

國立交通大學

工學院聲音與音樂創意科技碩士學位學程

碩士論文

時變性的音樂情緒成份分析研究



**Analytic research on the time-varying ingredients of
emotion evoked by the sound of music**

研究生：吳偉廷

指導教授：鄭泗東 教授

中華民國一百年七月

時變性的音樂情緒成份分析研究

Analytic research on the time-varying ingredients of emotion evoked by the
sound of music

研究生：吳偉廷

Student : Wei-ting Wu

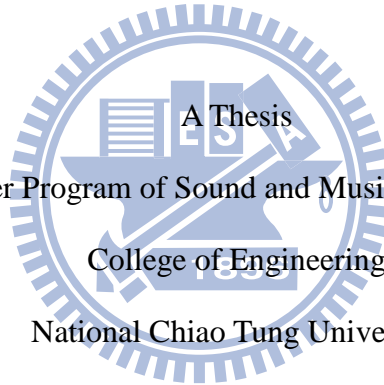
指導教授：鄭泗東

Advisor : Stone Cheng

國立交通大學

工學院聲音與音樂創意科技碩士學位學程

碩士論文



Submitted to Master Program of Sound and Music Innovative Technologies

College of Engineering

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Engineering

July 2011

Hsinchu, Taiwan, Republic of China

中華民國一〇一年七月

時變性的音樂情緒成份分析研究

學生：吳偉廷

指導教授：鄭泗東

國立交通大學工學院聲音與音樂創意科技碩士學位學程

摘要

人們的情緒隨著聆聽音樂的過程被牽動，以往音樂情緒檢索系統多半提出單一情緒分類。本研究以時變性的情緒變化為基礎並嘗試結合各種心理學家提出的情緒模型，包含二維情緒模型和類別式情緒模型，進而提出以 Content、Depression、Anxious、Exuberant 四種情緒為基礎成分並隱含二維情緒模型資訊的即時性音樂成分分析系統。系統中採用 Support Vector Machine 當作分類的演算法，針對各種特徵以 192 首訓練歌曲建立兩兩情緒間的分類模型以此考慮二維情緒模型中的兩個維度單獨分析的情況。特徵值上採用音樂特徵和音訊特徵兩大類，共計 11 個特徵。並依其特性以不同的音框長度作分析。最後再以音樂情緒的問卷調查來比對程式實驗結果和實際聆聽者感受是否符合。



Analytic research on the time-varying ingredients of emotion evoked by the sound of music

Student : Wei-ting Wu

Advisor : Stone Cheng

National Chiao Tung University Master Program of Sound and Music Innovative
Technologies

Abstract

While listening to music, people's emotions are affected. The common music emotion recognition systems used to provide only one emotion classification. This study presented a time-varying music emotion analysis, and tried to integrate several psychological emotion models, including two-dimensional emotion space and category type. And further to make a time-varying analytic system, based on four basic compositions: Content, Depression, Anxious and Exuberant, which also contains information from two-dimensional emotion space. This system uses Support Vector Machine as classified algorithm. This system uses Support Vector Machine as classified algorithm. Training in a variety of features by 192 music clips as training data to build emotional classification model between each two, in order to inspect the situations for analyzing the two dimensions separately. Feature extractions were divided into two categories: Music Features and Audio Features, total of 11 features. Each feature used different length of frame for analysis. In the end, this study performed a questionnaire survey to compare the program results with the actual listeners' experiences.

誌謝

研究所的學習方式與之前求學生涯大不相同，要靠自己去摸索和尋找才能得到問題的答案。在此首先要感謝指導教授 鄭泗東老師的帶領，在每個禮拜的進度會議上都能激發許多新的想法並給予我許多寶貴的建議讓我的研究能順利的進行。感謝就學期間修過的每堂課的老師，各式各樣的學門的知識讓我不被侷限在狹隘的思考中而能更廣泛地看待自己的研究。感謝俊傑、于恬、雲凱學長姐在研究上或其他各種相助，讓我的研究可以事半功倍。感謝我的父母、家人在就學過程中的支持。最後感謝立緯、昭成、芷伊、俊億、思賢、舒方、志翰、哲瑋、子衡、書寰、欣諭、佳渝、宜庭、育彬、正晨、亮志、信翰、本活、尚威、淳俊、瑋承(taco)、伊庭(小 ya)、亞雯、冠廷、兆凱、振洲、煦佳、冀昕、哲睿、冠吾、培鈞、岱容、明運、孟亨、所有幫助過我的人、所有一起玩過的人、我看過的所有表演。



目錄

封面	i
摘要	iii
Abstract	iv
誌謝	v
目錄	vi
表目錄	viii
圖目錄	ix
一、 緒論	1
1-1 研究動機	1
1-2 系統簡介	2
1-2-1 音樂情緒模型	2
1-2-2 音訊特徵值萃取	4
1-2-3 Mpeg-7 規格簡介	4
1-2-4 音樂內容分析	5
1-2-5 情緒分類演算法介紹	8
1-3 相關研究	9
二、 研究方法	11
2-1 研究內容探討	11
2-2 研究內容構想	12
2-3 系統架構及流程	13
2-3-1 音訊輸入	14
2-3-2 音框化分析	14
2-3-3 特徵值萃取	14
2-3-4 分類演算法	15
2-3-5 計分方式	15
2-4 音樂情緒訓練資料	15
2-5 特徵值計算	16
2-5-1 音樂特徵	17
2-5-2 音訊特徵	23
2-6 SVM 分類演算法	29
2-7 情緒計分	33
2-7-1 情緒分類模型建立	34
2-7-2 情緒判定	35
2-7-3 計分方式	37
2-8 音樂情緒成分表示法	38

三、	音樂情緒心理分析調查	40
3-1	問卷調查	40
3-2	問卷調查與實驗結果分析	40
四、	結論	44
參考文獻	45
附錄一.	訓練資料曲目	48
附錄二.	音樂情緒問卷	54
附錄三.	問卷調查受測者資料	55
附錄四.	測試音樂結果	56



表目錄

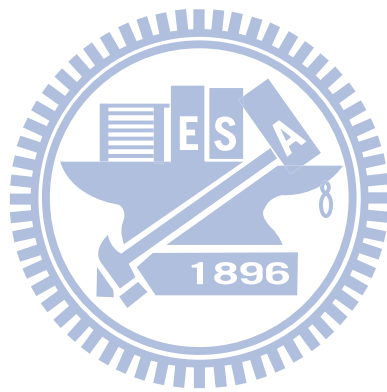
表 1. 一個八度間的音程表	6
表 2. C 大調的順階和絃	7
表 3. 各情緒成分的測試音樂數目	16
表 4. 採用的特徵值	16
表 5. 各特徵值採用的音框大小	17
表 6. 音樂特徵辨識率	34
表 7. 音訊特徵辨識率	35
表 8. 情緒判定舉例	36
表 9. 本研究測試音樂	40
表 10. Avenged Sevenfold - Dear God 結果分析	41
表 11. Celine Dion - My heart will go on 結果分析	42
表 12. Chopin - Nocturne opus 9 no 2 結果分析	43



圖目錄

圖 1. 各種 MIR 系統相對應的區域圖.....	1
圖 2. Hevner's adjective clock 進階版本	2
圖 3. 各式二維情緒模型	3
圖 4. Mpeg-7 中的音頻特徵描述	5
圖 5. (1)自然大調音階的組成(2)自然小調音階的組成.....	6
圖 6. SVM 分類示意圖	9
圖 7. 階級式情緒分類結構.....	10
圖 8. 二維情緒平面模型軌跡輸入程式圖	12
圖 9. 情緒成分元素示意圖	13
圖 10. 系統流程方塊圖	14
圖 11. MIRtoolbox 特徵萃取前流程.....	17
圖 12. Over the Rainbow 藤田惠美 - audio waveform.....	18
圖 13. Over the Rainbow 藤田惠美 - 取音框.....	18
圖 14. Over the Rainbow 藤田惠美 - 頻譜分析.....	19
圖 15. Over the Rainbow 藤田惠美 - 音樂色譜分析.....	19
圖 16. 主音萃取流程.....	20
圖 17. Over the Rainbow 藤田惠美 - 調性強度分析.....	20
圖 18. Over the Rainbow 藤田惠美 - 主音分析.....	21
圖 19. Over the Rainbow 藤田惠美 - 主音清晰度分析.....	21
圖 20. 調性萃取流程	22
圖 21. Over the Rainbow 藤田惠美 - 調性分析.....	22
圖 22. Over the Rainbow 藤田惠美 - 音調質心分析.....	23
圖 23. AudioSpectrumEnvelope 分類示意圖	24
圖 24. Over the Rainbow 藤田惠美 - ASE 分析	24
圖 25. Over the Rainbow 藤田惠美 - ASC 分析	25
圖 26. Over the Rainbow 藤田惠美 - ASS 分析.....	26
圖 27. Over the Rainbow 藤田惠美 - ASF 分析.....	26
圖 28. Centroid 示意圖	27
圖 29. Over the Rainbow 藤田惠美 -Centroid 分析.....	27
圖 30. Spread 示意圖	28
圖 31. Over the Rainbow 藤田惠美 -Spread 分析.....	28
圖 32. Over the Rainbow 藤田惠美 -Flatness 分析.....	29
圖 33. 六種情緒分類示意圖	34
圖 34. 情緒判定示意圖	37
圖 35. 各種衰減加權函數	37

圖 36. 實際影片結果畫面39
圖 37. 影片最後的情緒總結39



一、緒論

1-1 研究動機

音樂自古以來一直是日常生活中甚至是人文藝術上不可或缺的角色。而隨著時代的演進，科技的進步讓人們聆聽音樂的習慣也跟著改變，隨著音樂載體的不斷演進，黑膠、唱片到卡帶乃至於 CD，不但音質越來越好、容量越來越多也更方便取得和攜帶，但隨著 mp3 和網際網路的蓬勃發展使得音樂一口氣邁向數位化，甚至於踏入雲端資料庫的發展。快速的網路和小容量的 mp3 衝擊傳統唱片業，改變人們的消費習慣和聆聽習慣，線上的付費數位下載漸漸成為趨勢，類似的網站如：itunes、KKBOX、iNDIEVOX...等，購買後可以於任何地方隨時自網站資料庫中下載聆聽。於是當未來音樂都邁向數位化收藏時，除了現有的音樂標籤如：歌手、專輯、曲風、年份...等，音樂情緒分類也是熱門的研究主題。而在倫敦大學有研究[1]指出對於消費者而言，在音樂中情緒的感受是相對重要的訊息，如下圖：

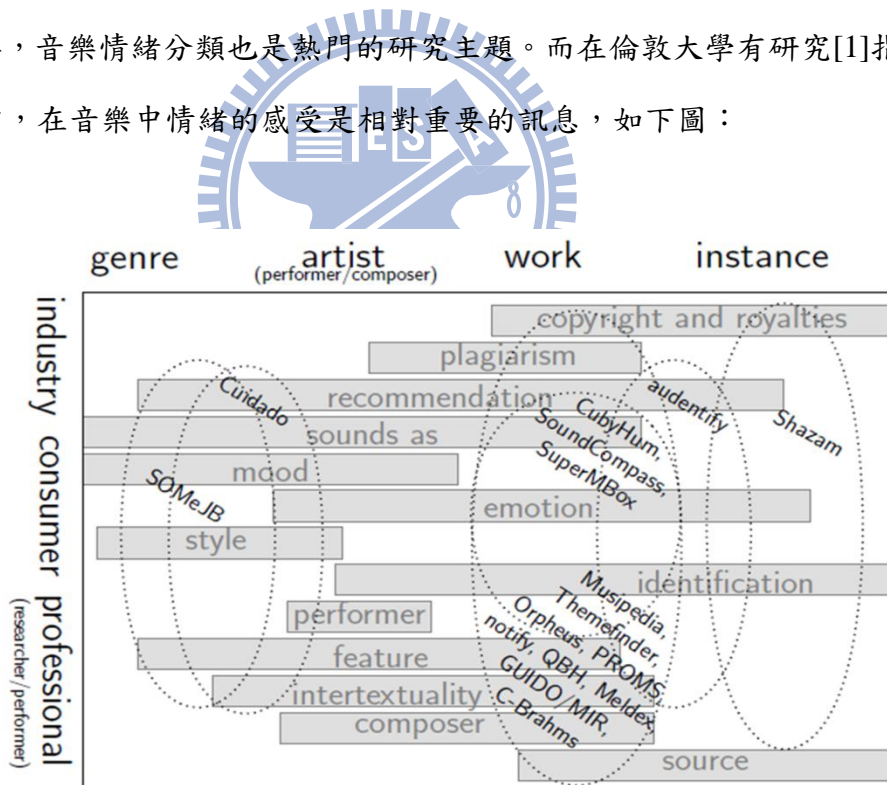


圖 1. 各種 MIR 系統相對應的區域圖

資料來源：[1]

近年來針對音樂訊號內容的檢索是熱門的研究，除了哼唱檢索(QBH)外，由於音樂內容的情緒分析的進步，更進一步的發展出情緒檢索(QBE)成了近期音樂的新型檢索方式。在各種平台上都可見到各種不同的音樂檢索系統，尤其在新式智慧型手機上也有依

據情緒的音樂檢索系統，但是都為將每一首歌視為單一情緒的分類，還尚未見到有針對時間變化、隨歌曲撥放引發情緒變化的相關系統。但是在歌曲的製作和譜寫上已有起承轉合的起伏安排，所以當在聆聽歌曲時應是有著情緒的起伏，這也是音樂之所以能感動人心的地方。因此本研究主要針對隨時間變化的情緒成分分析做探討。

1-2 系統簡介

在這一章節中會依序對音樂情緒分析(Music Emotion Recognition, MER)，相關研究內容做個介紹，首先 1-2-1 是 MER 中常見的核心理論—音樂情緒模型；1-2-2 則介紹關於 MER 常見的音訊特徵值；1-2-3 簡單介紹多媒體內容描述介面的標準—Mpeg-7；1-2-4 採用音樂分析的觀點探討樂理和情緒關聯；1-2-5 則介紹了相關研究中常見作為情緒分類的演算法。

1-2-1 音樂情緒模型

音樂與情緒間相關聯在心理學上已經有許多研究，詳細的發展文獻[2]中有詳細介紹，文中提到關於音樂情緒分類的研究主要可分為兩種：(1)類別式(2)連續式，其中類別式有重要影響力研究的來自 1936 年 Hevner 提出的 adjective clock，其中包含了 8 種情緒群聚，每個群組涵蓋 6 到 11 種情緒形容詞，這個模型也在 2003 年由 Schuert 結合二維情緒模型後提出進階版，如下圖：

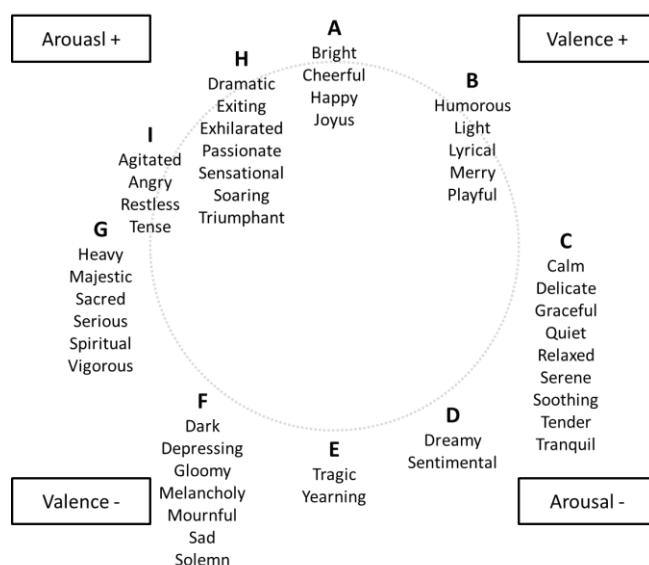
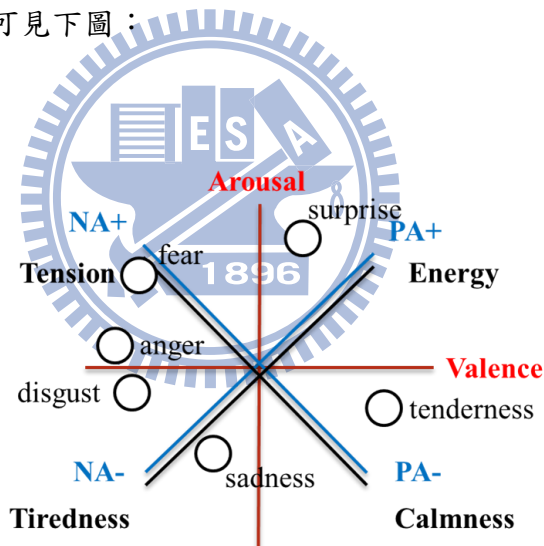


圖 2. Hevner's adjective clock 進階版本

資料來源：[2]

在 Hevner 之後陸續有許多學者發表相關研究，最近期的研究如 Zentner(2008)提出的階層式情緒模型 GEM-9，將四十種情緒依不同的權重彙總成九種情緒，最後再統合成三大類。

在連續性的研究上將情緒視為是連續的，因此現今大多時變性的音樂情緒研究大多採用這類模型。在這類型的早期研究中值得注意的是 Russell 提出的 Circumplex model(1980)，將情緒表示成 Arousal 和 Valence 兩個維度的二維模型；之後陸續有許多類似的二維情緒模型被提出，Watson 和 Tellegen 在 1985 年將 Circumplex model 旋轉 45 度同時考慮 Arousal 和 Valence，依此訂定兩個新維度 Postive Affective(PA) 和 Negative Affective(PA)；Thayer 在 1989 基於分離心理生物學系統(Separate psychobiological system)將 Arousal 分成 energetic arousal(EA)和 tense arousal(TA)兩個新維度；各情緒模型可見下圖：



Thayer's model (1986)
Watson Et Tellegen's model (1988)
Russell's circumplex model (1989)

圖 3. 各式二維情緒模型

資料來源：[2]

本研究嘗試結合類別式的成分表示和連續式的成分分析做為音樂情緒辨識的模型，詳細討論可見 2-1。

1-2-2 音訊特徵值萃取

由於現在人們聆聽的音樂大多以 mp3 為大宗，而不是 MIDI 這種內含音樂資訊的檔案格式，因此在分析音樂內容時大部分由兩個方面著手：(1)分析音訊波形的特徵，常見的特徵又可分為時域空間上的波形變化和頻譜分析後在頻域空間上的頻譜特徵。(2)逆向追蹤訊號產生的依據，在音樂訊號中即代表音樂本身的樂理分析。

音訊波形的特徵上常見的如：時域分析的過零率(Zero-crossing rate, ZCR)、波形(Envelope)、波形質心(Temporal Centroid)、發聲時間(Log attack time)、響度(Loudness)…等。頻域分析則有頻譜質心(Spectral Centroid)、頻譜散布度(Spectral Spread)、頻譜平坦度(Spectral flatness)、頻譜滑動(Spectral rolloff)、頻譜變遷(Spectral flux)、梅爾倒頻譜參數(Mel-frequency ceptral coefficient, MFCC)…等。

音樂內容中的特徵常見的如：音高(Pitch)、和聲(harmonicity)、和弦(Chord)、調性(Tonality)、主音(Key)、rhythm(節奏)、節拍(Tempo)…等。

上述特徵中並非皆適用於音樂情緒辨識上，本研究針對音樂內容的分析主要分為兩大方面：(1)樂理內容(2)音訊內容，依此分別去萃取各自類特徵為(1)音樂特徵(2)音訊特徵。將在 1-2-3、1-2-4 中介紹本研究採用的特徵並在 2-5 詳述特徵值計算過程。

1-2-3 Mpeg-7 規格簡介

MPEG(Moving Picture Experts Group)訂定出許多國際標準，如 Mpeg-1 和 Mpeg-2 提供了一套影音壓縮標準更在數位電視和影音光碟上實現互動式的影像，Mpeg-1 Layer3 的音訊壓縮技術也成為了現今廣泛流行的 MP3 檔案格式；Mpeg-3 原先為因應高畫質電視(HDTV)的應用設計，但隨後發覺 Mpeg-2 已足夠應付 HDTV 的應用之後就停止開發；Mpeg-4 提出了一套視訊的壓縮標準，其中第 10 部分的 H.264 更是應用在現今常見的網路影片上。為了因應日益複雜且大量的多媒體資料，提出一套多媒體內容描述介面的標準—Mpeg-7，跳過 Mpeg-5 或 6 直接發布 7 是由於

Mpeg-1+Mpeg-2+Mpeg4 的用意。Mpeg-7 的資料描述不在於資料的標籤而是資料的內容，而且支援眾多聲音音訊和視覺影像的描述，以影像為例，低階的描述包含形狀、顏色、紋理、運動、位置等。本研究主要採用 Mpeg-7 針對音訊內容方面的描述，在音訊內容方面 Mpeg-7 的低階描述主要可以分成六類共 18 種描述，如下圖：

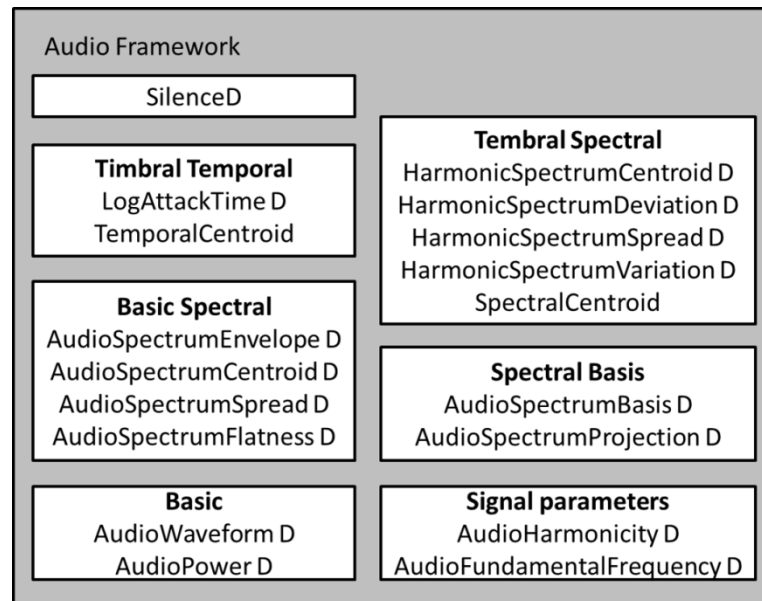


圖 4. Mpeg-7 中的音頻特徵描述

資料來源：Mpeg-7 Overview [3]

在本研究中只採用 Mpeg-7 作為基本的音訊內容分析，因此關於音色、旋律、和聲等地描述皆跳過再另外由音樂理論中設計相關特徵來描述情緒成分。因此在這邊只採用了 Basic Spectral 這個類別的描述子，相關詳細應用於 2-5-2 進行討論。

1-2-4 音樂內容分析

樂理可視為一般音樂的分析規則，其中包含節奏、調性、調式、和弦、和聲安排...等眾多方面的分析，一般常見的音樂大多是具有調性和固定節奏的音樂，其內容安排根據一般樂理也都是可以分析的。現今的流行音樂大多遵循西洋樂理，要由樂理來闡述音樂內容的話，最小的單位可以由每個音來開始討論，然後再由一連串的音組成不同的音階，這其中又包含著不同調性、調式等觀念。每個音依據不同的

音律下有不同的定義，古典音律如純音律是依據自然泛音列設計的音列，讓不同音組成的和聲最接近自然，但是不方便於轉調，同一樂器彈奏不同調性音樂時需重新調音。現今樂器大多採用十二平均律[4]的觀念設計，十二平均律中將一個雙倍頻率的八度音(Octave)分成 12 個半音(Semitone)以中央 A 定 440Hz，其他頻率和音高換算公式如下：

$$\text{semitone} = 69 + 12 \times \log_2 \frac{\text{frequency}}{440} \quad (1)$$

由於平均律為現今廣泛使用的音律，也決定各個音高的頻率分布，本研究的音樂特徵也是依此為基準。

決定了各個音高的定義後，接著分析由一連串的音列組成的音階，音階依照調性、調式也有許多不同定義，分別依據不同的音程(Interval)排列而成，其中音程是指兩個音之間的距離主要有：小、大、完全、增、減、倍增、倍減等型態，如下表：

表 1. 一個八度間的音程表

音程名稱	完全一度	小二度	大二度	小三度	大三度	完全四度	增四或減五度	完全五度	小六度	大六度	小七度	大七度	完全八度
半音差距	0	1	2	3	4	5	6	7	8	9	10	11	12

一般大家熟識的自然大調音階和自然小調音階以 C 當主音為例的音程分布如下圖所示：

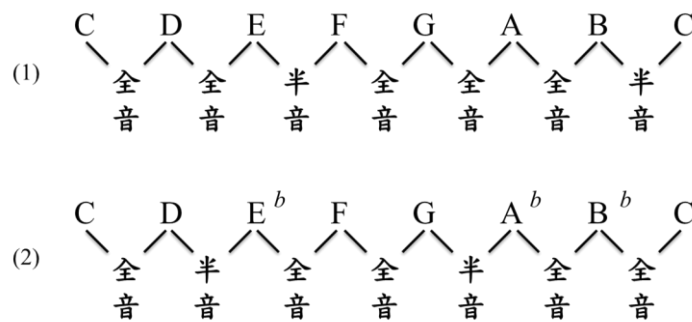


圖 5. (1)自然大調音階的組成(2)自然小調音階的組成

在此還有兩點需要被注意的，第一點就是關係大小調(Relative Major/Minor)，當圖

5.(2)的起始音換成 A 的話，此時音階的組成音和 C 大調將完全一樣；第二點就是平行大小調(Parallel Major/Minor)，當起始音相同，音階組成音不盡相同的兩組音階即為平行大小調，如圖 5. 中的 C 大調和 C 小調。當組成音相同時，若是由電腦來分析將無法判別此時是 C 大調或者是 A 小調，這時就需要仰賴主音(key)來分別，一個八度間的 12 個半音當主音的話，各大調音階將各自有其對應的 12 組自然小調音階作為平行大小調。除了自然大小調音階外各國家也有各自發展出來的民俗音階，現代音樂中也有更進一步分析的調式音階等。

音樂理論中除了音的橫向的流動組成的音階分析外，另一項重要的分析在於縱向的聲音疊加，也就是同發出兩個音以上的和聲分析。和絃的組成主要在於音與音之間音程的排列組合，但根據自然泛音列可以知道各音程間和諧度的不同，這也造就了各種和絃聽起來會有不同的感覺進而牽動聆聽者的情緒。三個音組成的三和絃(triads)是最常見的和絃，又可分為大和絃(major chord)、小和絃(minor chord)、減和絃(diminished chord)、增和絃(augmented chord)。大三和絃組成音是完全一度、大三度、完全五度，如若根音為 C 則組成音為 C、E、G；小三和絃組成音是完全一度、小三度、完全五度，如若根音為 A 則組成音為 A、C、E；減三和絃組成音是完全一度、小三度、減五度，如若根音為 B 則組成音為 B、D、F；增三和絃組成音是完全一度、大三度、增五度(音程同小六度)，如若根音為 E^b 則組成音為 E^b、G、B。由前述分析可看出大三和絃、小三和絃、減和絃都可以由自然音階的音階內音組成，若將一組自然音階依序將各個音當作根音可以發覺七個順階和絃，如下表：

表 2. C 大調的順階和絃

級數	I	II	III	IV	V	VI	VII
和絃	C	Dm	Em	F	G	Am	Bdim
組成音	C E G	D F A	E G B	F A C	G B D	A C E	B D F
根音	C	D	E	F	G	A	B

增三和絃的和絃內音無法完全配合自然音階的音階內音，是屬於和聲小調音階的三級和絃。本研究只針對基本的自然大小調做分析，所以在和絃方面也只考慮了自然大小調的順階和絃，這在流行音樂中也是最常被使用到的形式。

在林金慧提出的研究[5]中也解釋了一般廣為人知的音樂和情緒間的關聯，大音程、大和弦、大音階、大調四項觀念與正面情緒的關聯性和小音程、小音階、小調四項觀念與負面情緒的相關性。此外，現今的當代音樂也不乏無調性和無固定節奏的實驗性音樂，甚至於音高不在傳統音律中而是裡面採用電子合成的新音律，以打破傳統規律為主的內容安排，這些音樂並不在本研究的考慮範疇中。

1-2-5 情緒分類演算法介紹

在上述小節的討論過程中，在選擇了音樂情緒模型後確立了音樂情緒的分類，因此為了將音樂內容對應到音樂情緒上，音訊的特徵分析只是個媒介，而基於這些音訊特徵將音樂訊號分群分類的演算法則是連接兩者的工具。

在資料的分群和分類上有各種不同的演算法，常見的如： k 相近(k -nearest neighbor, k -NN)、高斯混和模型(Gaussian Mixture Model, GMM)、隱藏式馬可夫模型(Hidden Markov model, HMM)、倒傳遞類神經網路(Back-Propagation Neural Network, BPN)、支援向量機(Support Vector Machine, SVM)…等，本研究中為方便分析高維度的特徵值，採用了 SVM 作為分類的演算法。

SVM 演算法主要的概念是找出一個平面可以將兩個不同類別的集合分開，在多維的特徵空間下，這平面稱為超平面(hyperplane)。以二維空間為例，可見下圖：

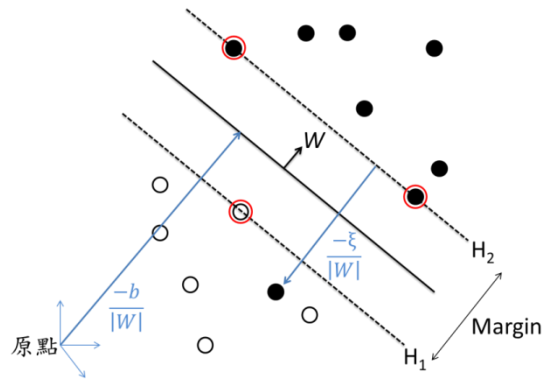


圖 6. SVM 分類示意圖

在已分類過的兩個資料群組但是不知道其分類的依據，SVM 嘗試在這兩個資料群組中找個超平面區隔開兩邊，兩邊的虛線代表的是支援用的超平面(Support hyperplane)，兩個支援超平面拉開的邊界(Margin)越大也就越能分開兩個群組，故最終目標是找出一個超平面有著最大的邊界稱之為最佳分割超平面(Optimal separating hyperplane，OSH)。

上述是可以理想分割的情況，但是資料會有誤差並不一定能夠完全的分割成兩種類別，所以在實際運用上會有更複雜的假設，在計算上會在導入個誤差值 ξ 並且給予懲罰的權重 C ，讓誤差的資料在判別上會多點成本(Cost)，詳細的推導見 2-6。

1-3 相關研究

現今關於音樂情緒分析(Music Emotion Recognition, MER)系統的研究有很多，各個情緒模型都有研究採用，也衍伸出各種不同型態的系統。在文獻[6]中採用了 Hevner 的情緒分組作為情緒標籤，採用了 17 種音樂音訊特徵，以 k -NN(k 選用 5) 為分類演算法，最後比較以各種不同降低維度的方法來降低特徵維度後的辨識率，如：FA(maximum likelihood common factor analysis)降低成 11 維、Ranker 降低成 6 維、GA(genetic algorithm of Goldberg)降低成 6 維、K-N-Match 降低成 13 維、PCA (Principal Component Analysis)降低成 10 維、Piv-Sel(local optimum selection of pivots)降低成 10 維。而在文獻[7]採用了 Tempo、Intervals、Loudness、Note、Density、Timbre、Chord 等七種特徵，直接對標記好的八種 Hevner 情緒類別訓練資料採用 GA-BP(back

propagation neural network and genetic algorithm)來建立一個音樂情緒辨識系統 (MER)。

採用二維情緒模型的如文獻[8]，利用了音色(Timbre)、強度(Intensity)、節奏 (Rhythm)三種類型的特徵值以一種階層式的情緒分類結構並採取 GMM 當作分類的演算法，如下圖：

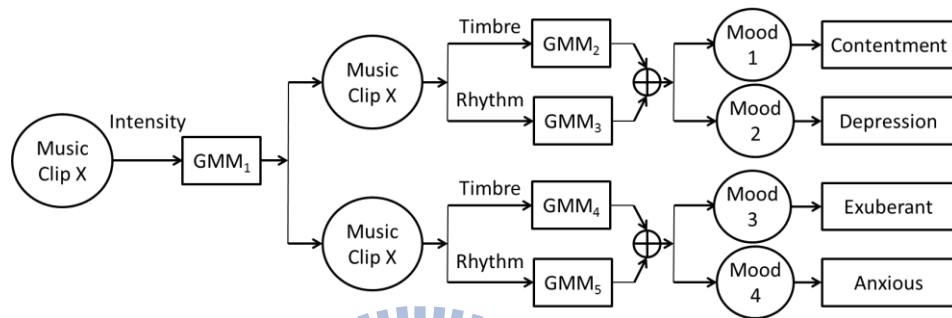


圖 7. 階級式情緒分類結構

資料來源：[8]

隨後也有文獻[9]以相似的架構但改採用 SVM 演算法作為分類器。

除了直接分析音樂音訊內容的研究外，也有以生物學上的特徵來分析音樂情緒對應到生理反應上的研究[10]，這也證實了音樂情緒的分類的二維模型其的可靠度。

二、 研究方法

2-1 研究內容探討

在音樂情緒辨識上有兩個主要的議題：

- (1)音樂情緒的時變性與否：關於音樂情緒的引發是隨著歌曲播放間經由音樂事件或音訊內容帶領變化；抑或情緒是由整首歌曲的變化所引發的單一結果，這是一大議題。
- (2)音樂情緒的量化：現在當前的音樂情緒研究多採用二維情緒模型（2-DES）為基底[8][9][10][11]，然後再嘗試將特徵萃取的結果對應到二維平面模型的兩個維度上去標記出當前情緒位置；但由於情緒是相當主觀的感受，要將其化作客觀的數值化表示則是二個大議題。

針對第一點議題，部分研究[12]為單一情緒的分析，但是歌曲在編寫間就有考慮到情緒起伏的鋪陳，故此聆聽者的感受勢必不會是整首歌都固定在一個情緒，而是隨著歌曲的音樂事件和音訊內容的激發而有所變動，而且近期研究[13]大多開始趨向連續變動情緒的方向發展。因此本論文主要研究主題將是在於隨時間變化的情緒成分分析。

針對第二點議題的處理方式現在有許多不同的做法，現今研究中有依自己的主觀感受對各特徵值設定權重計分將各音樂片段萃取的特徵值分別對應到二維情緒模型上兩個維度的值，以此作為量化[14]，也有採用測試人員直接畫下各自聆聽音樂的情緒感受軌跡圖，再將其量化結果和特徵值的數值計算一個數學模型做配對，測試程式畫面如下圖。

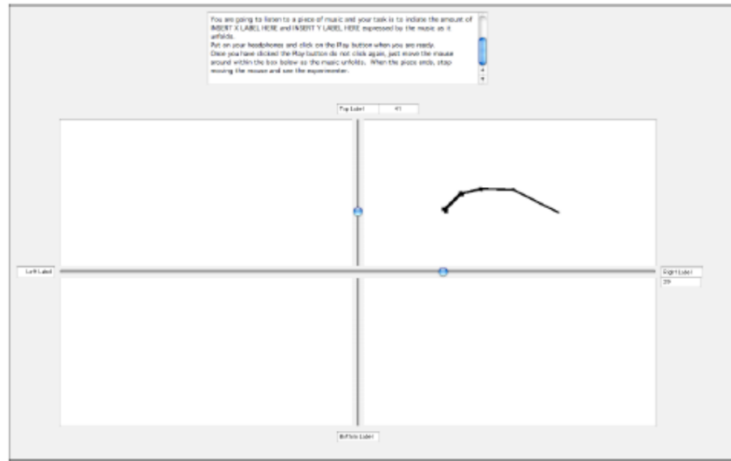


圖 8. 二維情緒平面模型軌跡輸入程式圖

由受測者手工繪入情緒軌跡模型。

資料來源：[15]

其中關於特徵值對應到二維模型的維度上的結果始終是令人質疑的，單一特徵是否只對應到二維情緒模型的單一維度和每個特徵間對情緒分類的影響是否各自獨立，都是值得研究的議題。因此，本研究將針對音樂情緒的成分來分析，而不是將音樂情緒量化表示在單一情緒上，在此假設情緒並不是單一獨立存在，而是由多種情緒混雜堆疊而形成。

2-2 研究內容構想

目前針對音樂情緒研究多以二維情緒模型為基礎[8][9][10][11]，少數一些研究提出的新式情緒模型也是延伸自二維情緒模型[16]，在 Zentner 和 Eerola 的書中[2]有針對二維情緒模型的演進做詳細介紹，心理學家不斷提出新的二維模型，以不同的維度來表示情緒的分布，但是綜觀各情緒模型，雖然有不同的維度表示卻有著相似的情緒分布，如圖 3。因此情緒模型的劃分依據各有出入，僅針對單一情緒模型為依據做量化有失客觀性。

近年 Zentner 提出新的情緒模型 GEM-9[17]中，表示出情緒是可以階層劃分的，將 40 個情緒形容詞加權歸類成 Sublimity、Vitality、Unease 三大類中。這也表示著個情緒形容詞間有一定的相關性，而早期二維情緒模型平面上的單點情緒應該也是可以歸類成簡單數個。

本研究結合上述研究提出一個對於情緒成分表示的假設，基於：

- (1) 二維情緒模型的四個象限皆有共通相近的情緒。
- (2) 各種情緒都可以簡化成簡單的情緒表示。

假設聆聽音樂所引發的情緒變化都是由四種基本情緒（Exuberant、Anxious、Depression、Content）成分組成，如圖 9 所示。如三原色的原理，這四種情緒間將在本研究視為組成情緒的基本要素，由這四個基本情緒依照不同比例組合進而表現出各式各樣的情緒形容詞，取代單一情緒形容詞的描述，藉以降低不同人之間對於情緒形容詞主觀的認知落差。

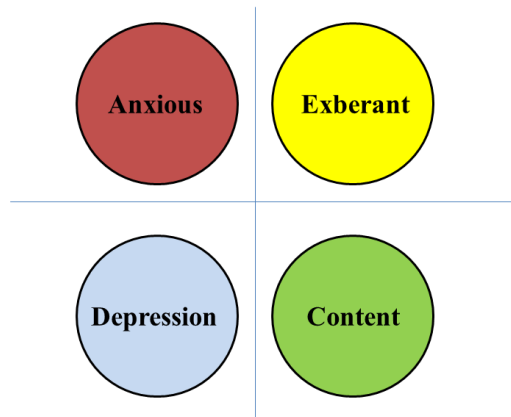


圖 9. 情緒成分元素示意圖

2-3 系統架構及流程

依據上一節中所提及的構想，本研究在 MATLAB 程式平台上設計了一套系統，系統的流程方塊圖見圖 10。在接下來的小節中會對於程式的流程一一簡單介紹，詳細的計算和說明則於之後 2-4 至 2-8 中探討。

ASF(AudioSpectrumflatness)、ASC(AudioSpectrumCentroid)，含 MIRtoolbox 中採用 Spread、Centroid、flatness 共計 11 個特徵值。各特徵詳細的內容於 2-5 討論。

2-3-4 分類演算法

本系統中的分類演算法採用 Chih-Chung Chang and Chih-Jen Lin[18]統整的 SVM(Support Vector Machine)演算法，這是一個最佳化分析的演算方式，可以在已分類好但是不知道分類依據的資料中建立一個分類模型，然後依此模型推測新進資料會屬於哪一個分類中，詳細流程將在 2-6 討論。

2-3-5 計分方式

當各特徵對於該音框的情緒判定後會在該特徵上累計，11 種特徵不一定全指向同一種情緒，可見得任一時刻的情緒並不是單一的。每個音框的情緒判定後也會累計到後面 3 秒的所有音框中，因為情緒的起伏並不是瞬間的反應而是各種音樂事件的累積，各時刻的影響會乘上一衰減加權累計進之後的音框計分，詳細計分步驟之後 2-7 會進一步討論。

2-4 音樂情緒訓練資料

訓練資料總共 192 首，內容皆為多聲部(polyphonic)的 wav 格式。詳細曲目見附錄一，音樂內容包含從古典到流行等各種音樂類型風格，情緒選取的方式採測試人員主觀判斷，情緒類別包含本研究提出的四種情緒成分：Content、Depression、Anxious、Exuberant，各情緒的歌曲數目見表 3。這邊採用主觀的選擇也是希望未來本系統在其他使用者的使用下可以依照各自的感受選取不同的訓練資料以建立針對個人有更佳分辨效果的分類模型，並藉此大大的降低不同人對情緒形容詞主觀的認知落差。

表 3. 各情緒成分的測試音樂數目

情緒	Content	Depression	Anxious	Exuberant
數目	47	48	48	49

2-5 特徵值計算

特徵值的選用主要依據音樂的音樂特徵和音訊特徵兩方面，採用 Olivier Lartillot、Petri Toiviainen 設計的 MIRtoolbox 1.3.2[19]，和 MPEG-7 [20]中關於音訊內容的描述子，詳細可見表 4。其中各特徵值計算時對音樂片段音框化的前處理也不盡相同，各特徵採用的音框大小可見表 5。

表 4. 採用的特徵值

萃取工具	特徵值	特徵類別
MIRtoolbox	Key	音樂特徵
	Key clarity	
	Mode	
	Tonal centroid	
	Spread	音訊特徵
	Flatness	
	Centroid	
Mpeg-7	Audio Spectrum Envelope	音訊特徵
	Audio Spectrum Centroid	
	Audio Spectrum Spread	
	Audio Spectrum Flatness	

表 5. 各特徵值採用的音框大小

特徵值	Key	Key clarity	mode	Tonal centroid	spread	flatness	centroid	ASE	ASC	ASS	ASF
Frame time	0.5 sec			0.075 sec	0.025 sec			0.05 sec			

2-5-1 音樂特徵

本研究採用 MIRtoolbox 作為音樂特徵萃取的工具，這個 toolbox 是以 MATLAB 為平台的音訊特徵萃取套件，由芬蘭 Jyväskylä 大學的 Olivier Lartillot、Petri Toiviainen 設計[21]，其中特徵值包含音樂動態、節奏、和聲、音色…等眾多方面，本研究採用版本為當時最新的 1.3.2，基於 1-2-4 的討論，在研究中將採用主音(key)、主音清晰度(key clarity)、調性(mode)、音調質心(tonal centroid)作為音樂特徵，在擷取這些音樂特徵之前，音訊資料進到這套系統裡都會先經過一套前處理用來分析音訊理的音樂事件，處理流程見下圖，

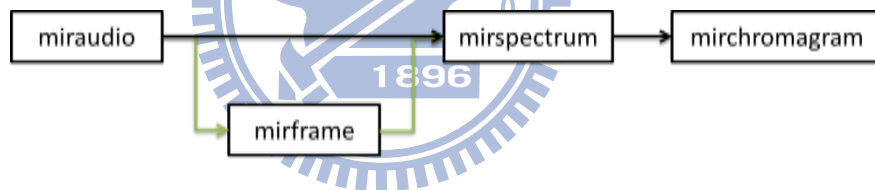


圖 11. MIRtoolbox 特徵萃取前流程

資料來源：MIRtoolbox user's manual[22]

由於本研究是針對 WAV 檔案格式而非 MIDI 格式有內涵音樂事件訊息，在已經經過複雜混音的音樂中，要回頭剖析該音樂片段的內容是非常困難的議題。

圖 11.的流程中，先以 miraudio 這程式讀入音檔，將 wav 的時域振幅波形讀入 MATLAB 中，如下圖：

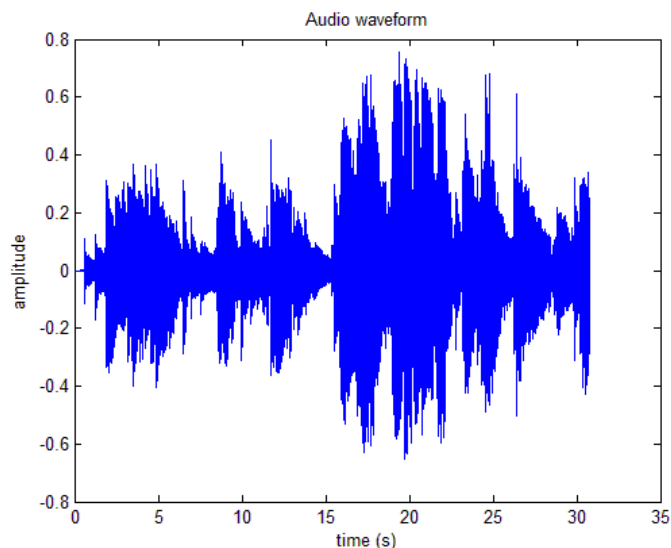


圖 12. Over the Rainbow 藤田惠美 – audio waveform

mirframe 是個可選用的變數，若是不經過這一環節則結果將會輸出整個音樂片段的分析結果，在這個環節針對不同的特徵質會給予不同的音框長度，以主音為例則是給予 500ms 的大小，取完音框後的結果如下圖：

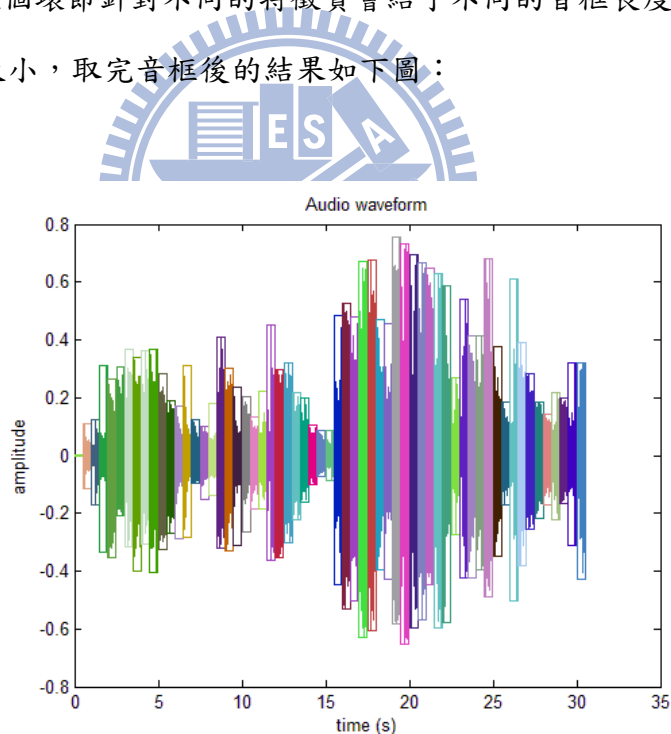


圖 13. Over the Rainbow 藤田惠美 – 取音框

音框化的處理完後接著是採用 mirspectrum 分析頻譜，由於在十二平均律中音高是根據頻率來定義，因此採用頻譜分析是反過來剖析樂理的重要方法，取完頻譜的結果如下圖：

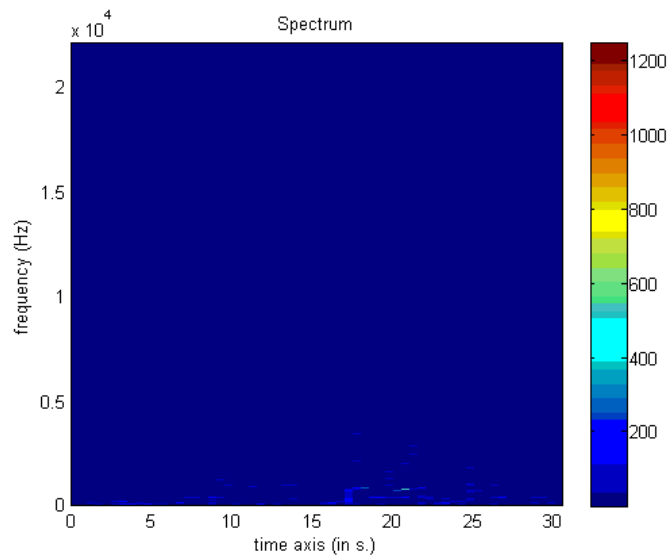


圖 14. Over the Rainbow 藤田惠美 - 頻譜分析

頻譜分析之後接著的 mirchromagram 是將頻譜的資訊對應到樂理上，是一個 12 維的特徵值，可以了解頻譜分析對應到 12 個半音上的成分多寡，如下圖：

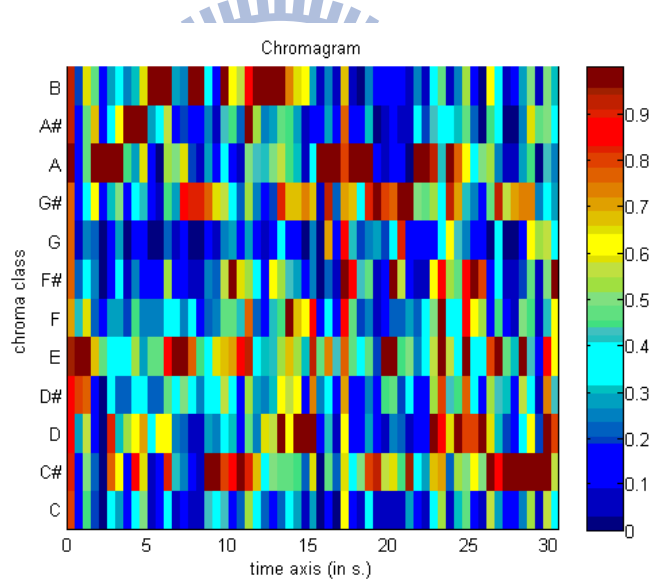


圖 15. Over the Rainbow 藤田惠美 - 音樂色譜分析

至此已將音訊波形對應到樂理的音高資訊上，接著再進一步的繼續下面的分析。

(1) 主音(Key)

如同 1-2-4 所討論的，要以音樂理論方面切入剖析音樂內容，該音樂主音是一個重要的特徵值。在本研究中，針對主音的分析會對輸入的音訊切割成

0.5 秒一個音框來處理。在 MIRtoolbox 中萃取 key 的流程圖如下：

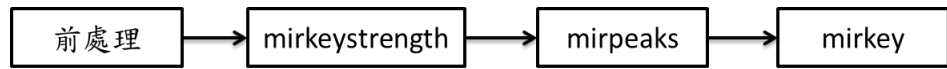


圖 16. 主音萃取流程

資料來源：MIRtoolbox user's manual[22]

在經過音訊內容對應到音樂色譜的前處理後，在 mirkeystrength 裡會針對 12 個主音構成的大小調做交叉相關函數(cross-correlation)來配對，最後輸出各個調性的相關度，如下圖：

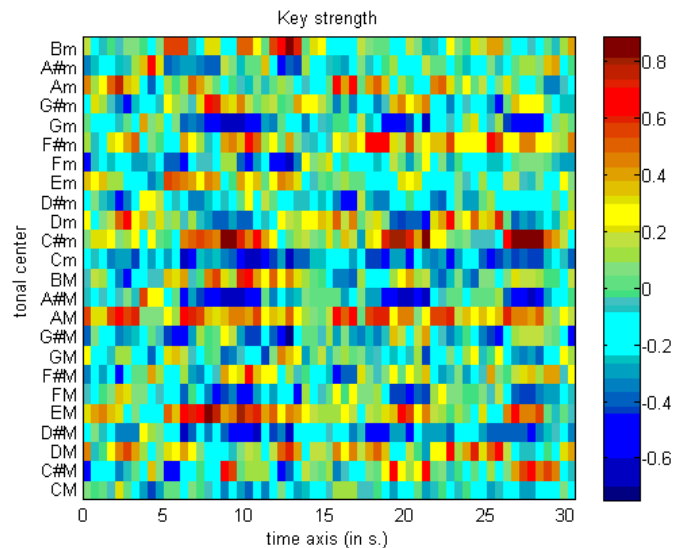


圖 17. Over the Rainbow 藤田惠美 - 調性強度分析

接著使用 mirpeaks 在調性強度的分布圖中找尋峰值即最有可能的調性分布，若是在平行大小調上都有著最高的調性強度，則將這組平行大小調的主音是為該音框的主音，最後主音的萃取結果如下圖：

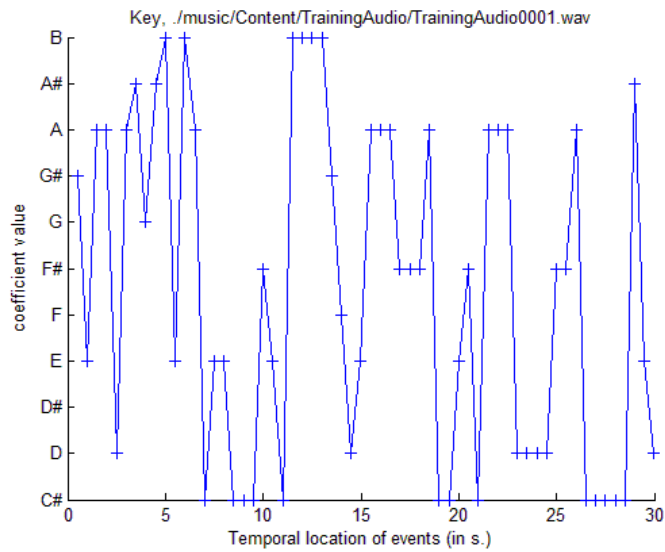


圖 18. Over the Rainbow 藤田惠美 - 主音分析

(2)主音清晰度(Key Clarity)

這項特徵是依據上一個特徵的內容同時計算的，在主音的特徵萃取中音樂色譜的分析後經由對於各種可能的大小調做交叉相關函數的計算，此時若是平行大小調皆有很強烈的峰值，則該音框有著清晰的主音也有著較高的主音清晰度的值，反之這項特徵有較低的值，結果如下圖：

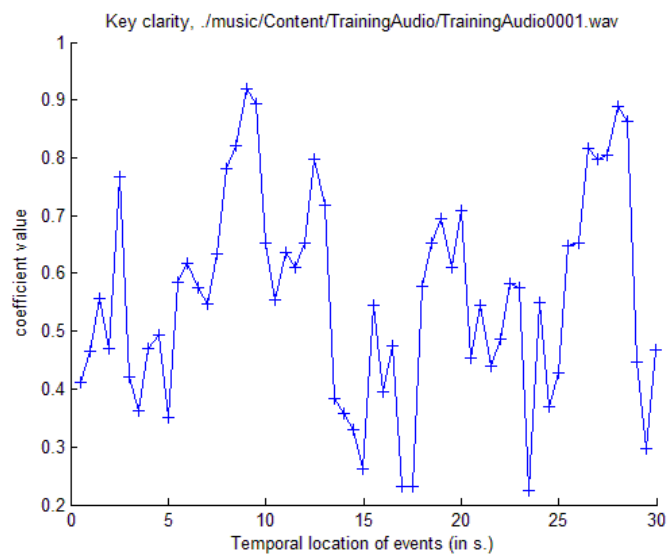


圖 19. Over the Rainbow 藤田惠美 - 主音清晰度分析

(3) 調性(mode)

相似於(2)中主音的萃取，對於調性的萃取流程如下圖：

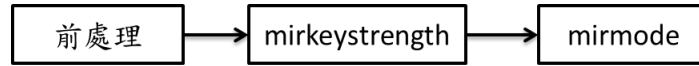


圖 20. 調性萃取流程

資料來源：MIRtoolbox user's manual[22]

調性和主音的差別在於分析完調性強度後，在 24 種不同調性結果中針對其中大調的峰值和小調的峰值做進一步的分析，最後的輸出是一個數值，以 0 為分界若為正值且越大的話代表是大調的強度越強，反之若為負值的話，則值越小其小調的強度越強，結果如下圖：

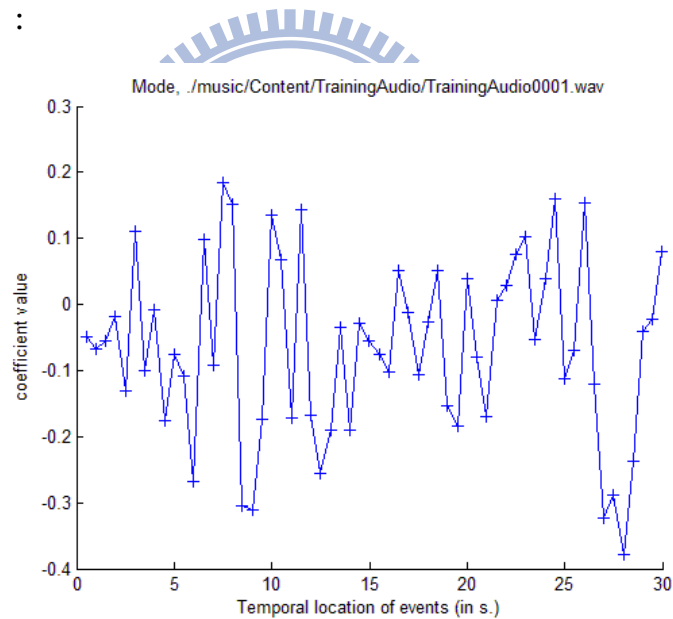


圖 21. Over the Rainbow 藤田惠美 - 調性分析

(4) 音調質心(tonal centroid)

音調質心在這邊是用來表達和弦的特徵，如 1-2-4 中討論的，大和弦可以帶給聆聽者正面的情緒，而小和弦則是會帶給聆聽者負面的情緒，這在對於音樂情緒的分析中也是個指標性的特徵。在 MIRtoolbox 中前處理後，除了十二平均律的分析更會加入音樂中和弦的組成推斷，包含五度圈、大三和弦、小三和弦的內容，最後

計算出一組 6 維的向量，其結果如下圖：

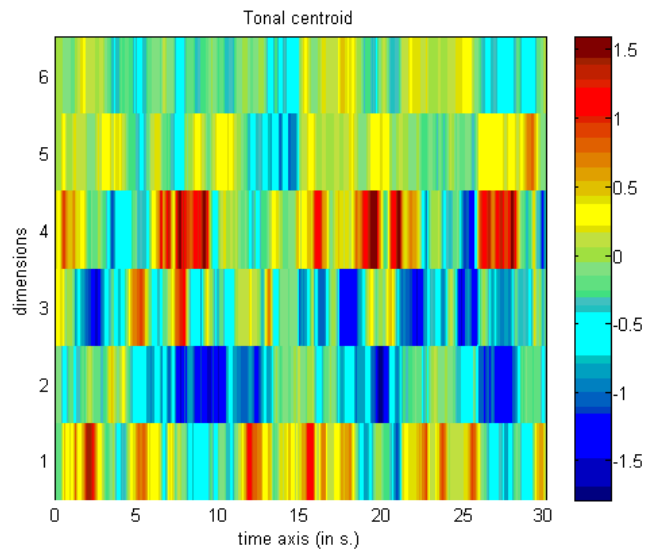


圖 22. Over the Rainbow 藤田惠美 - 音調質心分析

2-5-2 音訊特徵

在音訊特徵上採用了 Mpeg-7 中對於基本頻譜的特徵描述分別是 ASE(AudioSpectrumEnvelope)、ASC(AudioSpectrumCentroid)、ASS(AudioSpectrumSpread)、ASF(AudioSpectrumFlatness)，除了 Mpeg-7 還另外加入了 MIRtoolbox 中關於音訊內容的特徵 Spread、Centroid、Flatness，這些音訊特徵在情緒成分的辨識中都有著極高的辨識率，詳細的辨識內容可見 2-7，接著會針對各音訊特徵萃取作詳細介紹。

(1)AudioSpectrumEnvelope

這項特徵主要描述音訊經過頻率分析後在頻域中各頻率強度的波形，可以用來描述音訊輸入的頻譜圖，在 Mpeg-7 的描述子中會先將每個音框的分频段計算出 10 個參數，如下圖：

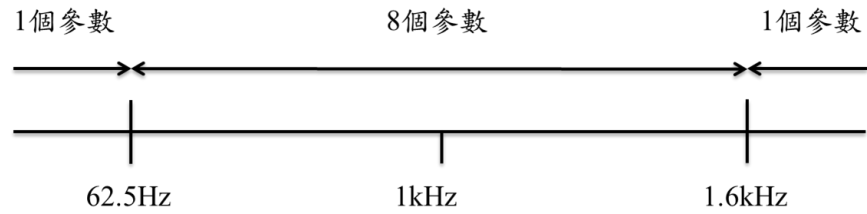


圖 23. AudioSpectrumEnvelope 分頻示意圖

各音框的計算公式如下：

$$ASE = \frac{|X(n)|^2}{lw \times NFFT} = P(n) \quad (2)$$

其中 $NFFT$ 表示 FFT(Fast Fourier Transform)計算的大小， lw 則是音框大小， X 是傅立葉轉換後的各頻率強度。運算結果產生 10 維的特徵和一組對應的權重，這 10 維的特徵可以根據對應的權重換算成頻譜，若將 10 維的結果乘上對應的權重則結果如下圖：

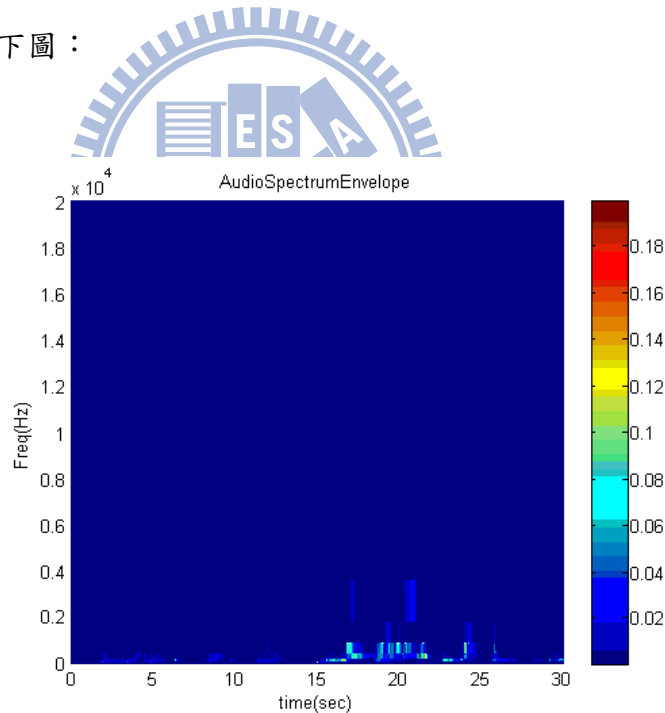


圖 24. Over the Rainbow 藤田惠美 - ASE 分析

其結果相近於 MIRtoolbox 中 mirspectrum 運算出來的圖 14，在本研究中採用 10 維的頻譜特徵值作為頻譜波形的特徵做運算。

(2)AudioSpectrumCentroid

ASC 是計算頻譜的質心作特徵，計算公式如下：

$$ASC = \frac{\sum_n \log_2 \left(\frac{f(n)}{1000} \right) P(n)}{\sum_n P(n)} \quad (3)$$

其中 $P(n)$ 是頻率 $f(n)$ 的頻譜強度。頻譜質心可以用來表達每個音框的頻率分布趨向，可視為一個簡單的頻譜能量分布表示，每個音框會計算出一個特徵值，其結果數值由公式中可見是以 1000Hz 為參考基準，若頻譜質心的落點大於 1000Hz 越多則為正值且越大，反之則小於 1000Hz 越多則為負值且越大，結果如下圖：

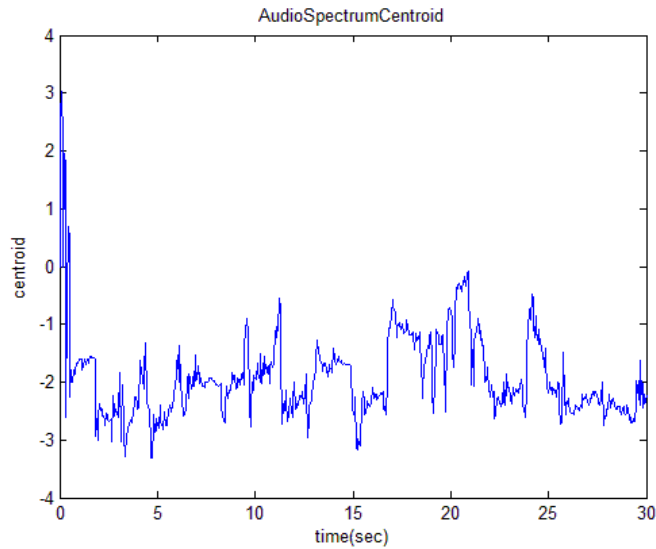


圖 25. Over the Rainbow 藤田惠美 - ASC 分析

(3) AudioSpectrumSpread

這項特徵延續 ASC 的分析，視為 ASC 的二階中心動差(second central moment)，就是所謂的變異數(Variance)，計算公式如下：

$$ASS = \sqrt{\frac{\sum_n \left(\left(\log_2 \left(\frac{f(n)}{1000} \right) - ASC \right)^2 P(n) \right)}{P(n)}} \quad (4)$$

其中各項變數同 ASC 中所計算的，這項特徵表現出頻譜的頻率分布是否集中，每個音框的計算結果為單一數值，計算的結果如下圖所示：

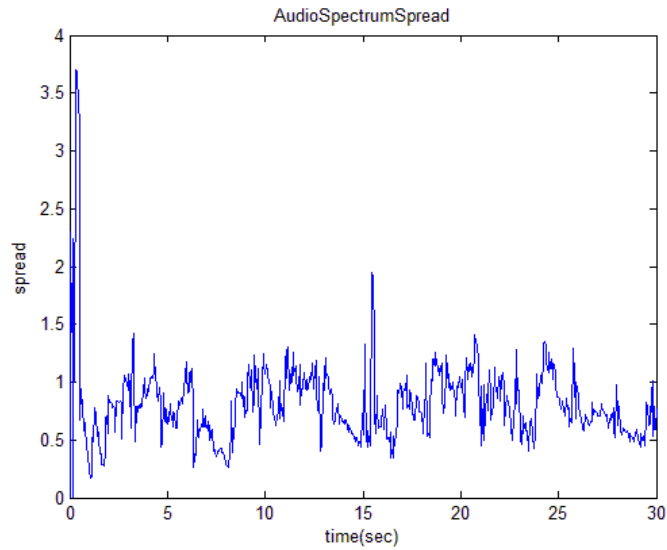


圖 26. Over the Rainbow 藤田惠美 -ASS 分析

(4)AudioSpectrumFlatness

這項特徵主要用來描述頻譜的平坦程度，在這邊會分成各頻段再進一步的計算各頻段間的頻譜平坦程度，每個頻段是 1/4 個 Octave，共計會產生 24 個頻段，每個頻段中的 ASF 計算公式如下：

$$ASF = \frac{\sqrt[b]{\prod_b c(i)}}{\frac{1}{b} \sum_b c(i)} \quad (5)$$

其中的 $c(i)$ 是第 b 個頻段內頻譜的參數。每個音框最後的計算結果是一組 24 維的特徵值，代表著該音框在 24 個頻段內的平坦程度結果如下圖：

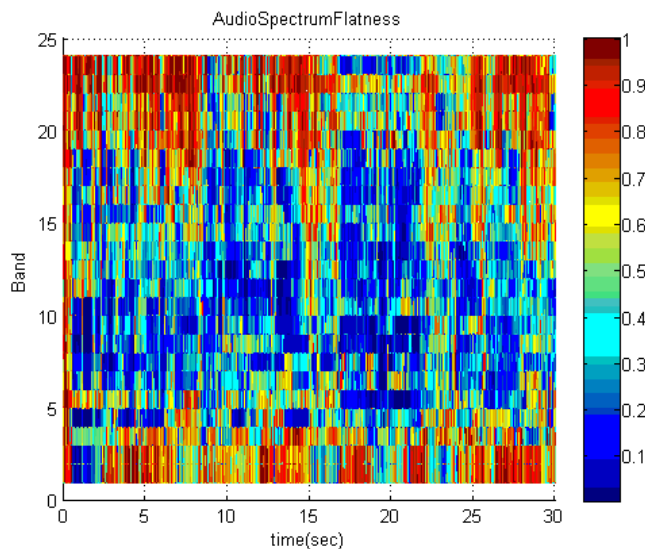


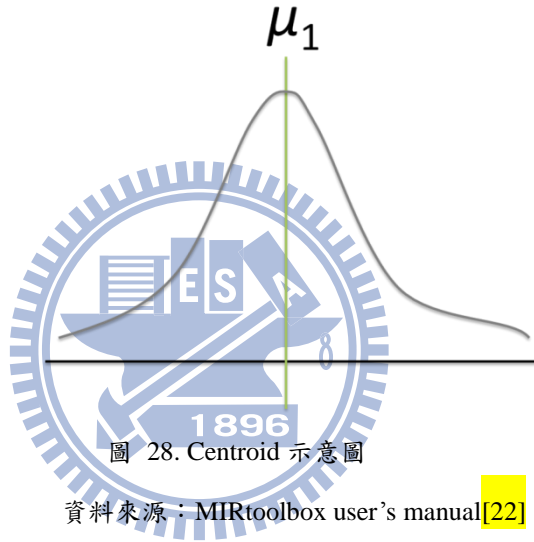
圖 27. Over the Rainbow 藤田惠美 -ASF 分析

(5)Centroid

除了採用Mpeg-7的音訊特徵外，本研究也採用MIRtoolbox中包含的音訊特徵，這邊的Centroid和Mpeg-7的頻譜質心的差別在於輸入的音框大小不同，計算頻域波形的一階動差(first moment)也就是期望值(mean)、資料分布的幾何中心可以表示頻譜的集中趨向，計算公式如下：

$$\mu_1 = \int xf(x)dx \quad (6)$$

μ_1 即為幾何中心，可以見下圖表示：



每個音框會取出單一個數值特徵，計算的結果如下圖所示：

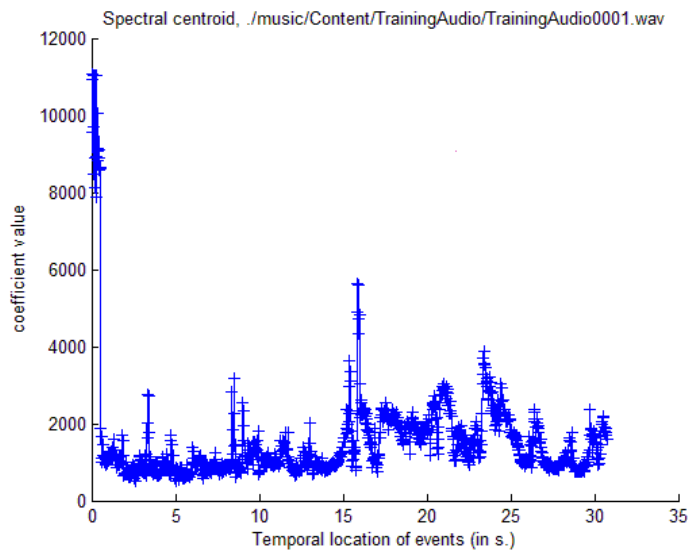


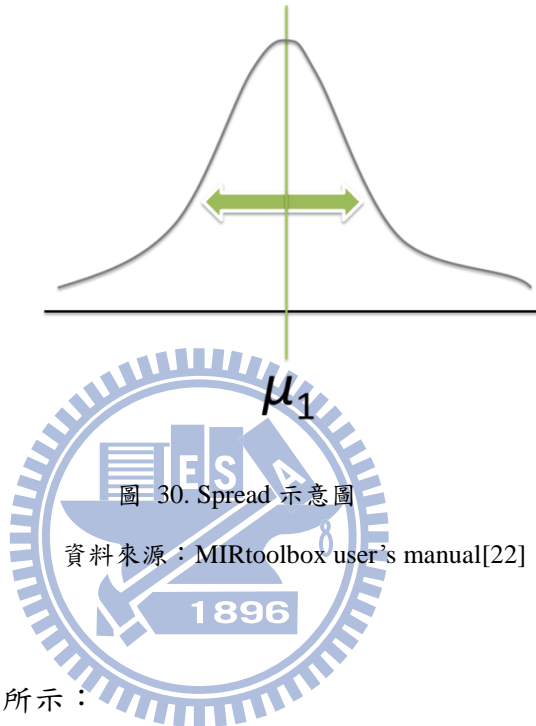
圖 29. Over the Rainbow 藤田惠美 -Centroid 分析

(6) Spread

如同 Mpeg-7 中的 Spread 概念，進一步的計算二階中心動差，其公式如下所示：

$$\sigma^2 = \mu_2 = \int (x - \mu_1)^2 f(x) dx \quad (7)$$

結果以 σ^2 表示，其計算的概念圖如下：



計算的結果如下圖所示：

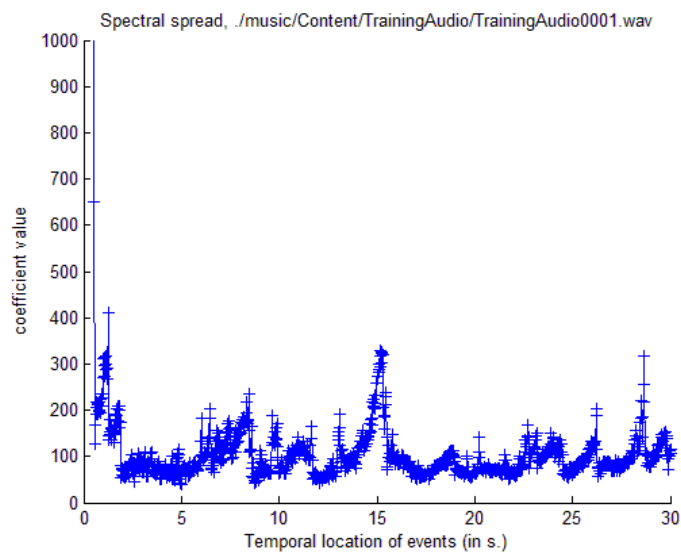


圖 31. Over the Rainbow 藤田惠美 -Spread 分析

(7) Flatness

這特徵表示音訊的頻域波形平坦度，可以顯示出頻譜是滑順的或者是有很多突起狀，計算的公式如下：

$$\text{flatness} = \frac{\sqrt[N]{\prod_{n=0}^{N-1} x(n)}}{\left(\frac{\sum_{n=0}^{N-1} x(n)}{N}\right)} \quad (8)$$

計算上是該音框內的幾何平均值和算術平均值的比率，結果如下圖：

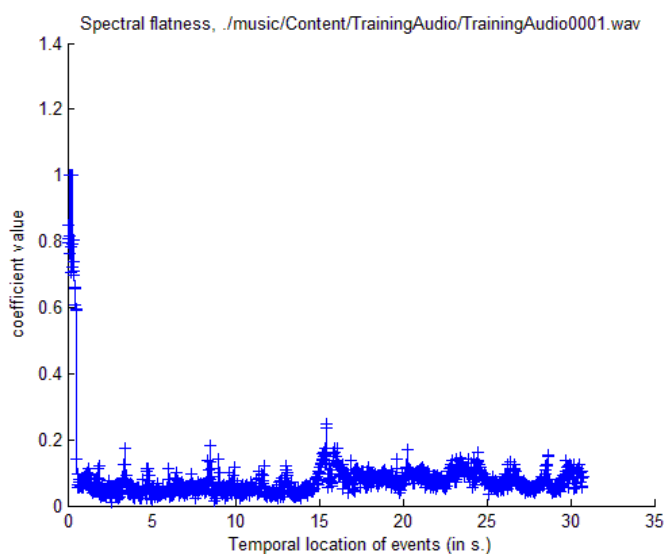


圖 32. Over the Rainbow 藤田惠美 -Flatness 分析

2-6 SVM 分類演算法

為了因應高維度的特徵值分類，本研究採用了 SVM 作為分類機制，關於 SVM 的詳細原理參考[23][24]，在 SVM 中我們假設有一堆點集合 $\{x_i, y_i\}$ ， $i = 1, \dots, n$ 且 $x_i \in R^d$ ， $y_i \in \{+1, -1\}$ ， x_i 即為測試資料，而 y_i 為各筆測試資料的類別標籤。這邊希望找到一個最佳超平面(Optimal separating hyperplane，OSH) $f(x) = w^T x - b$ 可以剛好分隔開兩個類別，使得 $y_i = -1$ 的點落在 $f(x) < 0$ 的這一邊，且 $y_i = +1$ 的都落在 $f(x) > 0$ 的那一邊，如此一來即可以 $f(x)$ 的正負來判定接著進來的測試資料類別。

為了達到最佳的分類效果要這邊要考慮與 OSH 平行的兩個支援超平面(Support

hyperplane)，即為圖 6.裡那兩條虛線，將支援超平面寫成下面式子表示：

$$w^T x = b + \delta \quad (9)$$

$$w^T x = b - \delta \quad (10)$$

這邊形成了一個過多待解參數的問題(over-parameterized problem)，假設 x 、 b 、 δ 乘上一任意常數等式依然成立，意即 x 、 b 、 δ 有無限多組解，則為了簡化問題在等式中皆乘上一個常數使得 $\delta = 1$ ，藉此消去一個待解參數，在接下來的討論中 x 、 b 皆已經過同樣的尺度調整。

求取 OSH 主要可以視為要找出有最大邊界(Margin)的支援超平面，這邊定義 Separating plane 到支援超平面的距離為 d 如下兩式：

$$d = \frac{(|b + 1| - |b|)}{\|w\|} = \frac{1}{\|w\|}, \quad b \notin (-1, 0) \quad (11)$$

$$d = \frac{(|b + 1| + |b|)}{\|w\|} = \frac{1}{\|w\|}, \quad b \in (-1, 0) \quad (12)$$

則要求的邊界等於 $2d$ ，

$$\text{margin} = 2d = \frac{2}{\|w\|} \quad (13)$$

由上式可見當 $\|w\|$ 越小則可以得到越大的邊界。由於之前的設計過尺度調整的，可以知道支援超平面和 OSH 的距離在 ± 1 之內，這個限制條件可寫成下列式子：

$$w^T x_i - b \leq -1, \quad \forall y_i = -1 \quad (14)$$

$$w^T x_i - b \geq +1, \quad \forall y_i = +1 \quad (15)$$

上兩式可以進一步的合併如下：

$$y_i(w^T x_i - b) - 1 \geq 0 \quad (16)$$

總和上面討論，可以得到下列的目標函式為 SVM 尋求 OSH 的主要問題(primal problem)，

$$\min_w \frac{1}{2} \|w\|^2$$

$$\text{並符合 } y_i(w^T x_i - b) - 1 \geq 0 \quad \forall i \quad (17)$$

為了求解上述的最佳化問題，採用拉格朗日法(Lagrange Multiplier Method)將上述式子轉乘對偶問題(dual problem)的二次方程式，找出可以使 L 為最小值的 w 、 b 、 α_i ，其中 α_i 是拉式乘數(Lagrange Multiplier)，式子如下：

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^N \alpha_i [y_i (w^T x_i - b) - 1] \quad (18)$$

接著求解 L 的最小值分別對 w 及 b 作偏微分，

$$\partial_w L_p = 0 \rightarrow w - \sum_{i=1}^N \alpha_i y_i x_i = 0 \Rightarrow w^* = \sum_{i=1}^N \alpha_i y_i x_i \quad (19)$$

$$\partial_b L_p = 0 \rightarrow \sum_{i=1}^N \alpha_i y_i = 0 \quad (20)$$

將(18)、(19)代回(17)可以得到新的目標函式為，

$$\max L_D = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{ij} \alpha_i \alpha_j y_i y_j x_i^T x_j$$

$$\text{並符合 } \sum_{i=1}^N \alpha_i y_i = 0, \forall \alpha_i \geq 0 \quad (21)$$

總合上述條件如下：

$$\partial_w L_p = 0 \rightarrow w - \sum_{i=1}^N \alpha_i y_i x_i = 0 \quad (18)$$

$$\partial_b L_p = 0 \rightarrow \sum_{i=1}^N \alpha_i y_i = 0 \quad (19)$$

$$y_i (w^T x_i - b) - 1 \geq 0 \quad (15)$$

$$\text{Largrange multiplier condition } \alpha_i \geq 0 \quad (22)$$

$$\text{complementay slackness } \alpha_i [y_i (w^T x_i - b) - 1] = 0 \quad (23)$$

上述條件被稱為 KKT 條件(Karush-Kuhn-Tucker condition)，當我們將訓練資料放進來時有些點會符合上述的 KKT 條件，由(15)和(22)可得知 $y_i (w^T x_i - b) - 1 > 0$ 的話， $\alpha_i = 0$ ，而當 $y_i (w^T x_i - b) - 1 = 0$ 時， $\alpha_i > 0$ ，訓練用且 $\alpha_i > 0$ 的樣本，這些點就是剛好落在支援超平面上的點決定了這個問題的解，一般狀況下這些點很少稱為稀疏 (sparse) 解，這些點稱之為 Support vector 其 $\alpha_i \geq 0$ 。

求得 Support vector 之後就可以用它們來判別新加入的測試資料是屬於哪一邊的集合，

$$f(x) = w^{*T} x - b^* = \sum_i \alpha_i y_i x_i^T x - b^* \quad (24)$$

並且由 complementay slackness 可以得到下式，其中 i 為 Support vector，

$$b^* = \left(\sum_j \alpha_j y_j x_j^T x_i - y_i \right) \quad (25)$$

上述討論中都是假設資料都位在線性空間，非線性資料在實際情況下會把資料映射到高維度的空間或是特徵空間中 $x \rightarrow \phi(x)$ ， ϕ 是映射函數，透過 ϕ 將 x 映射到特徵空間中，映射函數可能是一個非常複雜且不易求值的函數，但是內積的形式可以使其變得簡單，

$$\therefore x_i^T x_j \rightarrow \phi(x_i)^T \phi(x_j) \quad (26)$$

映射函數做內積得到的函數在 SVM 中稱之為核化(Kernel)。

最後討論資料不可完全分割的情形，真實使用上的情況通常比假設更為複雜，上述討論都是基於可以找到 OSH 對資料做完美分割，但在資料本身也有可能誤差下可能找不到一個 OSH 將資料做完美的切割，此時要再導入一個誤差項 ξ 來解決資料在邊界地方有重疊的情形。

$$w^T x_i - b \leq -1 + \xi_i, \quad \forall y_i = -1 \quad (27)$$

$$w^T x_i - b \geq +1 - \xi_i, \quad \forall y_i = +1 \quad (28)$$

$$\xi_i \geq 0 \quad \forall i$$

誤差項 ξ 允許原先的限制被違反，為了使限制條件依然存在意義，會額外再給誤差值一點懲罰性成本 C ，

$$\text{Cost} = C \left(\sum_i \xi_i \right)^k \quad (29)$$

當 $k=1, 2$ 時，這仍是一個二次規劃的問題，下面討論中選擇 $k=1$ ， C 的值將控制懲罰跟邊緣的平衡，要將一個訓練資料放置錯誤的一邊則 $\xi_i > 1$ ，因此 $\sum_i \xi_i$ 可以解釋違規的情形有多嚴重，同時也是違規的限制個數上限。於是新的主要問題變成，

$$\min \frac{1}{2} \|w\|^2 + C \sum_i \xi_i$$

$$\text{並符合 } y_i(w^T x_i - b) - 1 + \xi_i \geq 0 \quad \forall i \quad (30)$$

$$\xi_i \geq 0 \quad \forall i$$

接著一樣使用拉格朗日法引入限制條件，

$$L(w, b, \xi, \alpha, \mu) = \frac{1}{2} \|w\|^2 + C \sum_i \xi_i - \sum_{i=1}^N \alpha_i [y_i(w^T x_i - b) - 1 + \xi_i] - \sum_{i=1}^N \mu_i \xi_i \quad (31)$$

並依此得到新的 KKT 條件，

$$\partial_w L_p = 0 \rightarrow w - \sum_{i=1}^N \alpha_i y_i x_i = 0 \quad (32)$$

$$\partial_b L_p = 0 \rightarrow \sum_{i=1}^N \alpha_i y_i = 0 \quad (33)$$

$$\partial_\xi L_p = 0 \rightarrow C - \alpha_i - \mu_i \quad (34)$$

$$y_i(w^T x_i - b) - 1 \geq 0 \quad (35)$$

$$\xi_i \geq 0 \quad (36)$$

$$\text{Largrange multiplier condition} \quad \alpha_i \geq 0 \quad (37)$$

$$\text{Largrange multiplier condition} \quad \mu_i \geq 0 \quad (38)$$

$$\text{complementay slackness} \quad \alpha_i [y_i(w^T x_i - b) - 1 + \xi_i] = 0 \quad (39)$$

$$\text{complementay slackness} \quad \mu_i \xi_i = 0 \quad (40)$$

由上述條件可見，假設 $\xi_i > 0$ 則 $\mu_i = 0$ 因此 $\alpha_i = C$ 且 $\xi_i = 1 - y_i(w^T x_i - b)$ ，若當 $\xi_i = 0$ 則 $\mu_i > 0$ 因此 $\alpha_i < C$ 。只有當 $\xi_i = 0$ 時可以得到原先的結果：

$y_i(w^T x_i - b) - 1 = 0$ 時， $\alpha_i > 0$ ，否則 $y_i(w^T x_i - b) - 1 > 0$ 的話， $\alpha_i = 0$ 。跟原先的結論相同，若樣本不落在支持超平面上而是在正確分類的那一邊，則 $\xi_i = \alpha_i = 0$ ；如果落在支持超平面上，則 $\xi_i = 0$ 但是 $\alpha_i > 0$ ；但若是落在錯誤分類的那一邊，則 α_i 最大可以到 C ，但是 ξ_i 會平衡被違反的限制使得 $y_i(w^T x_i - b) - 1 + \xi_i = 0$ 。

在眾多分類器中選用了 SVM 在於可以進行多維度特徵分類且方便使用，本研究採用了 Chih-Jen Lin [18] 統整的 (lib)SVM 的 toolbox 是一套以 MATLAB 為平台統合各種 SVM 的演算法，這邊只採用了預設的 SVM 型態—C-SVC，內容包含非線性和誤差值的修正。

2-7 情緒計分

在萃取完特徵值和決定分類演算法後回到原先設計的情緒成分表示來討論，在情緒成分表示的圖 9. 中所示的四種情緒成分，其成分的分布依舊隱含各情緒模型的二維觀念，在判定當前音框的情緒時，主要先將該音框萃取一特徵值，利用特徵值在任兩情緒間做分類，產生六種分類結果最後再取眾數，採用六種情緒判定中被判定最多的情緒作為該音框在該特徵下的情緒。

2-7-1 情緒分類模型建立

在四種情緒成分下，任兩個情緒間的分類共有 12 種比較，扣除重複的比較就剩下圖中的 6 種比較：

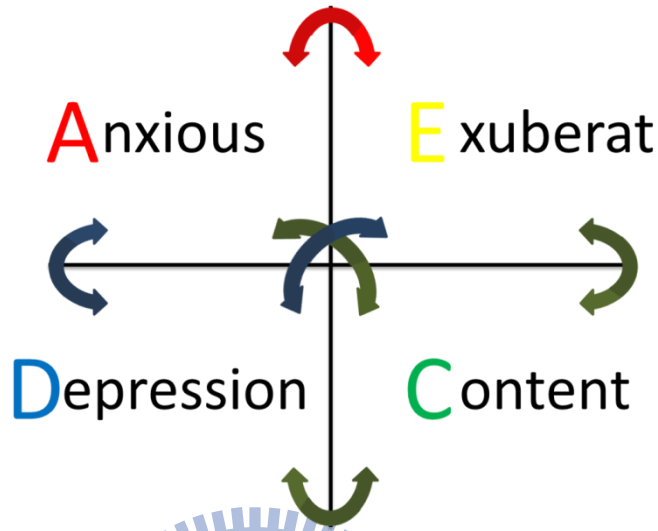


圖 33. 六種情緒分類示意圖

在這六種分類中除了是兩兩情緒間的分辨，對映到 Russell's circumplex model 上更是能表示在高 Arousal 下的比較 Valence 的情緒分布其況或者是高 Valence 下比較 Arousal 的情緒分布等。採用各特徵值採用 SVM 在六個比較下的辨識率可見下兩表：

表 6. 音樂特徵辨識率

	Content Depression	Content Anxious	Content Exuberant	Depression Anxious	Depression Exuberant	Anxious Exuberant
Key	45.82%	50.38%	50.47%	50.21%	55.39%	51.79%
Key clarity	56.86%	52.41%	58.97%	55.03%	50.57%	55.47%
Mode	55.94%	51.57%	57.50%	52.27%	63.18%	58.64%
Tonal centroid	52.23%	79.79%	70.93%	82.36%	73.78%	58.78%

表 7. 音訊特徵辨識率

	Content Depression	Content Anxious	Content Exuberant	Depression Anxious	Depression Exuberant	Anxious Exuberant
Spread	47.23%	89.41%	77.61%	90.39%	77.14%	64.40%
Flatness	64.95%	67.52%	65.76%	80.23%	77.66%	50.26%
Centroid	60.09%	77.86%	71.89%	85.04%	80.26%	54.40%
ASC	57.71%	63.31%	64.98%	51.96%	53.81%	51.03%
ASE	54.48%	79.98%	70.29%	87.41%	73.90%	51.46%
ASS	51.25%	88.37%	82.74%	88.93%	84.57%	57.57%
ASF	67.91%	79.10%	69.27%	85.17%	75.82%	68.81%

上兩表是由各情緒一半的訓練歌曲當成建立 SVM 模型用的訓量資料，並拿另一半的訓練歌曲當作測試資料去做訓練資料的內部測試，目的在於分析各個特徵值對於該配對情緒的分類的辨識能力。

最後在事前處理的部分會將 192 訓練歌曲依照這六種分類對各情緒間的分類建立 SVM 的分類模型，每種特徵值都會有其對應的 6 個情緒分類 SVM 模型，11 個特徵共計 66 組分類模型。

2-7-2 情緒判定

在事前的訓練中建立了分類的模型，於是當測試資料進來時將採用這個模型進行第一步驟的情緒判定，在第一步驟中會得到 6 個情緒判斷結果，之後再進行第二步驟的判定由前六種情緒判定結果去投票決定出當下的情緒，下表以 Celine Dion 的 My heart will go on 的各特徵第一個音框的情緒判定結果為例：

表 8. 情緒判定舉例

	Content Depression	Content Anxious	Content Exuberant	Depression Anxious	Depression Exuberant	Anxious Exuberant	result
Key	Depression	Anxious	Exuberant	Anxious	Exuberant	Anxious	Anxious
Key clarity	Content	Content	Content	Anxious	Exuberant	Anxious	Content
Mode	Content	Anxious	Content	Anxious	Exuberant	Anxious	Anxious
Tonal centroid	Content	Anxious	Content	Anxious	Exuberant	Anxious	Anxious
Spread	Depression	Anxious	Content	Anxious	Exuberant	Anxious	Anxious
Flatness	Content	Anxious	Content	Anxious	Exuberant	Anxious	Anxious
Centroid	Depression	Anxious	Content	Anxious	Exuberant	Anxious	Anxious
ASC	Depression	Anxious	Exuberant	Anxious	Exuberant	Exuberant	Exuberant
ASE	Depression	Anxious	Exuberant	Anxious	Exuberant	Exuberant	Exuberant
ASS	Depression	Content	Content	Depression	Depression	Exuberant	Depression
ASF	Content	Content	Exuberant	Anxious	Exuberant	Exuberant	Exuberant

因此每個特徵值都會在各自的音框中決定一種情緒，若是同一時間點上不同特徵值決定了不同的情緒，那也符合本研究的原先假設，聆聽音樂引發的情緒是由四種基本的情緒疊加而成的結果，而不只是單一情緒可以描述的。各特徵值分析的音框長度不同，因此情緒判定對應到時間軸上如下圖：

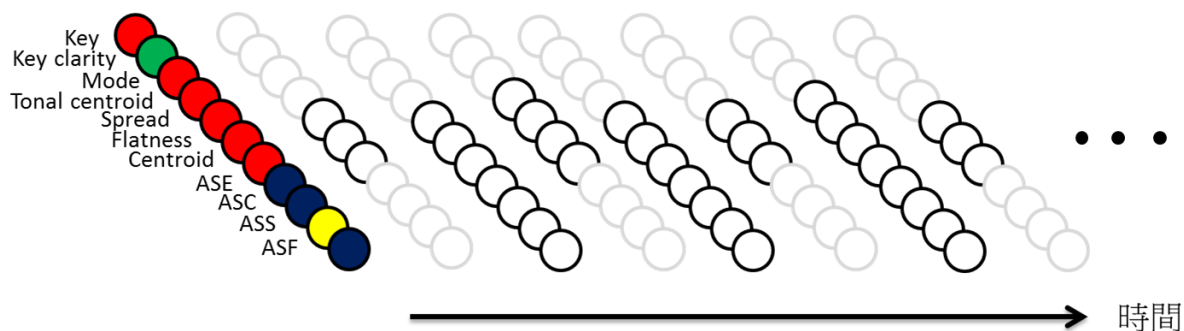


圖 34. 情緒判定示意圖

上圖中每一個圓圈代表一個音框，情緒變化時間以最小音框(0.025 秒)為基本，由於各特徵值採用的音框大小不盡相同，並非所有時刻都有採計所有特徵的情緒判定，如 Key、Key clarity、Mode 對情緒結果的影響每 0.5 秒累加一次；Tonal centroid 是每 0.075 秒累加一次；Spread、Flatness、Centroid 是每 0.025 秒一次也就是基本的情緒成分變化的時間單位；ASE、ASC、ASS、ASF 則是每 0.05 秒累加一次。

2-7-3 計分方式

情緒累積計分上將考慮人類的情緒響應並不是快速的切換，因此每個情緒的判定後的影響不會馬上結束，而是會持續的累積到接下來的音框判斷中，但是其影響會漸漸衰弱，因此會乘上一個衰減函數。在本研究中設定情緒的響應時間是 3 秒，意即在當下判定的情緒會持續影響接著 3 秒內的情緒成分多寡，依此設計了 3 種情緒的衰減加權方式：線性衰減、高斯函數衰減、自然指數衰減，如下圖：

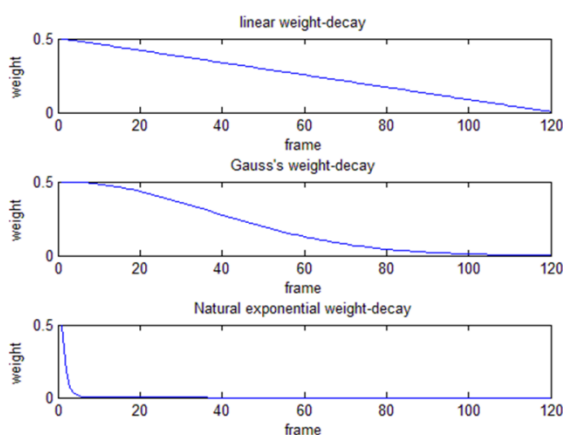


圖 35. 各種衰減加權函數

計分方式以數學公式表示如下：

$$\text{(linear weight - decay)} \quad E_i = \sum_j \sum_{t=1}^{120} 0.5 * \left(\frac{x}{120}\right) \quad (41)$$

$$\text{(Gauss's weight - decay)} \quad E_i = \sum_j \sum_{t=1}^{120} 0.5 * e^{\frac{-\left(\frac{5x}{120}-c\right)^2}{2\sigma^2}} \quad (42)$$

$$\text{(Natural exponential weight - decay)} \quad E_i = \sum_j \sum_{t=1}^{120} 0.5 * \frac{e^x}{e^{120}} \quad (43)$$

$$x = 121 - t, i = mode(P_{1-t})$$

其中 t 是情緒成分變化的最小音框即 0.025 秒， P_{1-t} 是訓練資料在各音框裡第一步驟的六種判定結果， i 則是表示情緒成分類別， j 表示 11 種不同的特徵值， E_i 為當下的各特徵值判定的情緒成分累積分數。由公式中可得知每個當下的情緒判定都將追溯前 120 個基本音框(意即 3 秒)的判定結果，並乘上不同的衰減函數再累計起來。在本研究中設定線性衰減加權函數為預設值。

2-8 音樂情緒成分表示法

在上述的流程處理過後會產生各情緒成分的向量，其中每個元素都代表每 0.025 秒音框大小的情緒判定結果，最後的結果採用四個圓圈表示並輸出成影片，每個時刻四個情緒的累計分數不同，因此四個圓圈也不同而是隨著時間變化的。由於最小的時間單位為 0.025 秒，所以情緒的解析度等於 40 fps，每秒內會有 40 次情緒判定，並隨時計算各情緒的比例，影片畫面如下：

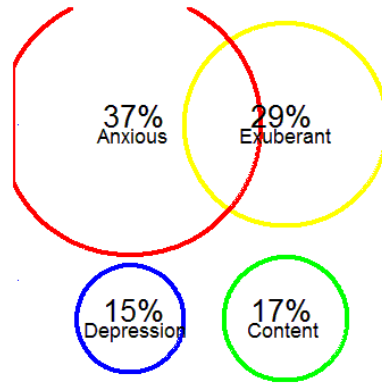


圖 36. 實際影片結果畫面

在整首歌曲播畢，會統計各個情緒成分在歌曲中總佔有的比例，如下圖：

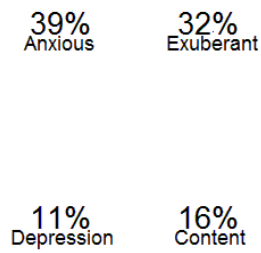


圖 37. 影片最後的情緒總結

完整影片的 DEMO 可至 YOUTUBE 觀賞，(網址：http://youtu.be/-e9G_viDD9U)。

三、 音樂情緒心理分析調查

3-1 問卷調查

前述章節中，本研究建立了一套視覺化即時性的情緒成分分析系統，音樂情緒是主觀的感受，因此為了驗證系統分析的結果是否具有客觀性，在研究的最後進行了針對音樂情緒的問卷調查。也藉此比對由數理演算分析的音樂情緒和心理學家建立的音樂情緒模型結果是否吻合。

在此基於 Zentner 的 GEM-9 情緒模型設計了一份問卷，問卷可見附錄二，來驗證結果的可靠度。測試音樂如下表：

表 9. 本研究測試音樂

演出者	曲名	長度
Avenged Sevenfold	Dear God	00:06:34
Celine Dion	My heart will go on	00:04:41
Chopin	Nocturne opus 9 no 2	00:04:08
Debussy	Claire de lune	00:05:06
Carpenters	Goodbye to love	00:03:56
Edith Piaf	La Vie En Rose	00:03:05

問卷調查對象共計 66 位，一年級到四年級的大學生且來自於各系所有不同的背景，詳細受測者資料可見附錄三。問卷調查過程為不事先告知這是一個關於音樂情緒辨識問卷的調查，而是讓聽者在沒有預設立場下先行聆聽過歌曲後再填問卷，問卷上針對各情緒形容詞有統一的翻譯，以降低聽者對情緒形容詞不認識或認知上落差的情形。

3-2 問卷調查與實驗結果分析

各測試歌曲的計分統計及情緒比例圖形可見附錄四，下面針對幾首結果分析，可以搭配觀看附錄中 DEMO 網址的影片來比較，Avenged Sevenfold - Dear God 的問

卷調查的結果和實驗結果比較如下表，

表 10. Avenged Sevenfold - Dear God 結果分析

問卷調查		程式結果	
情緒類別	情緒比例	情緒成分	情緒比例
Wonder	9.66%	Content	16.48%
Transcendence	9.66%		
Tenderness	10.90%		
Nostalgia	13.39%	Depression	11.71%
Peacefulness	7.94%		
Power	12.97%		
Joyful Activation	10.70%	Anxious	39.57%
Tension	14.70%		
Sadness	10.08%		
		Exuberant	32.24%

由上表中的問卷調查結果可以知道，懷舊、憂鬱(Nostalgia)和動力(Power)和激動(Tension)這三個情緒的比例較高，在 Dear God 這首歌之中大致可以分兩個段落，主歌觸人心弦的滑音吉他和鋪層的主唱旋律都是比較抑鬱的，在副歌激昂的歌聲中再一口氣爆發出來，後面一段破音吉他加進來的橋段再接木吉他的平靜，最後在漸漸進到吉他獨奏的橋段。歌曲的編曲安排非常巧妙，一般聆聽者的情緒會漸漸的被帶領到最後的高潮，因此在問卷上會顯示這三種情緒較多。

但在程式中追蹤的是即時的變化，這首歌除了較激動的片段也不乏細膩的橋段，不過在整個變化過程中 Anxious 和 Exuberant 的成分佔高比率的情緒結果抑是可由上表見到的，由於這是首樂團編制的歌曲，在有鼓組編制下大多較能煽動聆聽者激昂的情緒，這兩個情緒在能量上也都是偏高的，且 Anxious 的成分較 Exuberant 大這也體現了問卷中負面情緒 Nostalgia 的影響。

接著分析 Celine Dion - My heart will go on 問卷結果如下，

表 11. Celine Dion - My heart will go on 結果分析

問卷調查		程式結果	
情緒類別	情緒比例	情緒成分	情緒比例
Wonder	11.49%	Content	23.76%
Transcendence	13.05%		
Tenderness	12.40%		
Nostalgia	13.46%	Depression	30.93%
Peacefulness	10.18%		
Power	8.78%		
Joyful Activation	7.23%	Anxious	24.20%
Tension	11.33%		
Sadness	12.07%		
		Exuberant	21.10%

這首曲中也有著主、副歌的段落安排，編曲上以笛聲為主旋律，輔以電子的氛圍長音這也是問卷中超然(Transcendence)的分數較高的原因，當然在女主唱哀傷的腔調，就算聽不懂歌詞也可以受到那浩瀚的悲愴情感，不外乎懷舊、憂鬱(Nostalgia)的情緒是問卷中的最高分。這首曲子的構造和上一首不盡相同，除了音量大的時候會提高 Anxious、Exuberant 的比例外，其餘時間都表現在 Depression 上，超然的感覺也讓 Content 占有不少比例。由於這首歌是鐵達尼號(Titanic)電影主題曲，因此在問卷調查結果是否受聆聽者喚起電影劇情的記憶的影響，也是一個值得探討的議題。

最後看到針對古典音樂 Chopin - Nocturne opus 9 no 2 的問卷結果如下，

表 12. Chopin - Nocturne opus 9 no 2 結果分析

問卷調查		程式結果	
情緒類別	情緒比例	情緒成分	情緒比例
Wonder	10.29%	Content	16.66%
Transcendence	14.06%		
Tenderness	14.38%		
Nostalgia	11.33%	Depression	58.07%
Peacefulness	15.18%		
Power	7.87%		
Joyful Activation	11.73%	Anxious	13.16%
Tension	7.15%		
Sadness	8.03%		
		Exuberant	12.11%

在古典音樂中的音樂內容和流行的歌曲很不相同，比如這首蕭邦的鋼琴獨奏曲，整首歌曲中只有鋼琴的演奏，在內容上能量普遍偏低，因此情緒比例以 Depression 佔大宗，但是可以分出 Content 和 Depression 這可以預想是音樂特徵的功用，針對這點也可以是之後研究的議題。在問卷調查中超然(Transcendence)、溫和(Tenderness)、平靜(Peacefulness)為最大宗，在此與程式結果最大的共通點在於低情緒能量的感受。

其餘測試歌曲問卷結果和影片 DEMO 可查看附錄四。

四、 結論

本研究彙整各種情緒模型提出一種新的視覺化即時性音樂情緒表示方式，有別於一般即時性研究將音樂對只對應到單一情緒，這邊採用四種基本的情緒成分比例來表示，在強烈單一情緒的音樂中也能瞥見細微的其他情緒成分變動。

與 Zentner 的 GEM 模型比較最大問題在於，大家對於情緒形容詞的理解不同。這也同樣存在於受測者認知的情緒形容詞和我們預想的感覺間，也是分析結果造成差異的主因，受測者的對音樂的學習程度也同樣是尚未考慮的變數，對於有音樂學習經驗的人在聽音樂時候對音樂內容感受的重點可能不如一般平常人而是更具學習經驗而有不同感受。

本研究系統較心理學家的情緒模型優勢在於測試資料可以是使用者自行定義的歌曲庫，每個使用者可以以自己感受的情緒音樂片段輸入系統去建立屬於各自專屬的分類模型，這也可以大大降低主觀感受的認知落差。



參考文獻

- [1] Typke, R.; Wiering, F.; Veltkamp, R.C., "A survey of music information retrieval systems", Proc. 6th International Conference on Music Information Retrieval, pp. 153-160, 2005
- [2] Zentner, M. R. & Eerola, T. . "ch8 .Self-report measures and models." , "Handbook of Music and Emotion", Juslin, P. & Sloboda, J. (Eds). , pp. 187-221. Oxford University Press: Oxford, UK., 2009
- [3] José M. Martínez , "Mpeg-7 Overview", 2004 oct.
<http://mpeg.chiariglione.org/standards/mpeg-7/mpeg-7.htm>
- [4] 劉客養, " 初探鋼琴十二平均律調音" , Jewels of the Art World, Journal of Aesthetic Education, No.131, pp.88-95
- [5] 林金慧, "在蒙德里安架構下闡釋情緒的樂理", 台南科大學報, 第 28 期, 生活藝術類, 115-132 頁, 中華民國 98 年 10 月
- [6] Maria M. Ruxanda, Bee Yong Chua, Alexandros Nanopoulos, Christian S. Jensen, "Emotion-based music retrieval on a well-reduced audio feature space", International Conference on Acoustics, Speech, and Signal Processing - ICASSP , pp. 181-184, 2009
- [7] Bin Zhu, Kejun Zhang, "Music emotion recognition system based on improved GA-BP", International Conference On Computer Design And Applications (ICCCA 2010), Volume 2, PP.409-412, 2010
- [8] Dan Liu, Lie Lu, Hong-Jiang Zhang "Automatic Mood Detection from Acoustic Music Data", Music Information Retrieval, pp. 13-17, 2003.
- [9] Ei Ei Pe Myint, Moe Pwint, "An approach for mulit-label music mood classification", 2010 2nd International Conference on Signal Processing Systems, Volume 1, PP.290-294, 2010
- [10] Jonghwa Kim, Ande, E., "Emotion Recognition Based on Physiological Changes in

Music Listening”, Pattern Analysis and Machine Intelligence, IEEE Transactions on,
VOL. 30, NO. 12, Dec 2008

- [11] Sanghoon Jun, Seungmin Rho, Byeong-jun Han, Eenjun Hwang, “A fuzzy inference-based music emotion recognition system”, 5th International Conference on In Visual Information Engineering (VIE), pp. 673-677, 2008
- [12] Y.-H. Yang et al, “Mr. Emo: Music retrieval in the emotion plane,” Proc. ACM MM, 2008, accepted.
- [13] Mark David Korhonen, “Modeling Continuous Emotional Appraisals of Music Using System Identification”, A thesis presented to the University of Waterloo, Canada, 2004
- [14] Jun-Jie Fu, “Emotion Locus Tracking System for Automatic Mood Detection and Classification of Music Signals” A thesis Submitted to National Chiao Tung University, Taiwan, 2010
- [15] Emery Schubert , “Real Time Cognitive Response Recording”, The inaugural International Conference on Music Communication, Science 5-7, ,PP.139-142, Dec 2007
- [16] 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, 2010
- [17] Zentner, Marcel; Grandjean, Didier; Scherer, Klaus R., ”Emotions evoked by the sound of music: Characterization, classification, and measurement.” Emotion, Vol 8(4), 494-521. , Aug 2008
- [18] Chih-Chung Chang,Chih-Jen Lin, ”A Library for Support Vector Machines”, ACM Transactions on Intelligent Systems and Technology, Volume 2, Issue 3, Article 27, April 2011
- [19] MIR toolbox,
<https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox>
- [20] MPEG-7 Multimedia Software Resources, <http://mpeg7.doc.gold.ac.uk/index.html>

- [21] Olivier Lartillot, Petri Toiviainen , “A Matlab Toolbox for Musical Feature Extraction From Audio”, the 10th Int. Conference on Digital Audio Effects (DAFx-07), Bordeaux, France, September 10-15, 2007
- [22] Olivier Lartillot, ”Mir toolbox1.3.2 user’s manual”, Finnish Centre of Excellence in Interdisciplinary Music Research, University of Jyväskylä, Finland, January, 19th, 2011
- [23] 李根逸, "支持向量機教學文件(中文版)", 台灣大學通訊與多媒體實驗室
- [24] 林宗勳, ”Support Vector Machines 簡介”, 台灣大學通訊與多媒體實驗室, 2006



附錄一. 訓練資料曲目

#	Title	Artist	EmotionType
1	Over the Rainbow	藤田惠美	Content
2	Over the Rainbow	藤田惠美	Content
3	Don't know Why	Norah Jones	Content
4	Don't know Why	Norah Jones	Content
5	Melody Fair	藤田惠美	Content
6	Melody Fair	藤田惠美	Content
7	Depois do natal	小野麗莎	Content
8	Depois do natal	小野麗莎	Content
9	All My Loving	藤田惠美	Content
10	Jingle Bell Rock	小野麗莎	Content
11	White Christmas	小野麗莎	Content
12	White Christmas	小野麗莎	Content
13	White Christmas	小野麗莎	Content
14	Paz Azul	小野麗莎	Content
15	Paz Azul	小野麗莎	Content
16	The Christmas Song	小野麗莎	Content
17	The Christmas Song	小野麗莎	Content
18	Boas Festas	小野麗莎	Content
19	Boas Festas	小野麗莎	Content
20	Um Anjo Do Ceu	小野麗莎	Content
21	Um Anjo Do Ceu	小野麗莎	Content
22	The wee small hours of the morning	小野麗莎	Content
23	The wee small hours of the morning	小野麗莎	Content
24	leaving on a jet plane	藤田惠美	Content
25	leaving on a jet plane	藤田惠美	Content
26	silent night	小野麗莎	Content
27	not going anywhere	Keren Ann	Content
28	not going anywhere	Keren Ann	Content
29	Polly	Keren Ann	Content
30	Polly	Keren Ann	Content
31	Sit on the Sun	Keren Ann	Content
32	Right Now and	Keren Ann	Content
33	Right Now and	Keren Ann	Content
34	Seventeen	Keren Ann	Content

35	Seventeen	Keren Ann	Content
36	Stephanie Says	Keren Ann	Content
37	Stephanie Says	Keren Ann	Content
38	In My Life	藤田惠美	Content
39	In My Life	藤田惠美	Content
40	Moonglow	Noon	Content
41	Moonglow	Noon	Content
42	I like to recognize the tune	John Pizzarelli	Content
43	Close to You	Noon	Content
44	Mountain Greenery	John Pizzarelli	Content
45	Moon River	Fried Pride	Content
46	Moon River	Fried Pride	Content
47	Fun Out of Love	Noon	Content
48	沒有你的世界	河野伸	Depression
49	我的煩惱	河野伸	Depression
50	9 crimes	Damien Rice	Depression
51	A summer's day	久石讓	Depression
52	A summer's day	久石讓	Depression
53	Resphoina	久石讓	Depression
54	記得	鄭智薰	Depression
55	Symmetry	Mew	Depression
56	The Following Day	Jacques Stotzem	Depression
57	All I need	Radiohead	Depression
58	red sky	David Gilmour	Depression
59	red sky	David Gilmour	Depression
60	the hill	Marketa Irglova	Depression
61	the hill	Marketa Irglova	Depression
62	the hill	Marketa Irglova	Depression
63	view of silence	久石讓	Depression
64	view of silence	久石讓	Depression
65	Le dessous des cartes	Jacques Stotzem	Depression
66	accidental babies	Damien Rice	Depression
67	空房間	The Daydream	Depression
68	summer	久石讓	Depression
69	summer	久石讓	Depression
70	Vent D'Ailleurs	Jacques Stotzem	Depression
71	Gloomy Sunday	Loreena McKennitt	Depression

72	Sleep dont weep	Damien Rice	Depression
73	Secret Garden	Jacques Stotzem	Depression
74	Secret Garden	Jacques Stotzem	Depression
75	walking in the air	藤田惠美	Depression
76	Secret Garden	Jacques Stotzem	Depression
77	legend of wind	久石讓	Depression
78	Fragile dream	久石讓	Depression
79	Fragile dream	久石讓	Depression
80	paun haung	久石讓	Depression
81	寂靜	鄭智薰	Depression
82	Moonlight	Beethoven	Depression
83	who am i	FF7 OST	Depression
84	end of may	Keren Ann	Depression
85	by the cathedral	Keren Ann	Depression
86	ending song	Keren Ann	Depression
87	ending song	Keren Ann	Depression
88	do what i do	Keren Ann	Depression
89	do what i do	Keren Ann	Depression
90	walk real slow	Keren Ann	Depression
91	the morning after	Keren Ann	Depression
92	new born	Muse	Depression
93	Space Dementia	Muse	Depression
94	DarkShines	Muse	Depression
95	Paranoid Android	Radiohead	Depression
96	15 steps	Radiohead	Anxious
97	15 steps	Radiohead	Anxious
98	15 steps	Radiohead	Anxious
99	muscle museum	Muse	Anxious
100	Cave	Muse	Anxious
101	Cave	Muse	Anxious
102	Cave	Muse	Anxious
103	Showbiz	Muse	Anxious
104	Showbiz	Muse	Anxious
105	Showbiz	Muse	Anxious
106	Assasain	Muse	Anxious
107	Assasain	Muse	Anxious
108	Assasain major like	Muse	Anxious

109	Uno	Muse	Anxious
110	Uno	Muse	Anxious
111	CityOfDelusion	Muse	Anxious
112	CityOfDelusion	Muse	Anxious
113	CityOfDelusion	Muse	Anxious
114	Sober	Muse	Anxious
115	Sober	Muse	Anxious
116	Sober	Muse	Anxious
117	Knights of Cydonia	Muse	Anxious
118	Knights of Cydonia	Muse	Anxious
119	Knights of Cydonia	Muse	Anxious
120	Stockholm Syndrome	Muse	Anxious
121	Stockholm Syndrome	Muse	Anxious
122	Stockholm Syndrome	Muse	Anxious
123	ThoughtsofaDyingAtheist	Muse	Anxious
124	ThoughtsofaDyingAtheist	Muse	Anxious
125	ThoughtsofaDyingAtheist	Muse	Anxious
126	NewBorn	Muse	Anxious
127	NewBorn	Muse	Anxious
128	NewBorn	Muse	Anxious
129	NewBorn	Muse	Anxious
130	Space Dementia	Muse	Anxious
131	Space Dementia	Muse	Anxious
132	Hypermusic	Muse	Anxious
133	Hypermusic	Muse	Anxious
134	Hypermusic	Muse	Anxious
135	Citzen Erased	Muse	Anxious
136	Citzen Erased	Muse	Anxious
137	Micro Cuts	Muse	Anxious
138	Micro Cuts	Muse	Anxious
139	Dark Shines	Muse	Anxious
140	Paranoid Android	Radiohead	Anxious
141	Electioneering	Radiohead	Anxious
142	Electioneering	Radiohead	Anxious
143	Electioneering	Radiohead	Anxious
144	skyskysky	DEPAPEPE	Exuberant
145	skyskysky	DEPAPEPE	Exuberant

146	depa canon	DEPAPEPE	Exuberant
147	star	youha	Exuberant
148	star	youha	Exuberant
149	selfish jean	travis	Exuberant
150	selfish jean	travis	Exuberant
151	Tall Fiddler	Tommy Emmanuel	Exuberant
152	Tall Fiddler	Tommy Emmanuel	Exuberant
153	when your mind's made up	Glen Hansard	Exuberant
154	day tripper	Beatles	Exuberant
155	day tripper	Beatles	Exuberant
156	I can do better	Avril Lavigne	Exuberant
157	Electric de Chocobo	FF7 OST	Exuberant
158	Electric de Chocobo	FF8 OST	Exuberant
159	Fiddle de Chocobo	FF9 OST	Exuberant
160	Fiddle de Chocobo	FF10 OST	Exuberant
161	Highwind	FF11 OST	Exuberant
162	Ride on	FF12 OST	Exuberant
163	Ride on	FF13 OST	Exuberant
164	Kurumi	Mr. Children	Exuberant
165	Kurumi	Mr. Children	Exuberant
166	masa ume	Spitz	Exuberant
167	masa ume	Spitz	Exuberant
168	spring song	Spitz	Exuberant
169	spring song	Spitz	Exuberant
170	summer paeade	DEPAPEPE	Exuberant
171	wherever you will go	The Calling	Exuberant
172	wherever you will go	The Calling	Exuberant
173	new world	L'arc en Ciel	Exuberant
174	new world	L'arc en Ciel	Exuberant
175	stay away	L'arc en Ciel	Exuberant
176	stay away	L'arc en Ciel	Exuberant
177	free eat	L'arc en Ciel	Exuberant
178	free eat	L'arc en Ciel	Exuberant
179	star chaser	Spitz	Exuberant
180	star chaser	Spitz	Exuberant
181	Am I Wry No	Mew	Exuberant
182	Eurostar	T Square	Exuberant


183	Eurostar	T Square	Exuberant
184	Fightman	T Square	Exuberant
185	Birdland	Weather Report	Exuberant
186	Birdland	Weather Report	Exuberant
187	the man in the green shirt	Weather Report	Exuberant
188	jerry's breakdown	D'Agostino	Exuberant
189	tennessee rag	Chet Atkins	Exuberant
190	Higher	The Band Apart	Exuberant
191	when you wish upon a star	The Band Apart	Exuberant
192	I love you Wasted Junks & Greens	The Band Apart	Exuberant



附錄二. 音樂情緒問卷

姓名: _____ 學號: _____ 系所年級: _____

音樂情緒分析: 以數理方式分析由音樂內容喚起聽者的情緒。

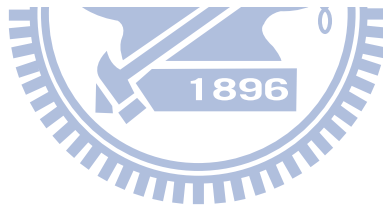
曲名: **La_Vie_En_Rose** 

演唱者: **Edith Piaf**

曲風: **Bossa nova, Pop.**

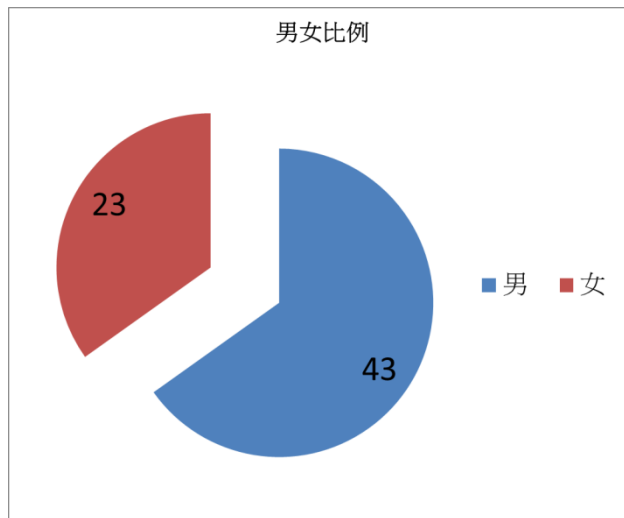
When providing your ratings, please describe how the music you listen to makes you *feel* (e.g., this music makes me feel sad). Do not describe the music (e.g., this music is sad) or what the music may be expressive of (e.g. this music expresses sadness). Keep in mind that a piece of music can be sad or can sound sad without making you feel sad. Please rate the intensity with which you *felt* each of the following feelings on a scale ranging from 1 (*not at all*) to 5 (*very much*).

	1	2	3	4	5	1	2	3	4	5
	Not at all	Somewhat	Moderately	Quite a lot	Very Much					
1	Wonder (驚異; 讚歎): Filled with wonder, Dazzled, Moved									
2	Transcendence (超然; 靈性): Fascinated, Overwhelmed, Feelings of spirituality									
3	Power (動力的; 喜悅的; 狂歡的): Strong, Triumphant, Energetic									
4	Tenderness (溫和; 親切): Tender, Affectionate, In love									
5	Nostalgia (懷舊之情): Nostalgic, Dreamy, Melancholic(憂鬱的)									
6	Peacefulness (平靜的; 寧靜的): Serene, Calm, Soothed									
7	Joyful Activation (高興的, 充滿喜悅的): Joyful, Amused, Bouncy									
8	Sadness (悲哀, 悲傷): Sad, Sorrowful									
9	Tension (激動的): Tense, Agitated, Nervous									

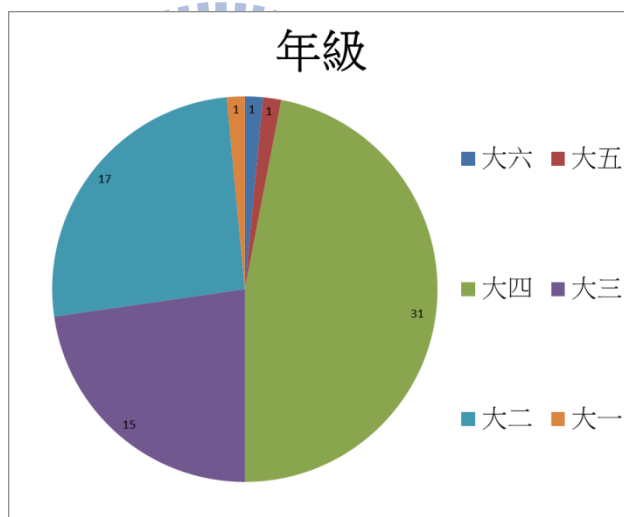


附錄三. 問卷調查受測者資料

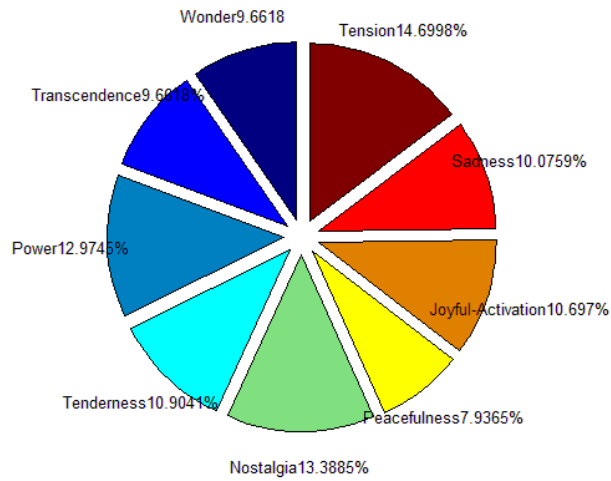
1. 男女分布



2. 年級分布



3. 系別分布



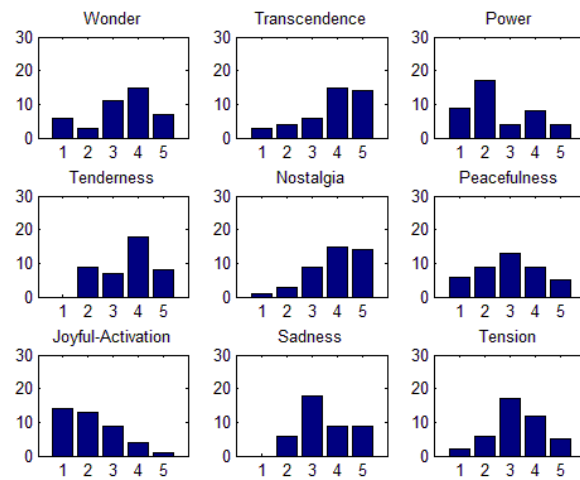
3. 結果分析

<i>Avenge Sevenfold - Dear God</i>		Demo Link: http://youtu.be/U-YI-xUf8lk			
Wonder	9.66%	Sublimity	41.29%	Content	16.48%
Transcendence	9.66%			Depression	11.71%
Tenderness	10.90%			Axious	39.57%
Nostalgia	13.39%			Exuberant	3224.21%
Peacefulness	7.94%	Vitality	23.14%		
Power	12.97%				
Joyful Activation	10.70%	Unease	17.42%		
Tension	14.70%				
Sadness	10.08%				

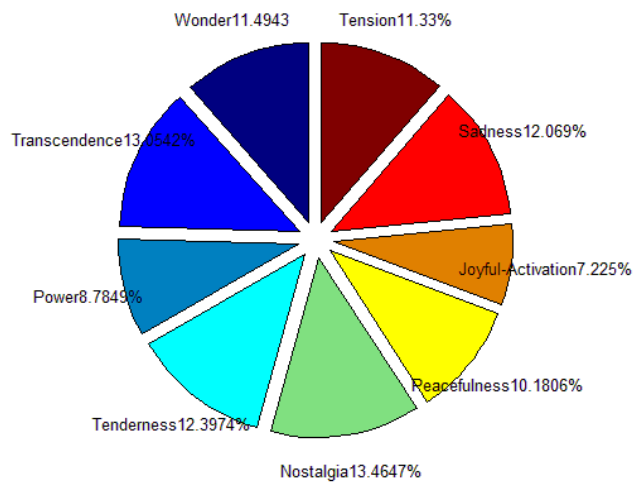


二、 Celine Dion — My heart will go on

1. 問卷計分統計



2. 問卷情緒比例

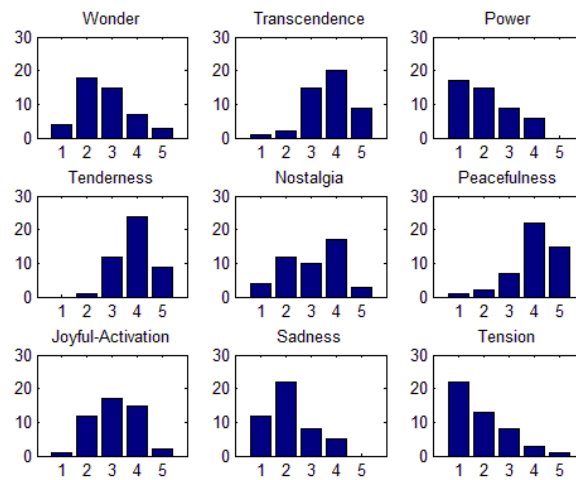


3. 結果分析

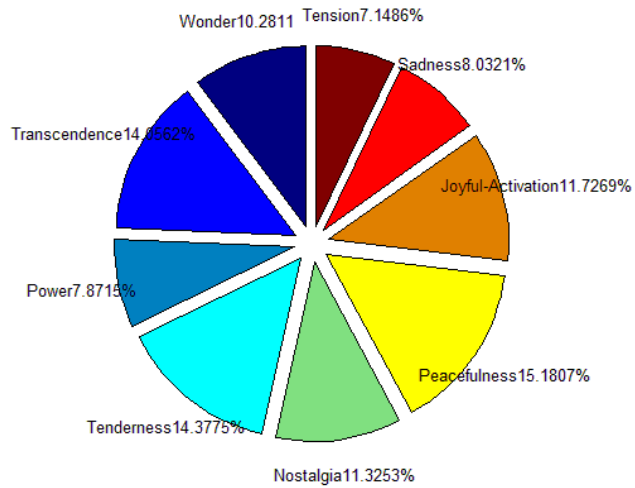
<i>Celine Dion My heart will go on</i>		Demo Link: http://youtu.be/-e9G_viDD9U			
Wonder	11.49%	Sublimity	48.01%	Content	23.76%
Transcendence	13.05%		Depression	30.93%	
Tenderness	12.40%		Axious	24.20%	
Nostalgia	13.46%		Exuberant	21.11%	
Peacefulness	10.18%	Vitality	15.65%		
Power	8.78%		Unease	14.59%	
Joyful Activation	7.23%				
Tension	11.33%				
Sadness	12.07%				

三、 Chopin — Nocturne opus 9 no 2

1. 問卷計分統計



2. 問卷情緒比例

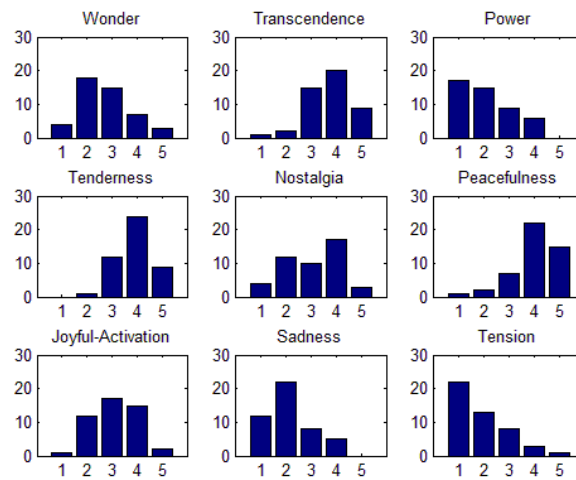


3. 結果分析

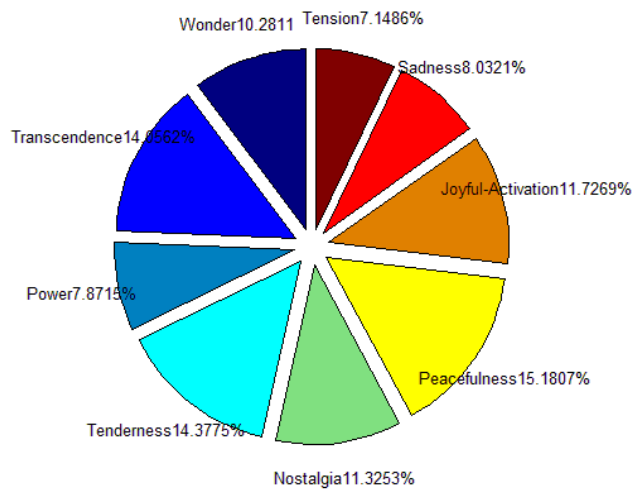
Chopin Nocturne opus 9 no 2		Demo Link: http://youtu.be/gvIsRUqmcze			
Wonder	10.29%	Sublimity	50.30%	Content	16.66%
Transcendence	14.06%		Depression	58.07%	
Tenderness	14.38%		Axious	13.16%	
Nostalgia	11.33%		Exuberant	12.11%	
Peacefulness	15.18%	Vitality	19.01%		
Power	7.87%		Unease	9.32%	
Joyful Activation	11.73%				
Tension	7.15%				
Sadness	8.03%				

四、 Debussy — Claire de lune

1. 問卷計分統計



2. 問卷情緒比例

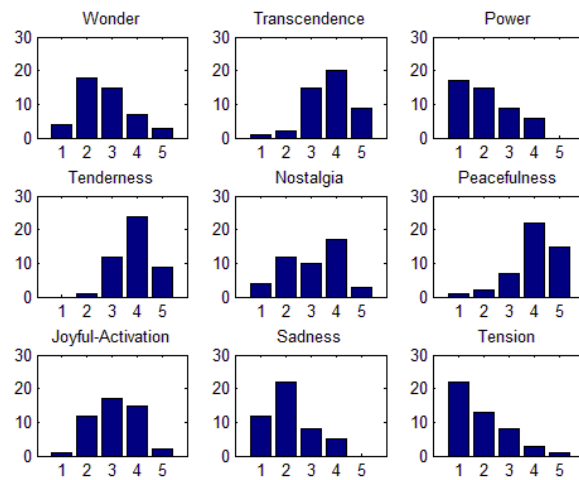


3. 結果分析

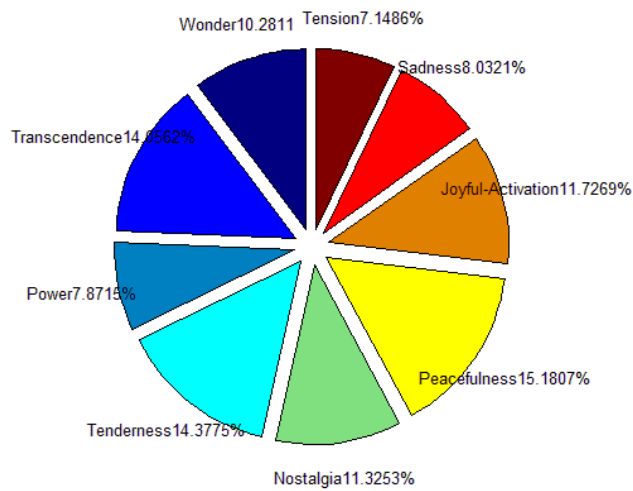
<i>Debussy Claire de lune</i>		Demo Link: http://youtu.be/47D-f82jEjE			
Wonder	11.20%	Sublimity	49.39%	Content	14.04%
Transcendence	13.63%		Depression	61.51%	
Tenderness	13.02%		Axious	12.50%	
Nostalgia	11.42%		Exuberant	11.94%	
Peacefulness	14.24%				
Power	8.23%	Vitality	17.99%		
Joyful Activation	10.28%	Unease	10.91%		
Tension	8.30%				
Sadness	9.67%				

五、 Carpenters — Goodbye to Love

1. 問卷計分統計



2. 問卷情緒比例

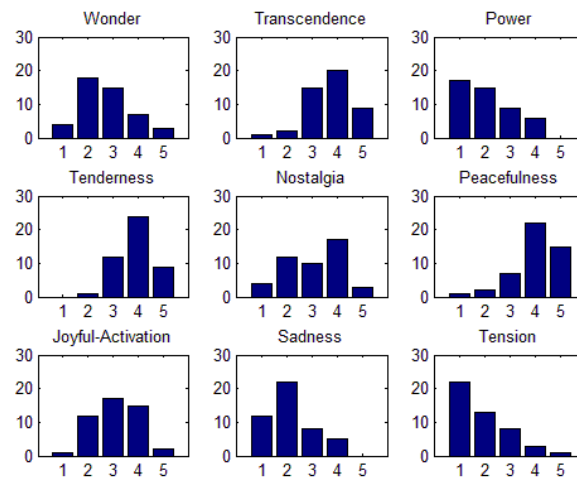


3. 結果分析

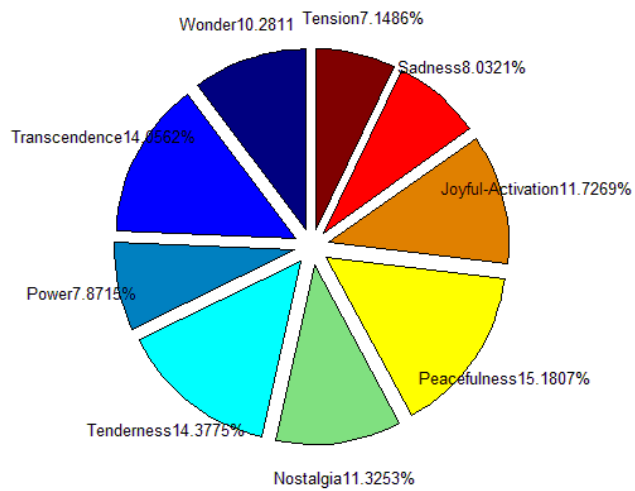
<i>Carpenters Goodbye to Love</i>		Demo Link: http://youtu.be/46FpnuTnWHY			
Wonder	8.61%	Sublimity	46.50%	Content	25.71%
Transcendence	10.88%		Depression	21.34%	
Tenderness	14.60%		Axious	25.79%	
Nostalgia	13.72%		Exuberant	27.17%	
Peacefulness	11.61%	Vitality	20.88%		
Power	9.93%		Unease	12.36%	
Joyful Activation	11.53%				
Tension	9.85%				
Sadness	9.27%				

六、 Edith Piaf — La Vie En Rose

1. 問卷計分統計



2. 問卷情緒比例



3. 結果分析

<i>Edith Piaf La Vie En Rose</i>		Demo Link: http://youtu.be/vagFnzio5ms			
<i>Wonder</i>	9.15%	Sublimity	48.91%	Content	8.43%
<i>Transcendence</i>	10.92%		Depression	45.00%	
<i>Tenderness</i>	14.46%		Axious	36.78%	
<i>Nostalgia</i>	14.46%		Exuberant	9.78%	
<i>Peacefulness</i>	13.58%				
<i>Power</i>	8.78%	Vitality	19.65%		
<i>Joyful Activation</i>	11.44%	Unease	10.46%		
<i>Tension</i>	7.97%				
<i>Sadness</i>	9.23%				