

資工系實務專題研究計畫成果報告

音樂資料庫比對系統

專題編號：097CSIE-S002

執行期限：96 年 1 學期至 97 年 1 學期

指導教授：尤信程 副教授

專題計劃參與人員：陳芝亦 張詩婷

一、中文摘要

本專題的主題是建構一個音樂檢索系統。主要的概念是透過 MPEG-7 音樂特徵值(Audio Descriptor)來判別某歌曲片段為歌曲資料庫的哪一首歌。我們將資料量龐大的音樂原始 wav 檔轉換成 MPEG-7 之聲音簽章特徵值(Audio Signature Descriptor)，有效地降低資料的維度的，以達到加快音樂比對速度的目的，更有效率地進行歌曲的比對辨識的工作。

關鍵詞：

MPEG-7 之聲音簽章特徵值
(Audio Signature Descriptor)

STFT(Short-Time Fourier Transform, STFT)

漢明視窗(Hamming Window)

二、緣由與目的

目前網路的音樂平台，大多是讓使用者可以透過關鍵字(例如:歌曲名稱)的搜尋，進而找到想購買的音樂，少有直接拿音樂的內容來當作索引的依據。而辨識歌曲最簡單的方法，是拿查詢資料與資料庫

中所有的歌曲，每一取樣點做比較，找出最相近的歌曲清單，但是這種方法非常沒有不切實際，因為必須花費相當大的時間成本，而儲存空間也是一大挑戰。所以，我們想到去分析歌曲本身的特性，去產生唯一的特徵值，進而把特徵值當作比對資料，才是更有效率的方法。更希望可以藉由此應用，讓喜好音樂的人，在不知道歌曲相關資訊的情況下，還能搜尋到此首歌曲之正確資訊。

三、系統架構與技術說明

我們實作的程式可以分為『維度降低』、『搜尋演算法』兩個部份。在維度降低的部分，是希望能將歌曲特徵值利用線性轉換，降低特徵值資料量，我們藉由將 wav 檔的資料，搭配漢明窗(Hamming Window)係數執行短時間傅利葉轉換(Short-Time Fourier Transform, STFT)，再計算頻譜平坦度大小(Spectrum Flatness Measure) 其公式如圖一：

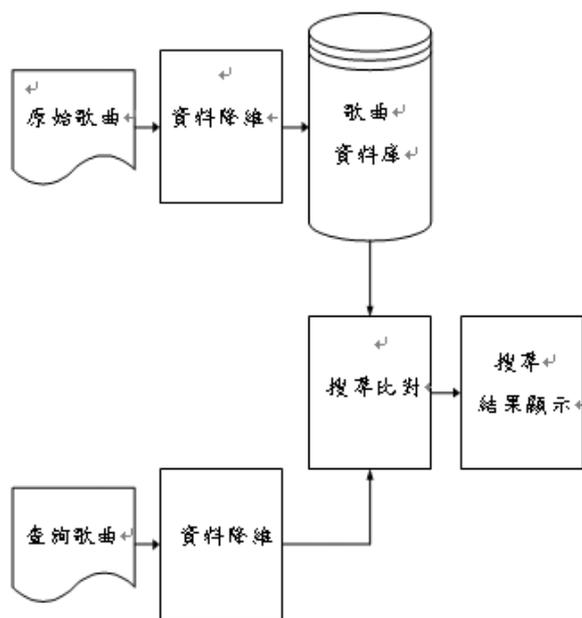
$$SFM_b = \frac{\sqrt{\prod_{i=il(b)}^{ih(b)} c(i)}}{\frac{1}{ih(b)-il(b)} \sum_{i=il(b)}^{ih(b)} c(i)}$$

(圖 1 計算頻譜平坦度大小公式)

最後，利用縮減比例(Scaling Ratio)來降低資料量，縮減比例的範圍是 2 到 128，我們是以 16 當作設定值，意思是說，將 16 點之資料取平均數及變異量，以此代替原始之 16 點資料，計算在同一頻帶但不同時間之平坦度的平均值(Mean)與變異量(Variance)。

在辨識策略的部分，則是透過降低維度所計算出的 Mean 及 Variance，進行資料庫搜尋比對，以找出最相近的歌曲並輸出其相關資訊。

圖二為系統架構說明：



(圖二-歌曲辨識系統架構圖)

四、成果與未來發展

對於比對為無失真歌曲之下，其正確率為百分之百。倘若音樂於錄音時產生音質不好的現象、音樂聲音較小或是音調高低起伏並不大時，那麼其鑑別率略有爭議性。

在歌曲資料庫辨識策略，我們假設所有查詢的歌曲都存在於資料庫中，因此著重在於如何以最低的維度，仍然可以保留最多的特徵，意即比較名次越前面越好。但於實際系統中，會有部分歌曲並未存在於資料庫中，此時需要有拒絕機制來判斷資料是否存於資料庫中，此部分值得後續進一步的討論。

五、參考資料

- [1] WAVE PCM soundfile format, <http://ccrma.stanford.edu/CCRMA/Courses/422/projects/WaveFormat/>
- [2] 李瑞育，利用 MPEG-7 支音樂特徵植作歌曲之辨識，國立台北科技大學碩論文，民國 93 年。
- [3] 陳威華，Music Retrieval System Using MPEG-7 Audio Descriptor，國立台北科技大學碩士論文，民國 96 年。
- [4] MPEG-7 overview(Version10), ISO/IE C JTC1/SC29/WG11 N6828, Palmade Mallorca, Oct. 2004.

