/8	75	3/6	50	6/6	15	100
6. 搖滾	7/9	78	2/7	29	6/7	24	86
7. 搖滾	6/9	67	3/6	50	5/6	18	83
8. 搖滾	6/9	67	4/6	67	5/6	9	83

9. 民謠	6/10	60	5/6	83	5/6	6	83
10. 民謠	7/11	64	5/7	71	6/7	8	86
平均		71		63		14	92

- 對於一首音樂，AG 表示受測者一致認定的分界點與所有分界點的比值。如 6/9 表示受測者總共認定 9 個分界點，其中只有 6 個是一致的部分。
- 以 K 代表受測者一致認定的分界點數，FB 表示我們方法所找的前 K 個分界點，有多少比例是受測者一致認定的。如 5/6 表示我們所找到的前 6 個分界點，只有 5 個是受測者一致認定的。
- BE 表示當我們方法固定找出 24 個分界點時，受測者一致認定的 K 個分界點，有多少比例會被找出來；MX 表示我們方法找幾個分界點就能達到這個 BE 值。

## 6. 結論與未來展望

音樂分析技術研究是一項嶄新的挑戰，首先要能深入瞭解音樂的低階特徵與其隱含的意涵，方能發展出有用的自動化分析技術。我們的研發成果，除了在資訊科技層面突破音樂資料分析的技術瓶頸外，未來更將結合原住民藝術之數位典藏及目前積極發展中的多媒體查詢系統，讓原住民珍貴的藝術資產，獲得更深入而豐富的分析與認知、及更廣泛的應用。

## 參考文獻

- [1] Cambouropoulos, E., The Local Boundary Detection Model (LBDM) and its Application in the Study of Expressive Timing, *Proc. of Int'l Computer Music Conf. (ICMC)*, 2001.
- [2] Chen, H.C., Lin, C.H., and Chen, A.L.P., Music Segmentation by Rhythmic Features

- and Melodic Shapes, *Proc. of IEEE Conf. on Multimedia and Expo*, 2004.
- [3] Dannenberg, R.B. and Hu, N., Pattern Discovery Techniques for Music Audio, *Proc. of Int'l. Conf. on Music Information Retrieval (ISMIR)*, 2002.
- [4] Foote, J., Automatic Audio Segmentation Using a Measure of Audio Novelty, *Proc. of IEEE Int'l. Conf. on Multimedia and Expo (ICME)*, I: 452-455, 2000.
- [5] Herrera-Boyer, P. et al., Towards Instrument Segmentation for Music Content Description: a Critical Review of Instrument Classification Techniques, *Proc. of Int'l. Conf. on Music Information Retrieval (ISMIR)*, 2000.
- [6] Hsu, C.T., Wu, Y.T., and Chen, A.L.P., Content-Based Image Retrieval by Feature Point Matching, *Proc. of Int'l. Conf. on Storage and Retrieval for Media Databases (SPIE)*, 2001.
- [7] Huron, D., The Melodic Arch in Western Folksongs, *Computing in Musicology*, vol. 10, 1995.
- [8] Jian, M.H., Lin, C.H., and Chen, A.L.P., Perceptual Analysis for Music Segmentation, *Proc. of Int'l. Conf. on Storage and Retrieval Methods and Applications for Multimedia (SPIE)*, 2004.
- [9] Kwak, N. and Choi, C.H., Input Feature Selection for Classification Problems, *IEEE Trans. on Neural Networks*, 13(1), January 2002.
- [10] Lee, H.M., Chen, C.M., Chen, J.M., and Jou, Y.L., An Efficient Fuzzy Classifier with Feature Selection Based on Fuzzy Entropy, *IEEE Trans. on Systems, Man and Cybernetics*, 31(3), June 2001.
- [11] Yu, C. F., *The new sound from ancestors: A hundred indigenous folk songs in Taiwan*, Taiwan Yuan-Yuan Indigenous Culture Troupe, 1998.
- [12] 江冠明, 回歸大地—原住民音樂發展與世界音樂趨勢, *山海文化雙月刊*, 民國 89 年 3 月.
- [13] 吳榮順, 傳統音樂的即興—以臺灣原住民音樂為例, *藝術評論*, 民國 89 年 10 月.
- [14] 哈尤·尤道, 臺灣原住民樂器介紹, *玉山神學院學報*, 民國 89 年 7 月.
- [15] 陳鄭港, 臺灣原住民族音樂文化及其發展, *山海文化雙月刊*, 89 年 3 月.
- [16] 謝衣霖, 臺灣原住民傳統歌謠的現代展現, *山海文化雙月刊*, 民國 89 年 8 月.