

以音樂內容之分析做為音樂資料分類之研究

羅有隆
朝陽科技大學
資訊管理系
yllo@cyut.edu.tw

林奕昌
朝陽科技大學
資訊管理系
s9714613@cyut.edu.tw

摘要

已被發展出的音樂分類技術可概分為兩個方向，其一為以學習機器為訓練基礎的分類方式，另一則為以音樂內容為基礎的分析分類方式，各有其優缺點。而其中，現有以音樂內容為基礎的分類方式，多以音樂之單特徵進行，在少數類型的音樂辨識可達 70% 左右的分類精確度，而多數類型的音樂，可正確被分類的精確度則低了許多。本研究報告在於分析與研究音樂資料內容，並利用音樂資料的多特徵性質，找出是否有每一類型音樂所特有的內容特徵，並以這些特徵設計出具更高精確度的音樂資料分類公式。在我們的實驗中，我們發現利用音樂內容之多重特徵為分類基礎，可將古典音樂分類出來達 85% 的精確度。

關鍵詞：多媒體資料庫、音樂資料庫、數位音樂、音樂分類、內容擷取

Abstract

The music classification techniques can be discriminated into two categories — based by music feature classification and training by learning machine classification. Both have their advantages and disadvantages. For music feature classifications, most of the approaches are based on single music feature, such as melody or chord, and the accuracy is about 70% in few genres of music. However, the accuracy for classification of most music genres is lower. In this paper, we study the music contents and use the multi-features of music to design equation for more accuracy music classification. Our performance study shown that more than 85% classic music can be classified correctly by using multi-feature of music content for classification.

Keywords: multimedia database, music database, digital music, music classification, content-based retrieval

1. 導論

通常多媒體的資料量大，媒體的特徵(features)多，相對於純文字與數字的傳統資料庫來說，多媒體資料庫是既龐大又難管理。而現今硬體和軟體技術發展快速，使得如何有效率的管理和存取大量的多媒體資料，已經是愈來愈迫切。因此，如何為多媒體資料建立有效的搜尋方式，在多媒體資料庫的研究中，便成為一個重要的課題。較早之前的多媒體資料庫大都採用字串的關鍵字來查詢，如影片名稱、影像檔名、歌曲的作者、歌詞、或演唱者... 等等。漸漸的，近年有了相當多的探討集中在以內容為主的多媒體資料擷取(content-based multimedia data retrieval)。不過在初期有關內容為主的擷取方式之議題又大多集中在文件(document)、圖像(image)、影片(video)上[8][9][12][13][14][15][20][29][30][31][32]。而在聲音與音樂的內容擷取之研究上，就相對的比較少。然而隨著行動式數位多媒體播放器(如：iPod、iPhone)的廣受歡迎，數位音樂的需求大增，對於數位音樂資料庫的應用也日趨熱門。音樂資料庫研究的相關議題有音樂的分類(classification)、特徵擷取(feature extraction)、多特徵索引(multi-feature indexing)、以及近似搜尋(approximate searching)等等，主要目的都是在幫助使用者能簡單而且快速的擷取到想要的樂曲[3][20][22][23][24][25]。

音樂可以陶冶性情，也可以安撫人心，使人心情獲得平靜。因此，有腦筋動得快的業者推出了所謂的音樂心情手機，例如：Sony Ericsson W508、W902 與 W910，就加入了心情點播器(SensMe)的功能，主要是將每一首歌依照歌曲的快節奏、慢節奏或是快樂、憂傷分為四個象限，聆聽音樂時，使用者就能以手機內心情點播器依照當時的心情，播放想要聆聽象限內的歌曲。但是，有關歌曲的分類，則還需要使用者自己建立音樂的分類清單，才能使用，使得人性化功能大打折扣。近年來有關音

樂資料自動分類的研究，也因此越來越受到重視[1][4][5][7][10][20][26][33][34][38]。

由於音樂資料的內容(content)提供了多樣的特徵(features)可做為分析與查詢使用，例如：主旋律(key melody)、節拍(rhythm)與和弦(chord)等等，皆可呈現出每一首音樂獨特的曲風與特性。因此，以內容為主的音樂擷取(content-based music retrieval)，多年來是音樂資料庫研究的重要方向之一，而且也被當做音樂分類的重要參考依據。相關研究例如：Lin 等學者的 music classification using significant repeating patterns [20]、Brecheisen 等學者的 hierarchical genre classification for large music collections [4]、以及 Cheng 等學者的 automatic chord recognition for music classification and retrieval [7]等，皆是探討以音樂內容來做音樂分類。不過，目前這些以內容為主的音樂分類方式，其分類之精確度，少部分音樂類型的分類精確度可達 70% 左右，但是多數音樂類型的分類是低於此精確度，由此也顯示出現有的音樂分類技術尚有不少的進步空間。

本研究報告主要是利用音樂內容特徵(music content feature)之分析，設計公式，用於音樂之自動分類，以提升目前以音樂內容為主的分類精確度。我們期望所提出的方案，不僅對於音樂資料的正確分類，可以進一步的提昇，同時也希望這個研究的成果，在其他多媒體資料庫相關領域也可能有其它應用的價值。

2. 文獻探討

近年來音樂的自動分類方式，主要可概分為兩個研究方向：一個是以音樂內容為基礎(music content-based)的分析分類方式，是利用旋律(melody)、節拍(rhythm)與合弦(chord)等等音樂特徵，作為分類的依據，如前所述 Lin 等 [20] 與 Brecheisen 等 [4]。另一個則是以 learning machine 為分類方式，利用貝式分析、線性分析和類神經網路被等來建立音樂分類，如：Loh 等學者的 Extreme Learning Machine (ELM) [26] 與 Mandel 等學者的 Multiple-Instance Learning [27]。此外，Cheng 等學者的 automatic chord recognition [7] 則是以音樂特徵之一的和弦為分類依據，但也透過訓練等的學習機制。以下則簡要的介紹各音樂分類的方式：

2.1 SRP-Based Classification

Lin 等學者於 [20] 中利用 significant repeating patterns 來做音樂資料的分類，且以兩

個於音樂內容擷取 (content-based music retrieval) 的特徵(features)—旋律(melody)與節拍(rhythm)來代表音樂，以做為分類依據。重覆片段(repeating patterns)是指音樂中連續的片段，如：melody sequence 或 rhythm sequence，它重覆出現於音樂資料中至少 2 次或以上，這常常是自動判斷主旋律的依據。在很多音樂學及音樂心理學的研究中也認同重覆片段在音樂結構中為一普遍性的特徵 [18][28]，因為重覆片段的長度較實際音樂短，因此若以重覆片段來表示實際音樂，則對以內容擷取式的音樂搜尋，將會使其效能大大提昇。作者將重覆片段做了一些的規範—maximum length、minimum length 與 minimum frequency 等，以產生 significant repeating patterns (SRP)，並提供方法來估算 SRP 於音樂分類的可用性(usefulness)，以及與音樂類別的相關性。而人類的感知(human perception)與音樂理論(musicology)也被併入考量，以為做 SRP 比對時的相似計量。在經過 Usefulness of SRP for Classification、Similarity Measure for SRP Matching、以及 Class Determination 三個階段，透過公式計算，以最高分來決定一段音樂(music piece)是屬於那一類的音樂。

2.2 Hierarchical Genre Classification

Brecheisen 等學者於 [4] 中提出了階層式的曲風分類法(hierarchical genre classification)，此方法可以處理音樂內容的多重特徵(multiple features)，以達到較高的音樂分類準確度。此分類方法的基礎是使用到三種聽覺上(acoustic realms)的特徵—音色(timbre)、節拍(rhythm)與音高(pitch)。如此，一首音樂就可以此三特徵的多維向量(multiple feature vectors)來表示，而一段音樂就可以由一組特徵向量的集合來代表。接著將 feature vectors 做 hierarchical instance reduction，之後 reduced descriptions 再透過兩層的分類程式(two layer classification process, 2LCP)來處理階層分類問題，這過程還利用了 Support Vector Machines (SVM) [36] 做為分類者(classifier)，經過訓練與歸類的程序。

2.3 Extreme Learning Machine, ELM

ELM 是用來解決傳統類神經網路(traditional neural networks)的問題，Loh 等學者於 [26] 提出了利用 ELM 來處理音樂資料的分類。ELM 技術處理出自於曲面的錯誤(error surfaces)的梯度下降(gradient descent)問題，而

其倒傳遞學習演算法(back-propagation learning algorithms)已知是緩慢的，特別是當學習率參數小(learning rate parameter small)，會增加收斂的時間。另一方面來說，較大的學習率，也可能形成無法收斂的網路。縱使全域的最低值(global minimum)比局部最低值(local minimum)都還低，Error Surface 演算法還是相當容易受到區域最小值(local minima)的影響，使得學習終止。ELM 也處理了過度學習(over training)的問題，增加了類神經網路的概括能力(generalization capabilities)。

2.4 Multiple-Instance Learning

Mandel 等學者於 [27] 中分析了 multiple-instance learning 的音樂資料分類法。作者未提出新的技術，而是分析與比較 mi-SVM algorithm[2] 與 Multiple-Instance Learning via Embedded Instance Selection (MILES)[6]，兩演算法應用於音樂分類之優劣。依據實驗的結果，其結論是 mi-SVM 優於 MILES，能較精確的做音樂資料的分類。

2.5 Chord Recognition

用來做音樂分類的 chord recognition system 是 Cheng 等學者於 [7] 所提出來的，它分為兩個階段。第一階段為 training phase，他們併用了 N-gram model 於 Hidden Markov Model(HMM)架構中 [16]，做為和弦編曲(chord transcriptions)的訓練，以習得和弦發展(chord progression)的通則(common rules)。第二階段為 testing phase，利用前置訓練獲得的模式(pre-trained models)，chord sequence 可從 input 中被解碼(chord decoding)，得以做音樂分類。同時作者也提出了兩個新的和弦特徵—longest common chord subsequence 與 chord histogram，它們對音樂的分析、管理與擷取(retrieval)，都相當的有幫助。

上述的音樂資料分類技術，缺點在於精確度多為 70% 左右或更低，或者需要先做費時的音樂資料訓練，以取得分類資訊。而本研究報告，即在於應用音樂資料的內容特徵，發展出更高精確度的音樂分類演算法。

3. 音樂分類公式設計

Lin 等學者於 [20] 中利用音樂資料的旋律(melody)與節拍(rhythm)等兩特徵，先找到 significant repeating patterns 來做音樂資料的分

類，部分類型之分類精確度最高可達 70% 左右，此驗證了利用音樂資料的特徵，也可以做為音樂資料分類的依據，然而精確度還有待提升。因此，我們也繼續對音樂內容特徵的多樣性做分析，期望能有更精確的音樂分類方式。

我們使用的方法是找尋各別類型的音樂(如：classic 或 pop)中可能有的音樂內容特徵性質，例如：那些音(notes)、那些節拍(rhythms)以及那幾種前後音差變化(pitch changes)很常出現在某類的音樂中。我們以取樣(sampling)來計算發生頻率最高的 n 個特徵值 x_1, x_2, \dots, x_n ，再以此 n 個特徵值當做多維空間中的中心座標 (x_1, x_2, \dots, x_n) 。如此，一首音樂可透過這相對的 n 個特徵值 y_1, y_2, \dots, y_n ，以計算此音樂距離中心座標遠或近，即可判斷其可能歸屬於那一類的音樂。而距離函式 $d(y_1, y_2, \dots, y_n)$ 可設計如下：

$$d(y_1, y_2, \dots, y_n) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad \dots (1)$$

例如：統計某類音樂，其所屬音樂的旋律中，發生最高頻率的前三個音為 Mi(30%)、Re(20%) 以及 So(20%)。我們可以此三個音的發生頻率做為中心座標 $(x_1, x_2, x_3) = (0.3, 0.2, 0.2)$ 。若有一首音樂之旋律音符(Mi, Re, So)的發生頻率為 $(0.25, 0.26, 0.24)$ ，則可以帶入公式(1)以求得距離如下：

$$\begin{aligned} & d(0.25, 0.26, 0.24) \\ &= \sqrt{(0.3 - 0.25)^2 + (0.2 - 0.26)^2 + (0.2 - 0.24)^2} \\ &= 0.08775 \end{aligned}$$

當此音樂的特徵嘗試過各類型音樂之距離運算後，與之距離最短的類型，我們將該音樂歸類於此類型。

此外，我們也可以考量同時以多個特徵來運算，加總於各特徵所計算的距離，以嘗試求得更精確的分類，假設有 k 個特徵，則多特徵的距離公式如(2)。其中 $md(k)$ 為 k 個特徵的距離加總， $y_{1j}, y_{2j}, \dots, y_{nj}$ 為第 j 特徵的 n_j 個特徵頻率。

$$md(k) = \sum_{j=1}^k d_j(y_{1j}, y_{2j}, \dots, y_{nj}) \quad \dots (2)$$

4. 精確度分析

於本節當中，我們嘗試先以公式(1)與公式(2)用來做古典音樂(classic)與流行音樂(pop)的分類實驗，並採用音樂旋律中的音符(note)、節拍(rhythm)以及前後音差變化(pitch change)來做為音樂內容的特徵，分析其所發生的頻率，以距離公式來判別可能歸屬的音樂類別。而會採用前後音差變化，是因為它比較不受使用者的升音或降音(key)所影響，較適合音樂擷取的以樣本查詢(Query By Example)。

於實驗的音樂資料庫中，我們先對古典音樂與流行音樂各建立 100 首真實音樂資料，使得資料庫中有完整 200 首真實音樂的音符、節拍與前後音差變化之資料。實驗的參數包括取樣 10%~40% 以統計最高頻率之 3~7 個特徵值(n) 以建立中心座標。實驗分析可分為三個部份：

- (1) 單特徵對分類的影響
- (2) 多特徵對分類的影響
- (3) 特徵值個數對分類的影響

4.1 單特徵對分類的影響

此實驗以音樂內容之單一特徵來做音樂分類，以探討其對分類之精確性的影響。首先我們分別對古典音樂與流行音樂之音符資料取樣 10% 做分析，分別統計出最高頻率的五個音符，以做為中心座標，如圖 1 為古典音樂裡出現頻率最高的前 22 音符與出現頻率。而表 1 為頻率最高的五個音與出現頻率，分別為高音 Do、高音 Re、中音 Si、高音 Mi 與中音 La。因此，中心座標為(0.0853, 0.0814, 0.0772, 0.0754, 0.0744)。由於古典音樂與流行音樂曲風上的差異，代表兩類音樂之中心座標之音符與頻率，也就會有所不同。因此，我們可以透過距離公式(1)，來計算一首音樂與中心座標的距離，以分析它可能較接近那一類型的音樂。

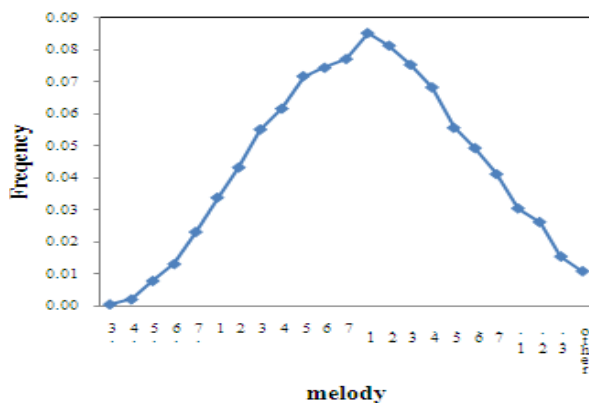


圖 1. 古典音樂音符分布圖

表 1. 古典音樂音符出現頻率前五名

排名	音符	出現頻率
1	高音 Do	0.0853
2	高音 Re	0.0814
3	中音 Si	0.0772
4	高音 Mi	0.0754
5	中音 La	0.0744

有了中心座標後，接著我們將資料庫中的 200 首音樂，逐一以音符的發生頻率來分類，統計可成功分類到正確音樂類型的精確度(Accuracy)是多少。統計方式是將每首音樂，逐一以同古典音樂最高頻之五個音符之發生頻率，以公式(1)計算與古典音樂中心座標的距離。同樣的，再將這首音樂以流行音樂之最高頻之五個音符，也計算出與流行音樂中心座標之距離。如此，可將此首音樂歸類於較短距離之音樂類別。之後，我們也繼續做取樣 20%、30% 與 40% 之實驗，統計分類之精確度。而每次實驗至少進行 5 次，以改變亂數取樣之參數重做實驗，並將實驗結果取平均值。以音符特徵為分類之實驗結果如圖 2 所示。圖中，對古典音樂之分類精確度，約在 80% 左右，明顯優於對流行音樂分類之精確度。而流行音樂精確度之起伏較大，且似乎有向下走的現象，表示流行音樂中之音符出現頻率特徵，變化較大，不若古典音樂來得穩定。

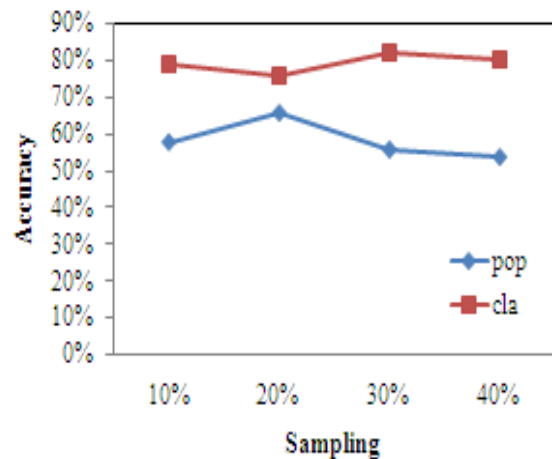


圖 2. 以音符分類之精確度

接著我們以相同的實驗方式，繼續以節拍與前後音差變化之音樂內容特徵來做音樂分類，實驗結果如圖 3 與圖 4。在圖 3 中，討論節拍的部份，其分類精確度隨樣本數之增加而提升，代表著若由較為充足之樣本數所取得中心座標，則可獲得較為精確之古典音樂分類，不過其精確度明顯不如以音符來做為音樂分

類。而流行音樂以節拍之分類精確度，與以音符分類之精確度，表現得相當的相似。

於圖 4，以前後音差變化來做音樂分類之精確度比較，古典音樂與流行音樂，則是差不多，兩曲線相當接近。但相對於以音符的分類方式，古典音樂表現略遜，流行音樂則稍優。整體而言，以單特徵來做音樂分類，於古典音樂效果較佳，且以音符之分類方式最佳。

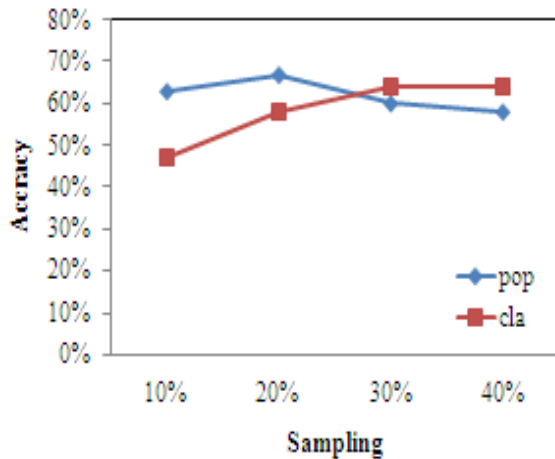


圖 3. 以節拍分類之精確度

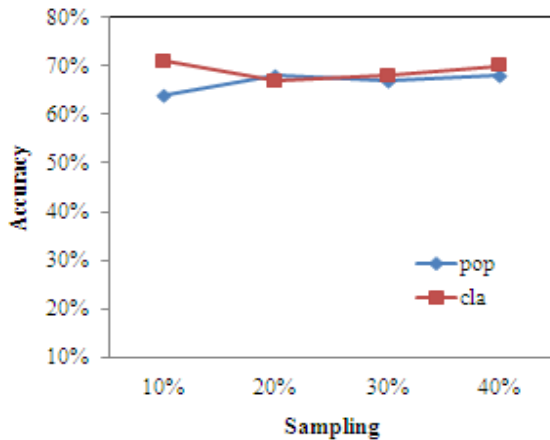


圖 4. 以前後音差變化分類之精確度

4.2 多特徵對分類的影響

在此實驗當中，我們探討了結合 2 個或 3 個音樂內容特徵(k)，是否可以提升音樂分類的精確度。實驗程序與參數設定，仍然與前一節(4.1 節)相同，但多用了公式(2)，以加總多特徵各別與中心座標的距離，以判斷可能的音樂類型。我們總共做了(音符+節拍)、(音符+前後音差變化)、(節拍+前後音差變化)以及(音符+節拍+前後音差變化)等四種特徵組合。實驗結果的精確度，則是分別呈現於圖 5~圖 8。

分析圖 5~圖 8，於圖 6 中以音符+前後音差變化來做古典音樂之分類，有較好之精確度，從單以音符特徵分類約 80% 的精確度，提升到約 85%。反觀流行音樂，採多特徵來做分類，精確度並未見提升，僅維持在 60%~70%。

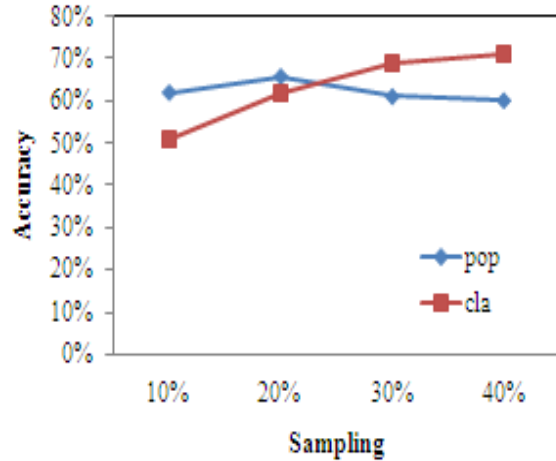


圖 5. 以音符+節拍分類之精確度

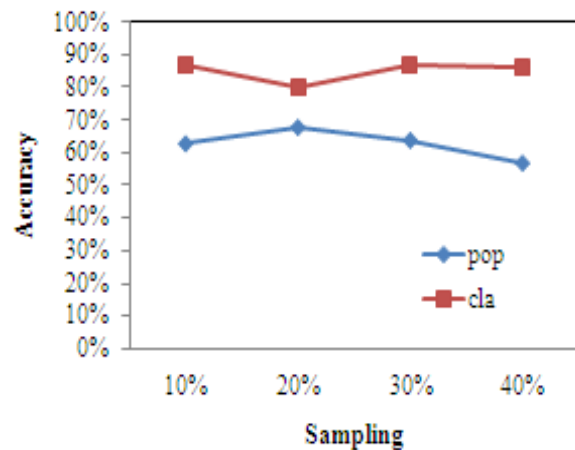


圖 6. 以音符+前後音差變化分類之精確度

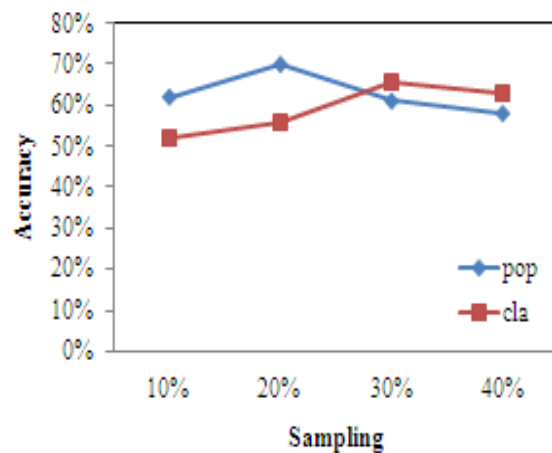


圖 7. 以節拍+前後音差變化分類之精確度

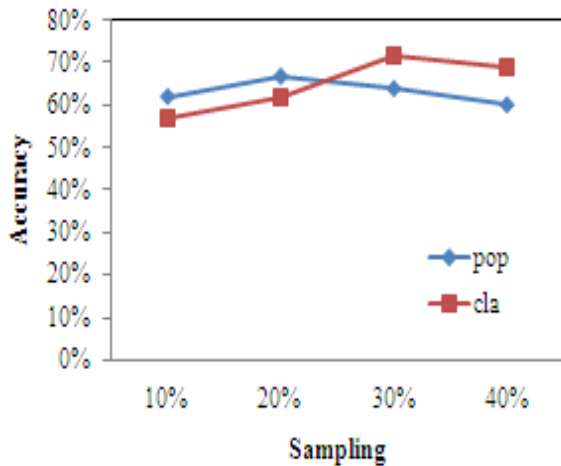


圖 8. 以音符+節拍+前後音差分類之精確度

4.3 特徵值個數對分類的影響

由於前兩節的實驗，特徵值的數目(k)皆是採用出現頻率最高的前 5 個，在本章節中，我們要討論特徵值的數目多寡是否會影響音樂分類的精確度。在此實驗中，我們以(音符+前後音差變化)兩特徵來做探討，古典音樂與流行音樂皆取 20% 的樣本來計算中心座標，而座標的維度(特徵值的數目)，採用 3~7 維，亦既分別以 3~7 個特徵值的數目來做音樂的分類，實驗結果如圖 9 所示。由圖裡可了解到，隨著特徵值數目的增加，可稍為增加音樂分類之精確度，但卻不是穩定成正向的增加。對古典音樂的分類來說，特徵值數目 4、5 與 7，有著差不多的精確度，但特徵值數目 3 與 6 之精確度則稍差。而對流行音樂的精確度來說，除了特徵值數目為 4 時的精確度偏低外，其餘特徵值數目的精確度則有稍微往上成長，逐步趨近於古典音樂。

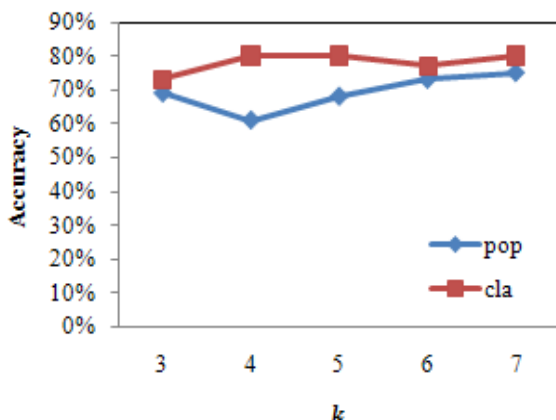


圖 9. 特徵值的數目(k)對精準度之影響

5. 結論

本研究報告提出了以音樂內容特徵(feature of music content)做為音樂分類之方式，以音樂特徵出現較高的頻率值，做為音樂分類的中心座標。對於一首音樂可能屬於那一類的音樂，則是計算其與中心座標的距離來決定。我們分別對古典音樂(classic)與流行音樂(pop)兩類音樂，搭配音符(notes)、節拍(rhythms)與前後音差變化(pitch changes)等三個音樂內容特徵，做音樂分類的初步實驗分析。而其中，以(音符+前後音差變化)兩特徵組合來對古典音樂做分類，準確度可達 85%，優於實驗中其它分類方式。但是對流行音樂的分類，則尚未發現有特別高的準確度。

未來我們將繼續推導新的分類公式或方法，除了持續對流行音樂的分類精確度做改善外，也將加入搖滾音樂(rock)與爵士音樂(jazz)進來做分類分析。

參考文獻

- [1] C. Anagnostopoulou and G. Westermann, "Classification in Music: A Computational Model for Paradigmatic Analysis," *Proceedings of the International Computer Music Conference*, 1997.
- [2] S. Andrews, I. Tsochantaridis, and T. Hofmann, "Support vector machines for multiple-instance learning," In S. Thrun and K. Obermayer, editors, *Advances in Neural Information Processing Systems 15*, pages 561-568, MIT Press, Cambridge, MA, 2003.
- [3] S. G. Blackburn and D. C. DeRoure, "A Tool for Content-based Navigation of Music," *Proceedings of the ACM Multimedia*, pp.361-368, 1998.
- [4] S. Brecheisen, H.-P. Kriegel, P. Kunath, and A. Pryakhin, "Hierarchical Genre Classification for Large Music Collections" *IEEE Seventh international conference on multimedia and Expo*, pp. 1385-1388, 2006.
- [5] W. Chai and B. Vercoe, "Folk Music Classification Using Hidden Markov Models," *Proceedings of International Conference on Artificial Intelligence*, 2001.
- [6] Y. Chen, J. Bi, and J. Z. Wang, "MILES: Multiple-instance learning via embedded instance selection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12):1931-1947, 2006.
- [7] H. T. Cheng, Y. H. Yang, Y. C. Lin, and H. H. Chen, "Automatic Chord Recognition

- for Music Classification and Retrieval,” *IEEE International Conference on Multimedia and Expo*, pp. 1505-1508, 2008.
- [8] G. Davenport, T. A. Smith, and N. Pincever, “Cinematic Primitives for Multimedia,” *IEEE Computer Graphics & Applications*, Vol. 11, No. 4, pp. 67-74, 1991.
- [9] Y. F. Day, S. Pagtas, M. Iino, A. Khokhar, and A. Ghafoor, “Objcet-Oriented Conceptual Modeling of Video Data,” *Proceedings of the IEEE Data Engineering*, pp. 401-408, 1995.
- [10] R. B. Dannenberg, B. Thom, and D. Watson, “A Machine Learning Approach to Musical Style Recognition,” *Proceedings of International Computer Music Conference*, 1997.
- [11] P. Dhanalakshmi, S. Palanivel, and V. Ramalingam, “Classification of Audio Signals Using SVM and RBFNN” *An International Journal Expert Systems with Application*, Vol. 36, No. 3, pp. 6069-6075, 2009.
- [12] E. A. El-Kwae and M. R. Kabuka, “Efficient Content-Based Indexing of Large Image Databases,” *ACM Transaction on Information Systems*, Vol. 18, No. 2, pp. 171-210, 2000.
- [13] S. T. Goh and K. L. Tan, “MOSAIC: A Fast Multi-Feature Image Retrieval System,” *Data & Knowledge Engineering*, Vol. 33, No. 3, pp.219-239, 2000.
- [14] R. Hjelsvold, and R. Midstram, “Modeling and Querying Video Data,” *Proceedings of the International Conference on VLDB*, pp. 686-694, 1994.
- [15] K. A. Hua, K. Vu, and J. H. Oh, “SamMatch: A Flexible and Efficient Sampling-Based Image Retrieval Technique for Large Image Databases,” *Proceedings of the ACM Multimedia*, pp. 225-234, 1999.
- [16] X. Huang, A. Acero, and H.-W. Hon, *Spoken Language Processing: A Guide to Theory, Algorithm and System Development*, Prentice Hall, 2001.
- [17] A. Holzapfel, and Y. Stylianou, “Rhythmic Similarity of Music Based on Dynamic Periodicity Warping,” *IEEE International Conference on Speech and Signal Processing*, pp. 2217-2220, 2008.
- [18] C. L. Krumhansl, *Cognitive Foundations of Musical Pitch*, Oxford University Press, New York, 1990.
- [19] F. F. Kuo and M. K. Shan “Looking for New, Not Known Music Only: Music Retrieval by Melody Style,” *International Conference on Digital Libraries*, Pages 243 – 251, 2004.
- [20] C-R. Lin, N-H. Liu, Y-H. Wu, and A. L. P. Chen, “Music Classification Using Significant Repeating Patterns,” *Lecture Notes in Computer Science*, Springer-Verlag, Vol. 2973, pp. 506-518, 2004.
- [21] C. C. Liu and Arbee L. P. Chen, “3D-List: A Data Structure for Efficient Video Query Processing,” *IEEE TKDE*, Vol. 14, No. 1, pp. 106-122, 2002.
- [22] C. C. Liu, J. L. Hsu, and Arbee L. P. Chen, “Efficient Theme and Non-Trivial Repeating Pattern Discovering in Music Databases,” *Proceedings of the IEEE Data Engineering*, pp.14-21,1999.
- [23] Y. L. Lo, C. H. Lee, and C. H. Wang, "Scalable Multi-feature Index Structure for Music Databases," *Information Sciences*, Elsevier, Vol.179, Issue 15, pp. 2662-2675, July 2009.
- [24] Y. L. Lo, W. L. Lee, and L. H. Chang, "True Suffix Tree Approach for Discovering Non-trivial Repeating Patterns in a Music Object," *Journal of Multimedia Tools and Applications*, Springer, Vol. 37, No. 2, pp. 169-187, April 2008.
- [25] Y. L. Lo and L. Y. Tsai , "Approximate Searching for Music Data in Real-Valued Feature Indexing," accepted by *Journal of Convergence Information Technology*, to appear.
- [26] Q. J. B. Loh, and S. Emmanuel, “ELM the Classification of Music Genres,” *ICARCV'06. 9th International conference on Control, Automation, Robotics and Vision*, pp. 1-6, 2006.
- [27] I. Mandel, and P. W. Ellis, “Multiple-Instance Learning for Music Information Retrieval,” *9th International Conference on Music Information Retrieval*, pp.577-582, 2008.
- [28] E. Narmour, *The Analysis and Cognition of Basic Melodic Structures*, The University of Chicago Press, Chicago, 1990.
- [29] S. Nepal and M. V. Ramakrishna, “Query Processing Issue in Image (Multimedia) Databases,” *Proceedings of the International Conference on Data Engineering*, pp.22-29, 1999.
- [30] J. H. Oh and K. A. Hua, “Efficient and Cost-Effective Techniques for Browsing and Indexing Large Video Databases,” *Proceedings of the ACM SIGMOD*,

- pp.415-426, 2000.
- [31] E. Oomoto and K. Tanaka, "OVID: Design and Implementation of a Video-Object Database System," *IEEE TRANSACTION ON Knowledge and Data Engineering*, Vol. 4, No. 4, pp. 629-643, 1993.
 - [32] S. W. Smoliar and H. J. Zhang, "Content-Based Video Indexing and Retrieval," *IEEE Multimedia*, Vol. 2, No. 1, pp. 63-75, 1994.
 - [33] M.-K. Shan, F.-F. Kuo, and M.-F. Chen, "Music Sytle Mining and Classification by Melody," *Proceedings. 2002 IEEE International Conference on Multimedia and Expo, 2002. ICME '02*, pp. 97-100, 2002.
 - [34] G. Tzanetakis, G. Essl, and P. Cook, "Automatic Musical Genre Classification of Audio Signals," *Proceedings of International Symposium on Music Information Retrieval*, Vol. 10, No. 5, pp. 293-302, 2001.
 - [35] G. Tzanetakis, A. Ermolinskyi, and P. Cook, "Pitch Histograms in Audio and Symbolic Music Information Retrieval," *Proceedings of International Symposium on Music Information Retrieval*, Vol.32, No.2, pp. 143-152, 2002.
 - [36] V. N. Vapnik, *The Nature of Statistic Learning Theory*, New York, Springer-Verlag, 1995.
 - [37] B. Whitman and P. Smaragdis, "Combining Musical and Cultural Features for Intelligent Style Detection," *Proceedings of International Symposium on Music Information Retrieval*, 2002.
 - [38] Yibin Zhang, and Jie Zhou, "A Study on Content-based Music Classification," *IEEE Seventh International Symposium on Signal Processing and Its Applications*, Vol. 2, pp.113-116, 2003.