

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

基於低音聲部特徵萃取的和弦行進  
與風格演算法之音樂情緒分類研究

Music Emotion Classification by Genre and Chord

Progression Based on Bass Line Feature Extraction

研究生：謝雲凱

指導教授：鄭泗東 教授

民國九十九年六月

基於低音聲部特徵萃取的和弦行進與風格  
演算法之音樂情緒分類研究

Music Emotion Classification by Genre and Chord  
Progression Based on Bass Line Feature Extraction

研究生：謝雲凱

Student : Yun-Kai Hsieh

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

June 2010

Hsinchu, Taiwan, Republic of China

中華民國九十九年六月

# 基於低音聲部特徵萃取的和弦行進與風格演算法之音樂情緒分類研究

研究生：謝雲凱

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

近年來伴隨著MP3此種以超高壓縮比與超小失真程度音樂格式的出現，以及網路快速的成長，帶動了數位音樂流行的風潮。從古至今的各種音樂型態，會由於樂曲編排、樂器編制、樂手風格、及地區表現手法等，產生出許多不同的音樂情緒。

市面上眾多的播放程式也設計了多樣化的歌曲分類方式，例如演唱(奏)者、專輯名稱、年份、音樂風格等分類法，用以幫助使用者整理歌曲清單，但卻鮮少看到一種針對「音樂情緒」的分類方式。因此本文則利用音樂中BASS此項樂器為基礎，利用數位訊號處理(DSP)的技術，根據BASS的音符使用，計算其音樂的和弦行進，大小調的運算，音樂風格的區別，來製作一圖像式的音樂情緒反應介面，讓使用者對該歌曲的內容有初步的了解。

關鍵字：情緒、低音聲部、和弦、音樂檢索、風格、音訊處理

# Music Emotion Classification by Genre and Chord Progression Based on Bass Line Feature Extraction

Student: Hsieh, Yun-Kai

Advisor: Dr. Stone Cheng

Master Program of Sound and Music Innovative Technologies

College of Engineering

National Chiao Tung University

## Abstract

Numerous style of music produces very different listening mood and emotion from song arrangement, musical instrument element, musician playing style, district feature etc. Many music players have designed the diverse song classification way, such as singer or performer, album name, years or style, it is music classification ways, and various metadata need to be created for each music piece. The content-based music information retrieval is under aggressive research. This study used MP3file, proposed a classification in music emotion based on the instrument “bass guitar”. The developed bass line feature extraction calculated the chord progression, music tonal, music genre, and established a real-time graphic user interface on music emotion while playing. It may help user realize the music they listening in time.

**Keywords:** Music emotion, music information retrieval, chord, music genre, Bass Line.

## 致謝

本篇論文能夠順利完成，首先要感謝我的指導教授鄭泗東老師、於我碩士生涯給予我細心的指導。在我做學問的過程中、提供了我許多寶貴的意見以及教導，適時的予我提醒，並在課餘之時分享許多做人處事上應該注意的事項與寶貴的社會經驗，不但讓我學習到做學問應有的嚴謹態度，也更加的使我在與人相處的接待上有所成長，使我能夠順利拿到碩士學位。也感謝白明憲主任的大力支持、曾毓忠與黃志方老師技術協助，推動本學程成立了全台灣第一個數位管弦樂團，並推薦學生擔任團長，使學程學生在研究上所學能於數位樂團之實地應用，推起數位音樂之新一代風潮。

我也感謝實驗室的同學俊傑、于恬，還有學弟立委、韋廷，平常大家能夠在實驗室暢談研究、課業以及音樂的話題，使我兩年的研究生涯多采多姿，在碩士過程中充滿了許多難得的回憶。

最後感謝我的父母親、家人、女友，平常給予我經濟與想法上的支持，在我低落失意的時候能夠幫助我度過各種難關，感謝有了你們，本篇論文能夠順利完成，感謝主。

# 目錄

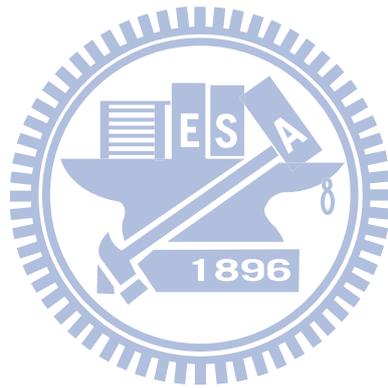
摘要.....	i
Abstract.....	ii
致謝.....	iii
目錄.....	iv
表目錄.....	vi
圖目錄.....	vii
一、緒論.....	1
1.1 研究背景簡介.....	1
1.2 相關研究.....	4
1.3 章節概要.....	6
二、研究方法.....	6
2.1 研究說明.....	7
2.1.1 心理聲學(Pschoacoustics).....	7
2.1.2 遮蔽效應(Masking Effect).....	8
2.2 研究動機.....	10
2.3 音訊內容分析.....	10
2.4 音樂特徵萃取.....	11
三、音樂特徵萃取與分類演算法.....	12
3.1 低音聲部特徵萃取.....	12
3.2 音高追蹤(Pitch Tracking).....	12
3.3 DTW 演算法.....	15
3.4 和絃研究與特徵.....	19

四、系統架構與實驗.....	26
4.1 音樂風格與情緒.....	26
4.2 演算法與系統架構.....	26
4.3 情緒分類器形式.....	35
五、結論與延伸應用.....	37
六、參考文獻.....	38



## 表目錄

表 1 頻譜特徵對應聲音特色.....	11
表 2 音樂風格與情緒之對照表.....	24



## 圖目錄

圖 1.1.1 Thayer 的情緒模型 .....	3
圖 1.2.1 Mood Cloud .....	5
圖 1.2.2 Mr.Emo .....	6
圖 2.1.2.1 Equal Loudness Curve .....	9
圖 3.2.1 LTP 與一階濾波器 .....	14
圖 3.3.1 兩向量之最短距離 .....	15
圖 3.3.2 DTW 示意圖 .....	16
圖 3.3.3 DTW-1 .....	17
圖 3.3.4 DTW-2 .....	18
圖 3.3.5DTW-3 .....	18
圖 3.4.1 由三音所構成的和弦 .....	19
圖 3.4.2 四音構成的和弦-加上 m6th 音的和弦 .....	20
圖 3.3.3 四音構成的和弦-加上 m7th 音的和弦 .....	21
圖 3.4.4 四音構成的和弦-加上 M7th 音的和弦 .....	21
圖 3.4.5 大和弦類的引申和弦 .....	22
圖 3.4.6 小和弦類的引申和弦 .....	22
圖 3.4.7 屬七和弦類的引申和弦 .....	22
圖 3.4.8 Power Chord (五度和弦)的結構 .....	23
圖 3.4.9 Add 和弦的結構 .....	23
圖 3.4.10 Slash Chord / On Chord(分割和弦) .....	24
圖 4.2.1 系統概念圖 .....	27
圖 4.2.2 系統流程圖 .....	28

圖 4.2.3 Dance Music 的 Bass 特徵.....	29
圖 4.2.4 Heavy Metal 的 Bass 特徵.....	29
圖 4.2.5 Jazz Music 常用的 Walking Bass.....	30
圖 4.2.6 Bossa Nova 的 Bass 特徵.....	30
圖 4.2.7 包含快樂情緒較大之風格系統起始值.....	31
圖 4.2.8 包含憤怒情緒較大之風格系統起始值.....	31
圖 4.2.9 包含哀傷情緒較大之風格系統起始值.....	32
圖 4.2.10 包含舒適情緒較大之風格系統起始值.....	32
圖 4.2.11 系統呈現圖.....	35



# 一、緒論

## 1.1 研究背景簡介

現今網路科技的發達，以及數位音樂的蓬勃發展，連帶的也衍生了多樣化的周邊產品，並也或多或少改變了現代人聆聽音樂的習慣。在 MP3 尚未出現以前，一般人聆聽音樂的習慣都是以一般家庭劇院，或車上音響為主，往往受限於空間地點，且在使用上只能依照 CD 燒錄的音軌依序播放。自從 MP3 出現以後，大大的改變了人們欣賞音樂的方式，從傳統的 CD 音響轉至個人電腦、筆記型電腦、甚至是數位音樂隨身聽上，在播放上也不受限於單片 CD，而是可以快速的編排自己想要的播放清單，十分自由。有鑑於此，市面上眾多的播放程式也設計了多樣化的歌曲分類方式，例如演唱(奏)者、專輯名稱、年份、音樂風格等分類法，用以幫助使用者整理歌曲清單，但卻鮮少看到一種針對「音樂情緒」的分類方式。

音樂的本質是傳達人類心底深處的原始本性，傳達人類情感的一種表達方式。可追隨自希臘哲學家柏拉圖、認為音樂具有陶冶身心靈、怡情養性的功能；其弟子亞里斯多德進一步提出曲調與節奏是性格的衍伸複製，可使人生氣或平靜，猛進或節制，亢奮或是安穩，且具有淨化情緒與昇華情感的作用[1]。亦即音樂能夠喚起強烈的情緒經驗（emotional experiences），在聆賞中構成每個人獨特的音樂過往經驗，影響其腦部（brain）、身體（body）與情感（feeling）的發展[2]。

情感一詞本身含有許多解釋，相同字有情緒、感情、心情等字詞，皆是由於受到外在刺激，進而投射出的一種外顯反映，並包含著若干生理變化。音樂所能引起的情緒反應之影響因素，可以歸納為以下四大類：(1)音樂整體架構：包含了音樂三要素：和弦、節奏、旋律，或是樂器組成、編制，時間長短等因素。聲學中之波型、震幅大小、頻率、相位等。(2)環境背景：包含聆聽場所(如歌劇院、音樂廳、音樂教室、Pub 等場所)，外在背景噪音(如自然聲、人聲、工地聲、交通噪音等)，或偶發事件等等。(3)演奏人員：

如演奏技術、外貌、個性、修養、人格特徵、服裝儀容、心情等外顯與內在要素。(4) 聆聽者：生長環境、家庭教育、音樂偏好、職業狀況、當時身心靈狀況等條件。這些因素皆是對音樂情緒的判定上有著直接的關係。

國外 Radocy 以及 Boyle[3]針對當代心理學及音樂功能之概念，提出將音樂分為激勵型（stimulative music）與鎮定型（sedative music）二類之看法，此兩類音樂，此兩種音樂也扮演著不同的角色。激勵型是只能夠激發情緒反應的音樂，具有快節奏與明顯力度變化的特色，如：(1)力度強，聲音響亮。(2)速度通常較快，能刺激人類的心跳、脈搏、血壓等生理變化。(3)切分音符與重音分明俐落；而鎮定型音樂是指能舒緩、平靜與鎮定情緒的音樂，使人感到放鬆，舒壓，音樂特徵為：(1)力度較為輕柔，旋律圓滑。(2)速度慢，蕭邦的夜曲或是搖籃曲等形式。(3)長音符與溫柔的節奏，如圓舞曲。這些音樂特徵給予後來評定音樂情緒的參考指標。

一般來說，音樂可能代表著人類各個不同的情緒反應，以之對應到各種音樂風格，例如：Pop 抒情歌曲往往聽了讓人傷感，黯然神傷；Blues & Jazz 歌曲聽來令人心情愉悅、放鬆；Rock 聽來讓人心情亢奮，熱血沸騰，有向心力；Heavy Metal 聽來令人瘋狂、焦慮不安，且麻痺神經；Classical 有著莊重、大方，使人穩定及舒壓的效果；Country 歌曲大部分為作者描寫個人悲傷的故事，所以往往富有反戰思想；Disco 適合在跳舞的場合播放，製造歡愉的氣氛；Hiphop 常常嘲諷許多社會現象，極盡嘲弄、揶揄之能事。我們可以推斷人類在欣賞各種歌曲時，可能會產生不同的情緒反應，也可以推回到各類歌曲帶有不同情緒的音樂特徵。本文根據 Thayer.R.E [4]在“*The Biopsychology of Mood and Arousal*,”這篇所提出的模型(如圖 1.1.1)，

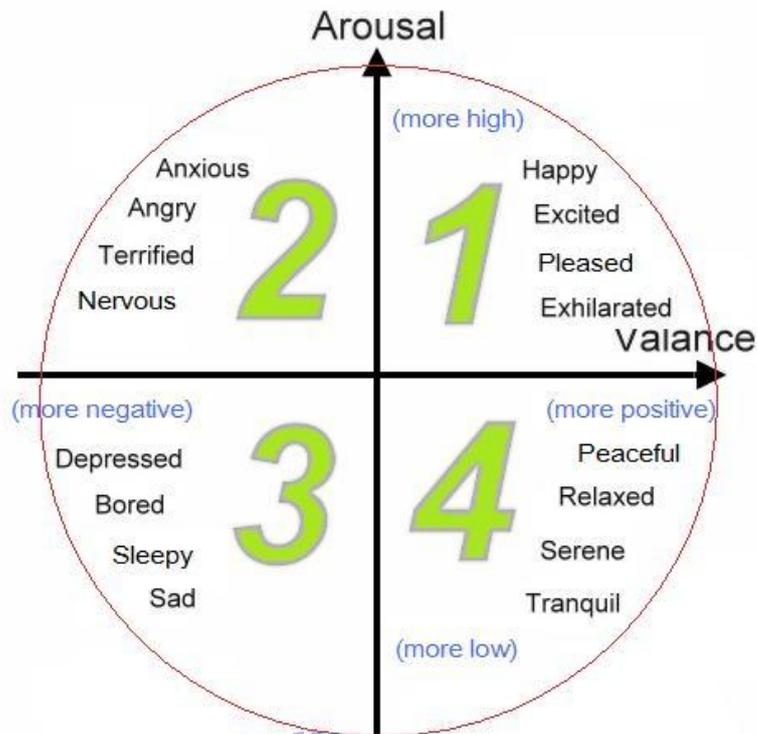


圖 1.1.1 Thayer 的情緒模型[4]

Thayer 的情緒模型主要著重在兩個特徵，壓力的大小以及心理能量的變化。由模型可看出右邊代表的是正面積極(Positive)的情緒反應、而左邊則代表負面消極(Negative)的情緒，上面代表著充滿活力、有能量(Energetic)的，下面代表的是平穩、寧靜(Silent)的。根據此圖約可分為四大類，包含「快樂輕鬆(Happy)」、「焦慮瘋狂(Anxious)」、「壓抑消沉(Depressed)」、「舒適安詳(Peaceful)」等四項。

對於評定出音樂情緒的標準而言，主要有兩種方法：一是運用文字描述表達 (verbal report)、二是利用物理測量的方式 (physiological measure)。前者為讓使用者自行選取適合的形容詞，描述聆聽音樂後的感覺，或是利用自動分類系統來實現。後者則是依據人類面對音樂而產生的生理現象反應，如腦波、血壓、心跳、汗腺等，利用儀器來加以測量及評估。

之於音樂情緒的另一項研究，為音樂風格的分類。在 MIR(音樂檢索)上的特徵萃取，除了基本如歌手、專輯等分類以外，常見尚有音樂風格之分門別類。

## 1.2 相關研究

在音樂情緒研究這塊領域，在國外 H.J Zhang [5]曾經提出將音樂分為三大特徵，分別是「強度」，「音色」及「旋律」，強度代表的是區段間能量的分布，音色是頻譜形狀的特徵，而旋律則包含了節奏的強度、節奏的規律性以及節奏本身為何。於國內已經有陳若涵[6]等二人提出可從古典音樂的調性組成，亦即分辨大小調，以及流行音樂中鼓組出現的時間來進行辨認，針對此兩種音樂皆有一定程度的辨識率。另外透過黃靜芳的研究，在聆聽賞析古典與非古典、器樂曲與聲樂曲、激勵型與鎮定型音樂時，聆賞者的情緒都有著明顯的不同。特別是激勵型音樂較易產生快樂的、激動的、幽默的、渴望的、壯闊的情緒反應；而鎮定型音樂則較易產生悲傷的、平靜的、嚴肅的、沮喪的、細緻的感受。而聽眾對於整首樂曲的熟悉程度將會影響其喜愛程度，兩者間具有顯著正相關，亦即熟悉度高則喜愛程度也高，熟悉度低則喜愛度相對較低，對樂曲的不同熟悉度與喜愛度，將會產生顯著不同的聆賞情緒反應 [7]。

於音樂風格分類研究，不得不提 G. Tzanetakis[8]的研究，他建構了一套有效的自動化音樂風格分類系統，針對音色成分、節奏型態、音高類別三項音樂特徵，設計及訓練分類器，使得對於十種音樂風格皆有較高程度的辨識效能。

在以 Bass Line 為特徵做 MIR 的領域裡，有 Y. Tsuchihashi[9]等人利用梅爾倒頻譜系數(MFCC)此技術進行辨認，總共分為 POP/Rock、Metal/Punk、Electric、Jazz/Blues、Classical、World 等五大類。K. Sumi[10]等人使用 Bass Line 做為參考輔助，進行自動化的和弦辨識[11][12][13][14]。

而在音樂情緒表示的系統介面設計上，C. Laurier and P. Herrer[15]發展的「Mood Cloud」系統，可以即時的將正在播放的音樂，分別顯示出五大情緒所呈現的比例，五

大情緒分別包含了「happy」、「sad」、「aggressive」、「relax」以及「party」，如圖 1.2.1。

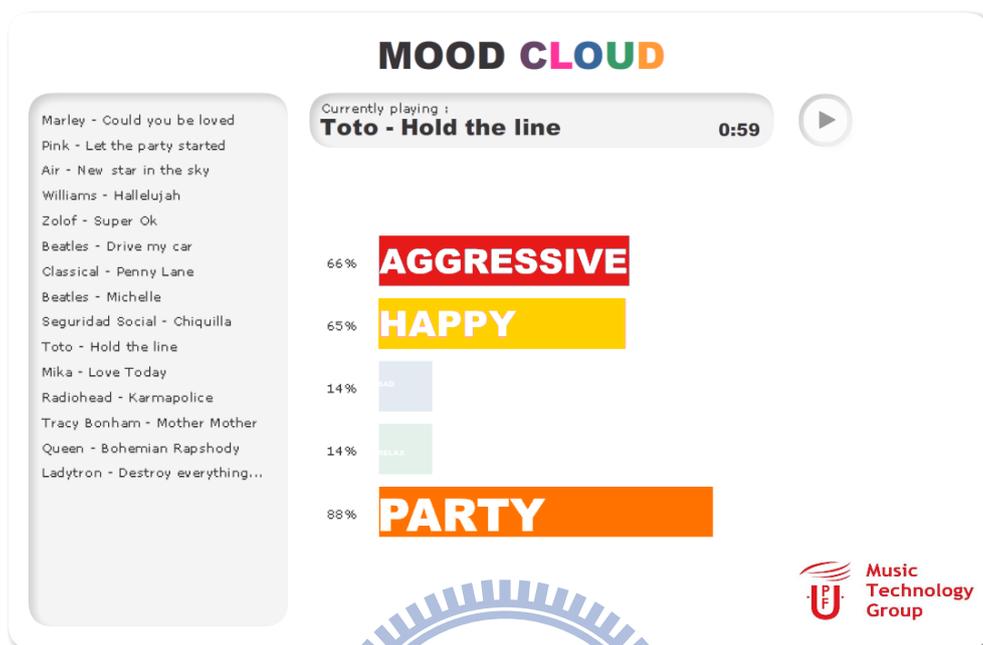


圖 1.2.1：Mood Cloud[15]

Yi-Hsuan 等人設計了一個名為「Mr.Emo」的系統[16]，此系統被設計為組織及瀏覽音樂的分佈集合狀況、並把情緒量化成各種程度的級別，再使每一個音樂樣本標示在此二維平面上。所以使用者可以簡便地找出這些樣本的中心音樂情緒。也可以以歌手為搜尋基礎、或是歌曲關鍵字搜尋，辨識出該歌手所演唱的歌曲，大致歸類在哪一情緒分布，如圖 1.2.2。

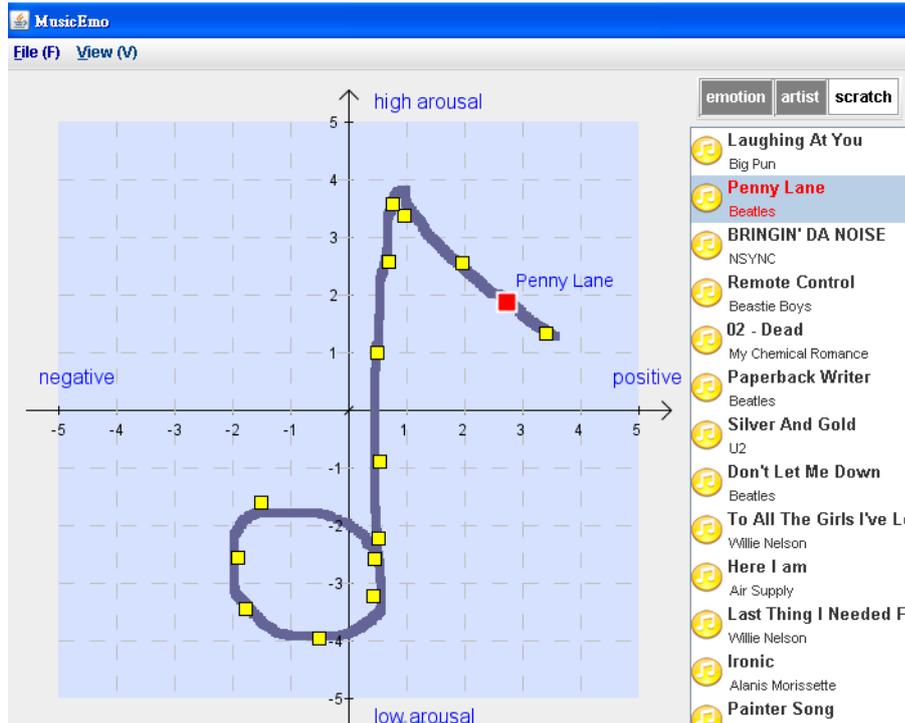


圖 1.2.2 : Mr. Emo[16]

## 1.3 章節概要

第一章為研究背景簡介與相關研究，第二章則提到了研究方法與音訊分析說明，第三章則為音樂特徵萃取與分類演算法，第四章為系統架構與實驗，第五章為結論。

## 二、研究方法

### 2.1 研究說明

本論文所針對的音樂格式為 MP3(MPEG-1 Layer3)格式，利用壓縮技術來使容量縮小，可以壓縮到原來的十分之一，而 MP3 所使用的壓縮技術、包含了心理聲學模型 (Psychoacoustics Model)，利用遮蔽效應(Masking Effect)來屏人耳所聽不到的極高頻與極低頻，達到不過多影響音質與減少檔案容量的優點。

對照前人所做的研究，大部分為採取 MIDI 做分析，但 MIDI 檔則與 MP3 檔完全不同。MIDI 檔案不過只是一堆數字的組成而已，它不是聲波檔因此沒有任何的波型及聲音，更不會有所謂的歌聲。也就是說，Midi 就是所謂的「電子音樂」。MIDI 訊息有如樂譜上的記號與音符，它記錄了發聲樂器的構成，音符長度與力度，用了什麼樂器（但是不包括樂器音色資料）、該在何時發出什麼聲音等(起始與結束時間)。所以，MID 檔的檔案大小，是與樂曲的音符個數、複雜度呈現正相關；而 MP3 檔案的大小，則只與時間長度成正相關，與紀錄的聲音內容毫無關係。因此，相對於 MIDI 而言，雖然利用 MP3 檔案來分析有較高難度，但也比較符合現今數位音樂生活化的趨勢。

另外，本文也參考了 E. Tsunoo[14]等人所提出的系統架構，分析 Bass-Line 此項音樂特徵。分析過程架構包含了：提取 BASS 特徵、Pitch Tracking、Low-Pass filtering、等技術，再藉由 DTW 演算法達成自動化分類器的訓練，區分出基本的音樂風格之後，再搭配作曲方法分析和弦的走向，用來判斷不同的情緒反應分布，詳細說明於後。

#### 2.1.1 心理聲學 (Psychoacoustics)

研究人類聽覺的心理反應稱為心理聲學 (Psychoacoustics)，就是所謂「人腦解釋聲音的方式」，心理聲學模型是對人聽覺感受的統計性質之數學表述模型，它解釋人各種聽覺感受的生理原理。而人耳的聽覺範圍即音域，大約是在 20Hz 至 20kHz。

聽覺的響度 (Loudness) 表示感受聲音的強度是根據頻率而變化的，以 phon 為單位；即在 1kHz 時的聲音強度。正常聽覺的範圍為 0-120phons。超出人耳的可聽頻率範圍的聲音，即使響度再大，人耳也不能聽到；此時響度即等於零 phon (threshold of hearing)。當響度達 120phon 時便是疼痛點 (threshold of pain)。這一組曲線稱為 Robinson-Dadson equal-loudness contours 是根據心理聲學研究人耳對純正弦音調 (sine tone) 而推論出來並經試驗證實的。

由於人耳聽覺系統複雜，人類到目前為止對它的機能和聽覺特性的某些問題總是還不能完全解釋清楚。所以，對人耳聽覺特性的研究目前僅限於在心理聲學和語言學內進行。而人耳對於不同強度和不同頻率聲音的一定聽覺範圍稱為聲域。在人耳的聲域範圍內，聲音聽覺心理的主觀感受主要有響度(Loudness)、音高(Pitch)、音色(Timbre)等特徵和掩蔽效應(Masking Effect)等特性。其中響度、音高、音色可以在主觀上用來描述具有振幅、頻率和相位三個物理的複雜聲音，故又稱為聲音“三要素”；而對於多種音源場合的人耳掩蔽效應等特性尤為重要，它是心理聲學的基礎。經常在充滿噪音環境底下工作或是常用耳機聽音樂的人對高頻的敏感度也會下降。

### 2.1.2 遮蔽效應(Masking Effect)

聲音訊號實際是一種能量波，在空氣或其他媒介中傳播，人耳對聲音能量的多少即響度或者聲壓最直接的反應就是聽到這個聲音的大小，我們稱它為響度，表示響度這種能量的單位為分貝 (dB)。即使是同樣響度的聲音，人們也會因為它們頻率不同而感覺到聲音大小不同。人耳最能明顯分辨的就是 4000Hz 的頻率，不管頻率是否增高或降低，即使是響度在相同的情況下，大家都會覺得聲音在變小。但響度降到一定程度時，人耳就聽不到了，每一個頻率都有著不同的值，如圖 2.1.2.1 所示。

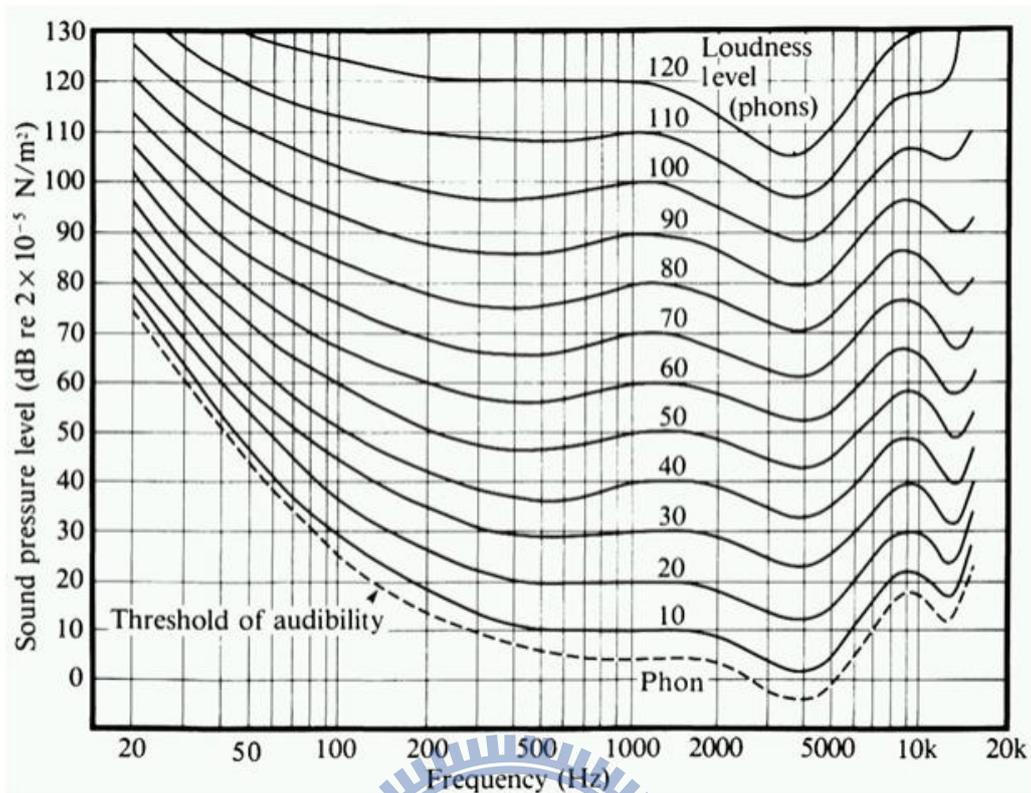


圖 2.1.2.1 Equal Loudness Curve[17]

代表的意義是說，人耳要能聽到該頻率的聲音、所需要的最小分貝值(dB)，亦即為當某個頻率的能量低於曲線中所對應出的臨界值時，就不會被人耳所察覺。一個聲音訊號如果與一種雜訊同時出現，它將變得微弱或完全聽不清楚。這就是遮蔽效應，雜訊掩蔽了訊號。遮蔽效應的大小取決於雜訊和訊號在頻率上的關係。一般來說，訊號與雜訊的頻率越接近，遮蔽也越大。這個會遮蔽其他頻率的訊號源，我們稱為遮蔽者 (Masker)，而遮蔽者對相鄰頻率的影響範圍和程度，則和遮蔽者本身是位在哪個臨界頻帶有關，不同臨界頻帶內的遮蔽者，對同一個頻帶內的其它訊號或相鄰頻帶內的訊號，會有不同的遮蔽效果，且低頻雜訊對高頻訊號的遮蔽常大於高頻雜訊對低頻訊號的遮蔽。

而遮蔽效應也存在時域上，依時間發生的前後次序，時域遮蔽效應又被細分成前遮蔽效應(Pre-masking)與後遮蔽效應(Post-masking)兩種。後遮蔽效應的意思是說，前一段時間如果聽到的是強度較大的聲音，在它之後的數百微秒(ms)內，所接收到強度較弱的

聲音，將會被遮蔽，不會被我們所聽到。而前遮蔽效應的意思是說，當人耳接收到比較微弱的聲音訊號後，緊接著接收到強度較大的聲音訊號，這時候先前所接收到比較微弱的聲音訊號，會被後來接收到強度較大的聲音訊號所遮蔽。

## 2.2 研究動機

本文嘗試以樂器中的BASS這項樂器來做為特徵，而Bass Line在音樂中扮演相當重要的角色，因為必須同時兼顧大鼓的拍點，也要跟著吉他的和絃行進，是節奏樂器與和弦樂器之間的溝通橋樑，同時也是低音聲部中不可或缺的樂器，為音樂中的靈魂。Bass Line有以下幾點特性：

1. 帶領著和絃行進。
2. 多數情況之下同時間只彈奏一個音符，相對於吉他而言，訊號分析較容易。
3. 不易察覺的低音曲線往往影響聽眾的情緒。
4. 在不同的風格中，存在著許多特定的表現手法。



根據以上特性，因此選用Bass Line做為分析的依據，歸類正確的音樂風格，分類音樂情緒。

## 2.3 音訊內容分析

一般針對聲音訊號的處理，將流程說明如下：

- 1 對聲音訊號做數位化取樣
- 2 將聲音訊號(Audio Signal)切割為許多音框(Frame)。
- 3 對音框內容做分析，包含主要三個聲音特徵:音量(Volume)，音高(Pitch)，音色(Timbre)，其他像是頻譜(Spectrum)，過零率(Zero Crossing Rate)等等。

- 4 根據以上所取得的特徵參數，可以分析音樂中所包含的資訊。
- 5 分析節奏(Tempo)，旋律(Melody)，和絃(Chord)等音樂要素。
- 6 將分類的資料使用分類器加以訓練，得到我們所需要的資料庫資料。
- 7 最後使用分類器分析音訊，建構出適合的人機介面。

## 2.4 音樂特徵萃取(Music Feature Extraction)

由音樂訊號中來判斷，可以直接從音樂中的物理現象、對應到人類的心理質量感知，如表 1 所示。但在一般文獻中提到的特徵萃取，對於特徵之定義並沒有一定的標準，根據多篇文獻歸納得到，特徵有分量值的大小。量值小的度由音框內容的分析計算得到，通常和人類感知沒有直接關係，如過零率，短時距的頻譜，倒頻譜等 [18][19][20]。大的量值通常就是人耳所能查覺到的特徵，如節奏、旋律、和聲等。

Physical Quantity	Perceptual Quality
Intensity(強度)	Loudness(響度)
Fundamental Frequency(基頻)	Pitch(音高)
Spectral Shape(頻譜形狀)	Timbre(音色)
Onset/Offset Time(起末時間)	Timing(時間)
Phase Difference(Binaural Hearing相位)	Location(聲音位置)

表 1 頻譜特徵對應聲音特色

## 三、音樂特徵萃取與分類演算法

### 3.1 低音聲部特徵萃取

通常在一首有完整編制的歌曲裡，可以略分為旋律部與節奏部。節奏部有常見的爵士鼓組(大鼓、小鼓、中鼓、落地鼓)、定音鼓、太鼓、康加鼓等。旋律部又可細分為管樂、弦樂、聲樂等部分，以聲音的頻率而言、分別控制著高、中、低頻三大部分。低頻的部分主要由低音樂器所負責，如電貝斯、低音大提琴，本文針對的特徵為低音聲部。

針對低音聲部做特徵萃取，本文設計了一個低通濾波器(Low-Pass Filter)，用來過濾音樂中的低頻訊號，除去不必要的高頻部分。低通濾波器容許低頻訊號通過，但減弱(或減少)頻率高於截止頻率的訊號的通過。對於不同濾波器而言，每個頻率的訊號的減弱程度不同。當使用在音訊應用時，它有時被稱為高頻剪切濾波器，或高音消除濾波器。

### 3.2 音高追蹤(Pitch Tracking)

所謂的Pitch，是音樂訊號裡一個很重要的特徵，一般來說，Pitch代表聲音頻率的高低，在這裡指的頻率是Fundamental Frequency，也是Fundamental Period的倒數，至於從基本頻率轉換到半音的公式如下：

$$\text{semitone} = 69 + 12 \cdot \log_2 \left( \frac{\text{Frequency}}{440} \right) \quad (3.2.1)$$

意及為，當基本頻率是440Hz時，對應到半音就是69，而每個全音階包含12個半音，每相隔一個全音階(相差八度)，頻率就是原來的兩倍，A4為440Hz，相隔八度的A5為880Hz，以此類推。

音高追蹤的基本流程如下：

- 1.將整段音訊訊號切成很多段音框 (Frames)，相鄰音框之間可以重疊。

- 2.算出每個音框所分別對應的音高。
- 3.剔除不穩定的音高值。
- 4.對整段音高進行平滑化處理。

本文使用的方法為長期模組預測(Long-Term Prediction)，是用來移除訊號中長期的重複特性，主要目的為獲得編碼增益。方法是在時間軸上做預測，而在頻譜上取變化值。觀察訊號變化為緩和或穩定，或是樂器部分具有明顯可辨識音高性質者，其前後時間之波形相當類似、時間頻率關連性很大，可利用預測的方法，移除這樣的重複性，而獲得編碼上的增益，提升效能，LTP 便是為了這樣的目的而產生。其在時間軸上計算預測值，兩者皆轉至頻譜上後再做數值相減。

一般的做法是從訊號中移除音高區間的重複、在M個延遲區段的取樣之後、利用短時距的FIR濾波器預測，定義M為音高之衰退，所以LTP預測失誤或殘差可定義如下：

$$d(n) = x(n) - \sum_{k=0}^q b_k x(n-M-k) \dots\dots\dots(3.2.2)$$

階數q+1是常態分佈於區間{1, 2, 3}，考慮單一係數濾波器的狀況，可簡化為：

$$d(n) = x(n) - b_0 \cdot x(n-M) \dots\dots\dots(3.2.3)$$

為了縮小d(n)在長度為N的區塊裡，推導過程中將b<sub>0</sub>設為0，可以得到最佳的濾波係數：

$$b_0 = \frac{\tilde{r}_{xx}(M)}{r_{zz}(M)} \dots\dots\dots(3.2.4)$$

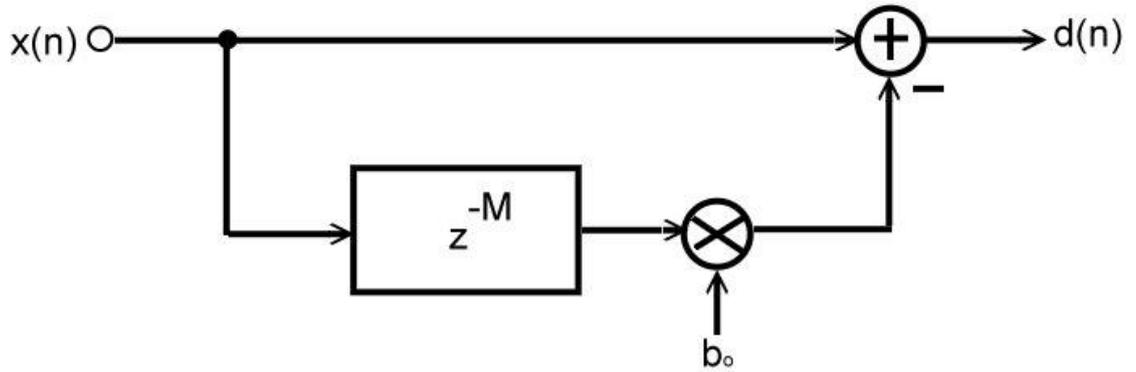


圖3.2.1 LTP與一階濾波器

而  $\tilde{r}_{xx}(m)$  可以定義為：

$$r_{xx0}(m) = \sum_{n=0}^{N-1} x^2(n-m) \dots\dots\dots(3.2.5)$$

當區間能量延遲達到  $m$  個取樣，設定結果帶回(3.2.3)式，可以得到以下：

$$E_d = \sum_{n=0}^{N-1} x^2(n) - r_{xx, norm}(M) \dots\dots\dots(3.2.6)$$

相關係數  $M$  為：

$$r_{xx, norm}(m) = \frac{r_{xx}^2(m)}{r_{xx0}(m)} \dots\dots\dots(3.2.7)$$

當  $r_{xx, norm}(m)$  從最大到最小化的剩餘能量，則  $r_{xx, norm}(m)$  的延遲為  $m=M$ 。考慮長度為  $N$  的區段，則(3.2.7)式呈現自相關係數的平方來計算、在分母的區間延遲能量為  $M$  個取樣的情況下。式子  $r_{xx, norm}(m)$  表示一種正規化自相關係數只包含正值，如果使用來偵測音高的週期， $r_{xx, norm}(m)$  不需要將  $m=M$  做全域最大化，但可以預期在這種情況下，可能會是區域最大的。

為了找到  $M$  音高延遲的可能者，在  $r_{xx, norm}(m)$  的第一個區域最大值做搜尋，第二步是

從這些最大值中找出自相關係數  $r_{xx}(m)$  為正值者。函式  $r_{xx, norm}(m)$  也有可能出現最大值在  $r_{xx}(m)$  為最小的時候。第三步將考慮  $b_0 m$  的值， $b_0$  的係數值可能像是語音或是在噪音訊號中的 0 值。因此  $b_0$  可以作為推估音高延遲的一種質量偵測，接著結合線性預測演算法，可以找到分析訊號中的基頻值，最後對應到標準音高。

### 3.3 Dynamic Time Warping (DTW) 簡介

DTW 是 Dynamic Time Warping 的簡稱，中文可以翻譯成「動態時間扭曲」或是「動態時間校正」，這是一套根基於「動態規劃」（Dynamic Programming，簡稱 DP）的方法，可以有效地將搜尋比對的時間大幅降低。

DTW 的目標就是要找出兩個向量之間的最短距離。一般而言，對於兩個  $n$  維空間中的向量  $\mathbf{x}$  和  $\mathbf{y}$ ，它們之間的距離可以定義為兩點之間的直線距離，稱為歐基里得距離（Euclidean Distance），如圖 3.3.1

$$\text{dist}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|^2 \tag{3.3.1}$$

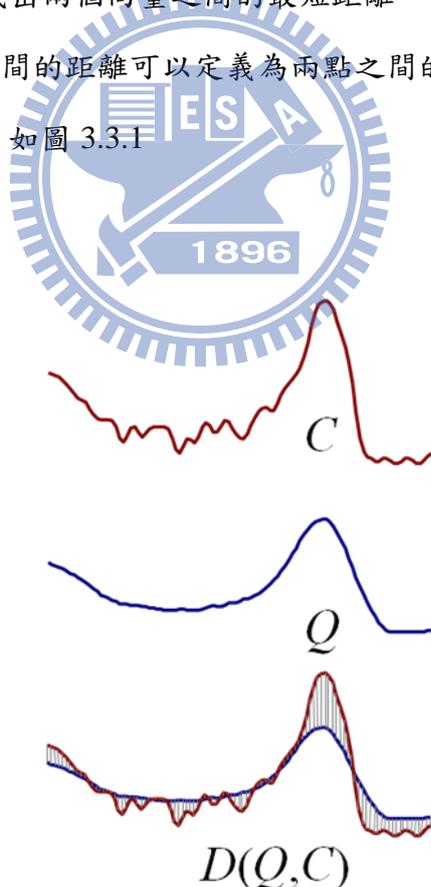


圖 3.3.1 兩向量之最短距離

$$D(Q,C) \equiv \sqrt{\sum_{i=1}^n (q_i - c_i)^2} \quad (3.3.1)$$

但是如果向量的長度不同，那它們之間的距離，就無法使用上述的數學式來計算。一般而言，假設這兩個向量的元素位置都是代表時間，由於我們必須容忍在時間軸的偏差，因此我們並不知道兩個向量的元素對應關係，因此我們必須靠著一套有效的運算方法，才可以找到最佳的對應關係。

- 假設有兩個向量  $t$  和  $r$ ，長度分別是  $m$  和  $n$ ，那麼 DTW 的目標，就是要找到一組路徑  $\{(p_1, q_1), (p_2, q_2), \dots, (p_k, q_k)\}$ ，使得經由上述路徑的「點對點」對應距離之和  $|t(p_i) - r(q_i)|$  為最小( $i=1 \sim k$ )，而且，此路徑必須滿足下列條件：
- 端點關係： $(p_1, q_1) = (1, 1)$ ,  $(p_k, q_k) = (m, n)$ 。此端點關係代表這是「頭對頭、尾對尾」的比對。
- 局部關係：假設最佳路徑上任一點可以表示成  $(i, j)$ ，那麼其前一點路徑只有三種可能： $(i-1, j)$ ,  $(i, j-1)$ ,  $(i-1, j-1)$ 。此局部關係定義了路徑的連續性，而且也規定了  $t$  的任一個元素至少對應一個  $r$  的元素，反之亦然。

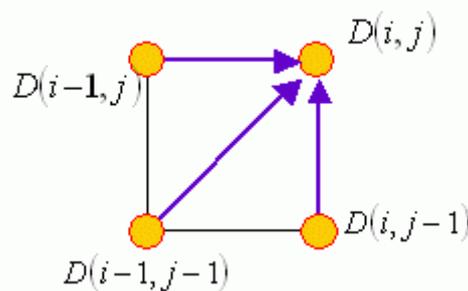


圖 3.3.2 DTW 示意圖

- DTW 基本四步驟：
  1. 目標函數：定義  $D(i, j)$  是  $t(1:i)$  和  $r(1:j)$  之間的 DTW 距離，對應的最佳路徑是由  $(1,1)$  走到  $(i,j)$ 。

2. 遞迴關係： $D(i,j) = |t(i) - r(j)| + \min\{D(i-1, j), D(i-1, j-1), D(i, j-1)\}$
3. 端點條件： $D(1,1) = |t(1) - r(1)|$
4. 最後答案： $D(m,n)$

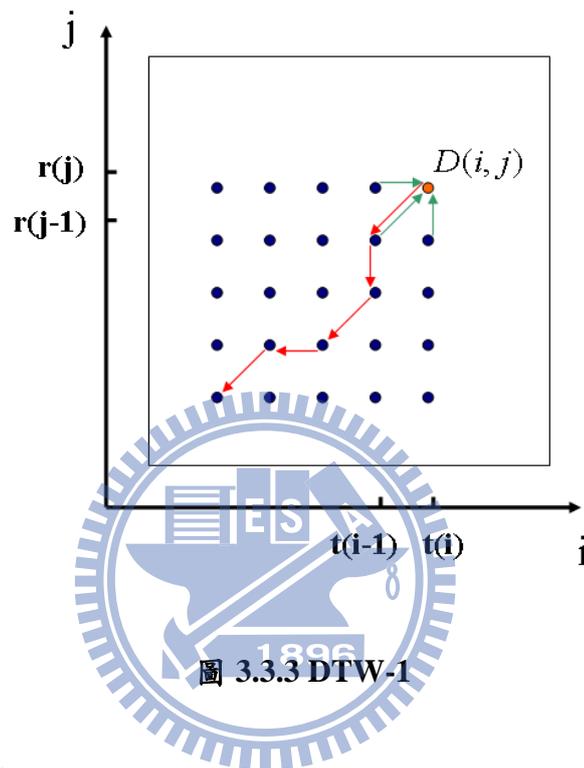


圖 3.3.3 DTW-1

另外包含以下幾種型態：

型態二：

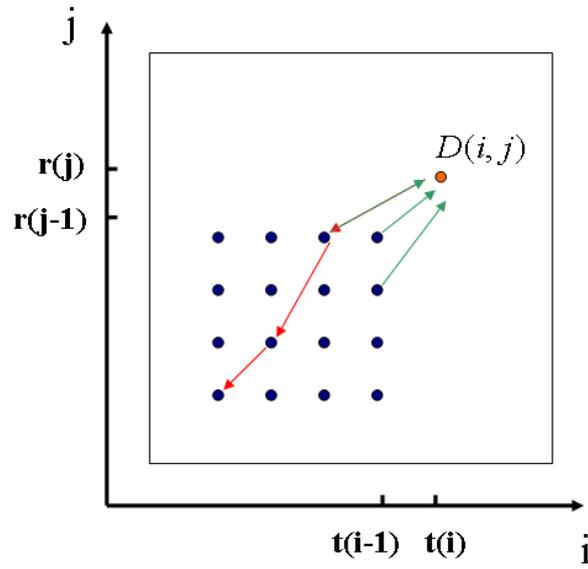


圖 3.3.4 DTW-2

$$D(i, j) = \|t(i) - r(j)\| + \min \begin{cases} D(i-1, j-2) \\ D(i-1, j-1) \\ D(i-2, j-1) \end{cases}$$

(3.3.2)

型態三：

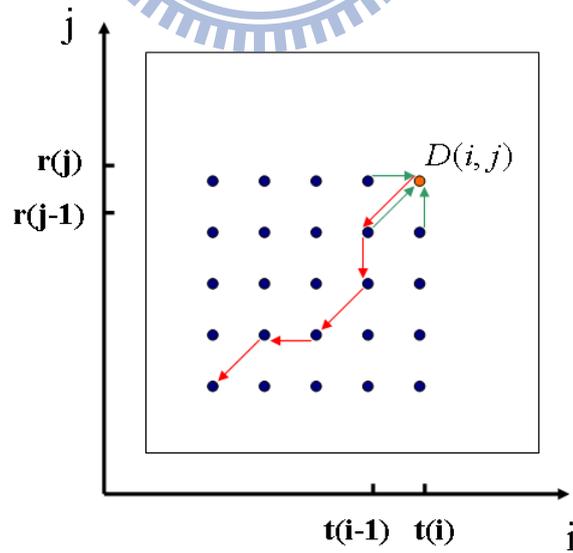


圖 3.3.5 DTW-3

$$D(i, j) = |t(i) - r(j)| + \min \begin{cases} D(i, j-1) \\ D(i-1, j-1) \\ D(i-1, j) \end{cases} \quad (3.3.3)$$

以上參考[21][22][23]等資料來源。

### 3.4 和絃研究與特徵

關於和絃的結構及種類，浦田泰宏在「和絃進行活用與演奏秘笈」中將代表性和絃的種類及特徵作了詳細的說明。本系統則對和絃分析說明於本章節。

和絃的定義就是「同時發出兩個音以上合成起來的音」。代表性和絃的種類及特徵部分共分為三音構成的和絃、四音構成的和絃、引伸和絃及其他和絃等4大部分。詳細說明如下，參考至[24]：

#### 一、三音構成的和絃部分：

這是各種和絃中最基本的和絃形式，就是三度一個音符，總共有三個音構成一個和絃。這種類型的和絃代表有大和絃(Major chord)、小和絃(Minor chord)、掛留和絃(Suspended chord)、增和絃(Augment chord)等4種(如圖)。大三和絃的音色特點是明亮且安定，是最常被廣泛應用的和絃；小三和絃的音色特點與大三和絃相較起來，音色特徵較為暗淡；掛留和絃其音色特徵比較空靈；增三和絃其音色特徵是帶有不安定的感覺，在使用上較有所限制，如圖3.4.1所示。



圖3.4.1 由三音所構成的和絃

## 二、四音構成的和弦部分：

此處主要介紹在三音構成的和弦上，加上「M6th(六度音)、m7th(小七度)、M7th(大七度)」這三個音符而成。

### 1. 加上M6th(六度音)的和弦：

加上M6th音的和弦，其和弦名稱就是在上面多寫一個「6」的數字。在音色的特徵上，會比原本的大和弦更暗更柔；而以小和弦為基本的「小六和弦」，聽起來會強調出小和弦的陰柔感；以減和弦(Diminish chord)加上M6th(六度音)，在音色的特質上，聽起來會有種不安感的獨特效果，如3.4.2所示。

(a) 6和弦

(b) 小6和弦

(c) 減和弦

圖3.4.2 四音構成的和弦-加上m6th 音的和弦

### 2. 加上m7th(小七度音)的和弦部分：

加上m7th音的和弦，其和弦名稱就是在上面多寫一個「7」的數字。七和弦(Dominant 7<sup>th</sup> chord)在音色的特徵上，比原本的大和弦聽起來更不安定，除了具有終止式的功能外，經常被運用在藍調類的和弦進行中；屬七掛留和弦(Dominant 7<sup>th</sup> sus4 chord) 其音色特徵比較空，常被運用在和以同一音為根音的七和弦樂句之句尾；小七和弦(Minor 7<sup>th</sup> chord) 會將小調樂句較暗的音色特質削弱，聽起來會有比較流行感，和小和弦一樣經常被使用。小七減五和弦(Minor 7<sup>th</sup> Flat 5 chord) 其音色特徵為不安感，如圖3.4.3所示。



圖3.4.3 四音構成的和弦-加上m7th 音的和弦

3.關於加上M7th(大七度音)的和弦部分：

加上M7th音的和弦，其和弦名稱就是上面多寫一個「M7」的記號。大七和弦(Major 7th chord)比原本大三和弦聽起來，更帶有洗鍊過的透明感。小三大七和弦(Minor Major 7th chord)聽起來更有強調出小和弦特質的感覺如圖3.4.4所示。



圖3.4.4 四音構成的和弦-加上M7th 音的和弦

三、引申和弦(Tension chord)部分：

在四音構成的和弦中，加上一個與根音相離一個八度上的和弦外音，這個和弦稱為「引申和弦」，其音色特徵聽來有種獨特的張力。可添加在「引申和弦」上的音符，要視其原和弦的種類而定。引申和弦可分為大和弦類的引申和弦、小和弦類的引申和弦及屬七和弦類的引申和弦等3類。說明如下：

1. 大和弦類的引申和弦部分：

在大和弦上加上 M6th 或M7th 音符的和弦，都屬於大和弦類，在這一類和弦上，可加「9th、#11th、13th」音符，就變成引申和弦，因添加了引申音，讓整個音色聽來更有深度的層次感如圖3.4.5所示。



圖3.4.5 大和弦類的引申和弦

2. 小和弦類的引申和弦部分：

以小和弦為基礎，再加上「9th、11th」的音作為引申音（11th 只加在小七和弦上），音色比原本的和弦更添複雜性，如圖3.4.6所示。



圖3.4.6 小和弦類的引申和弦

3. 屬七和弦類的引申和弦部分：

雖然屬七和弦可以添加更多豐富的引申音，但實際上常被運用的引申音以 9th 及以 9th 音符作上下半音的變化(#9th、b9th)延伸（altered tension）為主，如圖3.4.7所示。



圖3.4.7 屬七和弦類的引申和弦

#### 四、其他和弦部分：

以常用的 Power Chord(五度和弦)、Add 和弦(添加和弦)及Slash Chord / On Chord(分割和弦)等三部分作為說明，並注意其結構和音色特徵。

##### 1. Power Chord (五度和弦)部分：

Power Chord 僅以「根音、完全五度音(P5th)」兩個音符所構成的和弦，即被稱為「Power chord」或「五度和弦」。因為這種和弦內，沒有可決定是大和弦或小和弦的第三音，除了根據曲子的調性來找出和弦名稱之外，在記譜上也以「根音音名+omit3」(省略3)的形式作記錄，其音色特徵較為中性感如圖3.4.8所示。

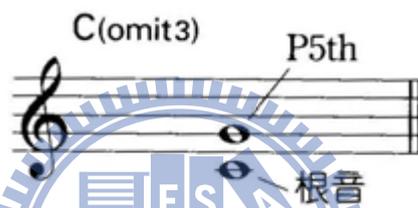


圖3.4.8 Power Chord (五度和弦)的結構

##### 2. Add 和弦(添加和弦)部分：

Add 和弦(添加和弦)就是直接在三音構成的和弦上添加音符，就是所謂的「Add 和弦」(添加和弦)，記載表示方式即是「add+添加的音符度數」。以這9th 音符和弦為代表，其音色為更有深度的層次感及複雜性。其解說譜例如下如圖3.4.9所示：



圖3.4.9 Add 和弦的結構

##### 3. Slash Chord / On Chord(分割和弦)部分：

Slash Chord / On Chord 就是將根音以外的音指定為最低音(Bass 音)的和弦，共分為兩種，一種是像轉位和弦一樣，將根音之外的和弦音當作是最低音；另一種是將和弦外音的音符指定為最低音。Slash Chord / On Chord 常用於通俗音樂bass line 之行進中，使曲子的和弦進行更加順暢無阻如圖3.4.10所示。

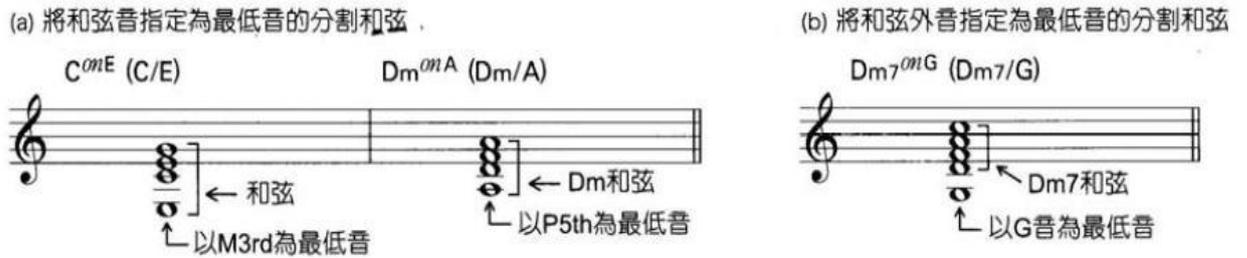


圖3.4.10 Slash Chord / On Chord(分割和弦)

而所謂的和弦級數是相對於調式而言的。簡單的說，每一個大調都有一個 Do Re Mi Fa Sol La Si 的音階。換成簡譜的話，就是每一個大調都可以被記成：1 2 3 4 5 6 7

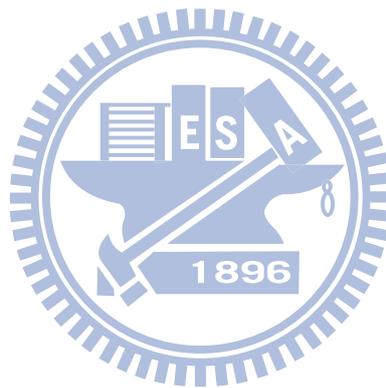
一個基本的三和弦（三個音組成的和弦）是由根音、加上三音與五音構成的。一般的和弦記法都是記根音，比如 Cm 就是以 C 這個音為根音的小和弦。那麼如果記成 1m，這個意思就是以一組音階中的 Do 這個音為根音所組成的小和弦。好比，在 A 大調中 1m 就是 Am 和弦，而 5m 就是 Em 和弦。

此外，這裡所講的這些和弦級數，所代表的和弦都是各調的「自然和弦」。所謂「自然和弦」就是該和弦的組成音都是大調音階的「音階音」。換句話說，以 C 大調的自然和弦為例，所有 C 大調的自然和弦的組成音都是鋼琴上的白鍵，沒有任何黑鍵。所以 C 大調的一到七級和弦。分別是：

C Dm Em F G Am B 減。

這七個和弦就是 C 大調的自然和弦。構成這七個和弦的音，都是 C 大調的大調音階中的音階音。也就是所有鋼琴的白鍵的音。[25]

由於本文針對Bass Line做和弦辨認，辨認方法為針對該和弦的根音為主。Bass於歌曲中在大多數情形下往往只彈奏和弦的根音，因此我們只辨認出該首歌曲的和弦級數，由和弦級數的行進作為情緒的變化指標。



## 四、系統架構與實驗

### 4.1 音樂風格與情緒

所謂的音樂風格，指的是把所有音樂做一個分門別類的方法。從古到今，根據各地區民族特性、社會經驗、樂器組成，進而衍生各式各樣的音樂風格，界定了不同的音樂流派。細數世上的音樂形式種類繁多，無法一一列出，本系統參照了常見的幾種音樂風格，對應到四種情緒類別。舉例來說：Dance, Latin, Rock等風格，常常可以炒熱現場氣氛，讓聽眾隨音樂起舞，因此歸類於Happy/Excited第一種情緒;Heavy Metal, Punk, Funk比較容易使人激動，挑起人們亢奮的情緒，因此歸類於Angry/Anxious第二類情緒;而Soul, Jazz, Blues等風格，比較傾向於較悲傷，令人黯然神傷，因此歸類於Depressed/Sad 第三種情緒;Bossa Nova, Country等，讓人放鬆，心情舒適，因此歸類於 Peaceful/Relaxed 第四種情緒，如表二所示。



<i>Music Genres</i>	<i>Emotion</i>
Dance/Latin/Rock	Happy/Excited/Pleased/Joyful
Heavy Metal/Punk/Funk	Angry/Anxious/Nervous
Soul/Jazz/Blues	Depressed/Sad/Bored/Sleepy
Bossa Nova/Country	Peaceful/Relaxed/Serene

表2 音樂風格與情緒之對照表

### 4.2 演算法與系統架構

本文的系統架構描述如下：當音樂訊號進入系統後，萃取音樂中的 BASS 訊號，並使用音高追蹤(Pitch Tracking)的方法分析音高及音長，接著利用 DTW(Dynamic Time Warping)演算法將音樂中的低音聲部、也就是辨識出的音符、與系統中提供的風格樣版

來比對，找出最相近的音樂風格，對應到最為相近的音樂風格之後，給予該風格相匹配的音樂情緒，如表 2 所匹配之相對應情緒。而系統介面是由四個圓形所組成，圖 4.2.1 為系統概念圖，並且根據該音樂之風格、設定該情緒較大的起始值及較大的加權係數，相反的情緒則設定較小的啟始值及較小的加權係數，之後再分析的和絃行進，藉由和弦的走向給予不同情緒述職的改變，即時的改變介面的圓圈半徑大小，系統流程圖如圖 4.2.2 所示：



圖4.2.1 系統概念圖

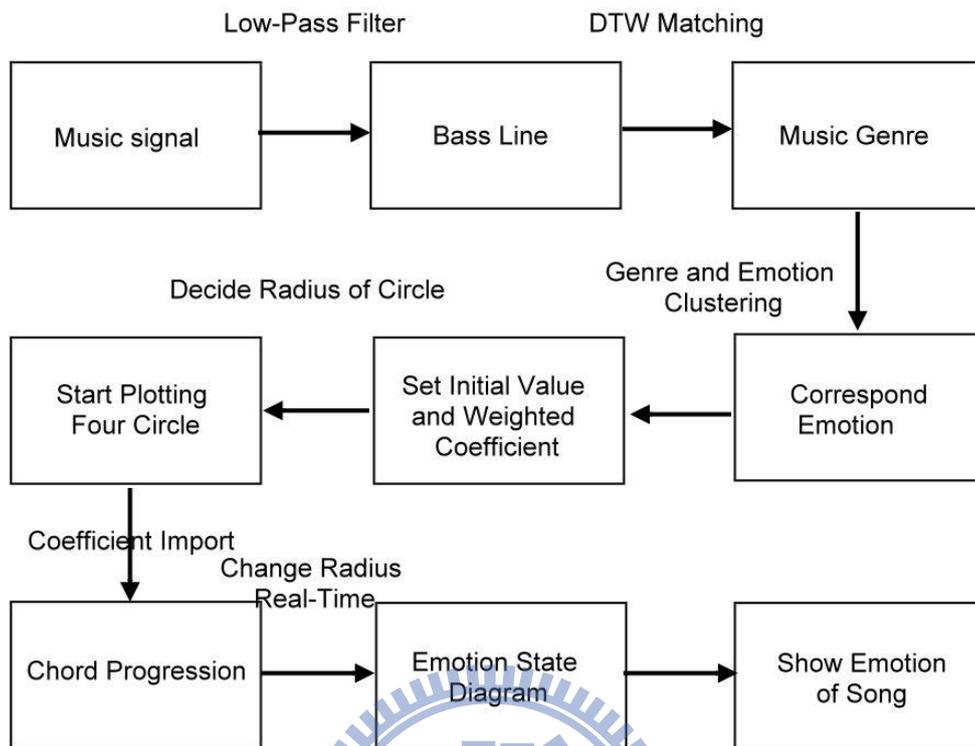


圖4.2.2 系統流程圖

本系統使用Matlab R2008a作為系統開發軟體，開發主機CPU為Intel Centrino2 Duo P8400，記憶體為4G，Windows Vista 32位元。

於測試的音樂檔，在此選用一段30秒的音樂片段來做為測試，依序通過低通濾波器，音高追蹤，比對相似音樂風格，細部方法如第三章所提。以下說明Bass-Line於各音樂風格所出現的特徵。

- **Dance Music**：在 Bass 彈奏中，時常使用根音與八度音交替使用的八分音符行進，如圖 4.2.3 所示，頻率呈現一比二的整數比，音符區間為八分音符。

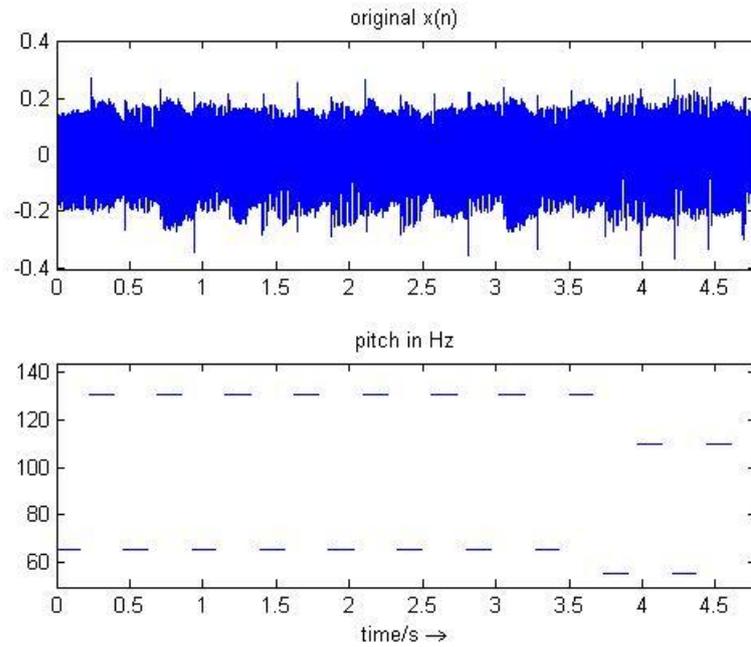


圖4.2.3 Dance Music的Bass特徵

- Heavy Metal :Bass 常常使用的是密集連續的十六分音符根音彈奏，如圖 4.2.4

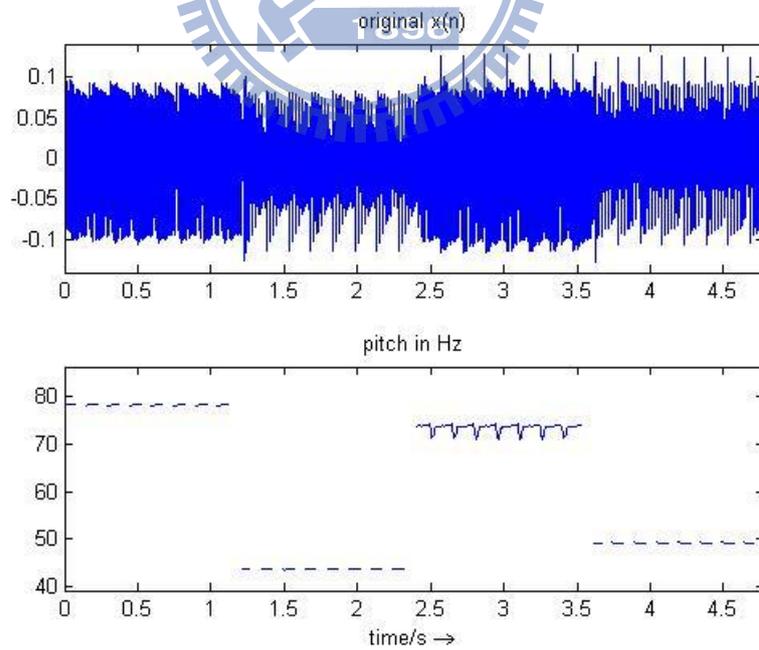


圖4.2.4 Heavy Metal的Bass特徵

- **Jazz:** Bass 通常是由許多等長的四分音符所組成，一般稱做 Walking Bass，由和絃內音及經過音所組成，如圖 4.2.5 所示

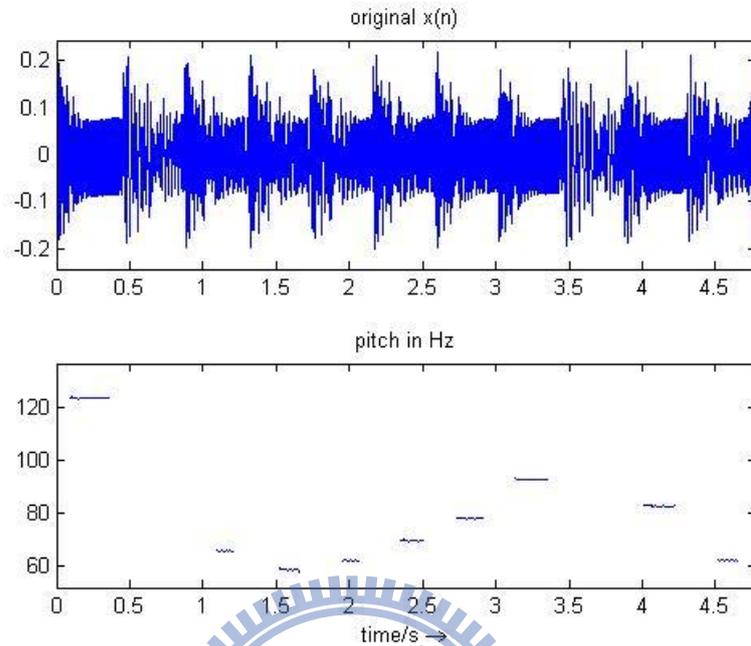


圖4.2.5 Jazz Music常用的Walking Bass

- **Bossa Nova:** Bass 的彈奏通常為根音與上五度音或下五度音交替使用，如圖 4.2.6

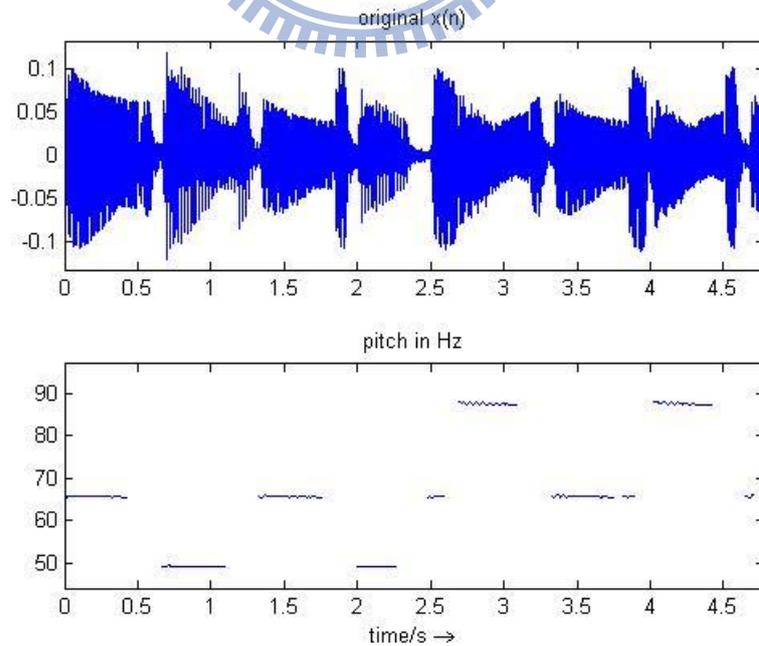


圖4.2.6 Bossa Nova的Bass特徵

由以上分類圖可歸納得知，Bass-Line在各風格常見的伴奏手法。當歸類為該音樂風

格之後，給予給風格對應之音樂情緒較大分起始半徑20，對角線之相反情緒半徑給予8，其餘兩者半徑則為10。

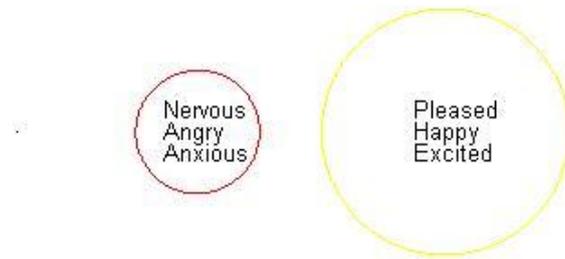


圖4.2.7 包含快樂情緒較大之風格系統起始值



圖4.2.8 包含憤怒情緒較大之風格系統起始值

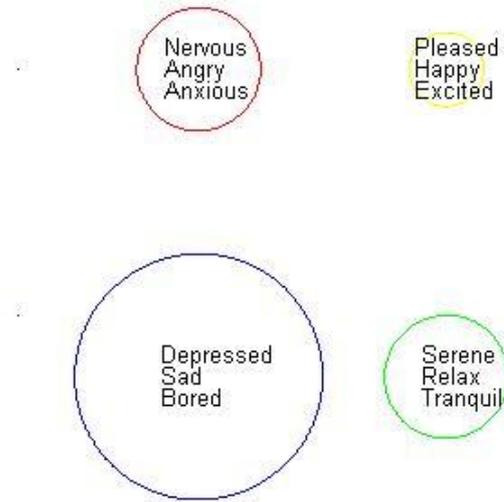


圖4.2.9 包含哀傷情緒較大之風格系統起始值

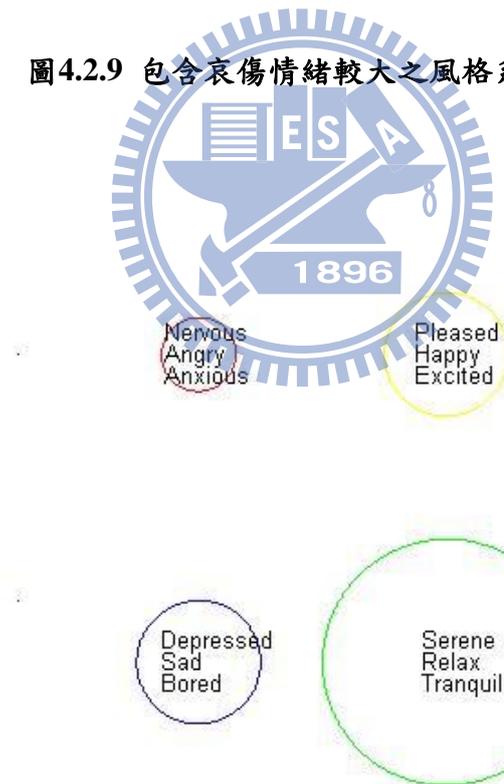


圖4.2.10 包含舒適情緒較大之風格系統起始值

本系統以四個圓形來呈現，圖中的四個圓形會隨著音樂的播放，時間的行進，改變

半徑、放大或縮小，圓形的大小也代表著情緒能量的反應強烈，能量越大，圖型越大，反之則越小。

接著會分析和絃的行進，給予適當的加權值，和絃分析參考至allmusic.com網站上提供的”Moods”分類[25]，分以下幾種情況討論，所列的羅馬數字為和絃之級數，數字的增減代表該情緒圓圈半徑的改變：

### Happy

Case 1: I → IV → V → I

Case 2: I → VI → II → V

Case 3: I → V → I → V

以上係數給予: Happy+0.5, Angry+0.1, Relax+0.3, Sad-0.2

### Angry

Case 4: 重複同樣和絃四小節以上

以上係數給予: Happy+0.2, Angry+0.5, Relax-0.3, Sad-0.1

Case 5: I → V → II → II

Case 6: I → V → 降VII → IV

Case 7: I → V → VI → IV

Case 8: VI → V → VI → V

以上係數給予: Happy+0.2, Angry+0.3, Relax-0.1, Sad+0.1

### Sad

Case 9: 順階和絃下行 I → VII → VI → V or VI → V → IV → V

Case 10: 三和絃的和絃內音運行

以上係數給予: Happy-0.1, Angry+0, Relax+0.2, Sad+0.3

Case 11: I → V → II → VI

Case 12: I → I → IV → V

Case 13: 藍調十二小節 I 7(\*4) → IV7(\*2) → I 7(\*2) → V 7 → IV7 → I 7 → V 7

以上係數給予: Happy-0.1, Angry+0, Relax+0.1, Sad+0.3

### Relaxed

Case 14: 順階和絃下行 I → VII → VI → V → IV → IV → V → V

以上係數給予: Happy+0.2, Angry-0.2, Relax+0.5, Sad+0.1

Case 15: I → IV → V → I

以上係數給予: Happy+0.2, Angry-0.2, Relax+0.3, Sad-0.1

### Others

以上係數給予: Happy+0.1, Angry+0.1, Relax+0, Sad-0.1



根據以上的幾種組合，每隔若干時間取一次新的半徑，及時的改變圖形大小，使用者可以觀看音樂情緒的及時分布狀態圖，最後輸出為 avi 影片檔，配合音樂來播放，如圖 4.2.11。

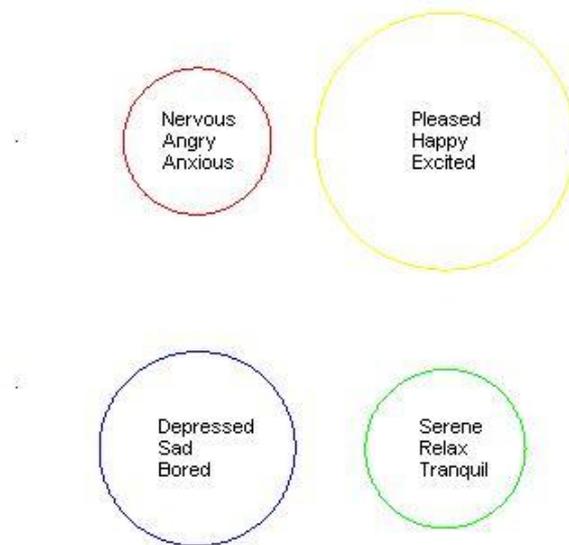


圖4.2.11 系統呈現圖

### 4.3 情緒分類器形式

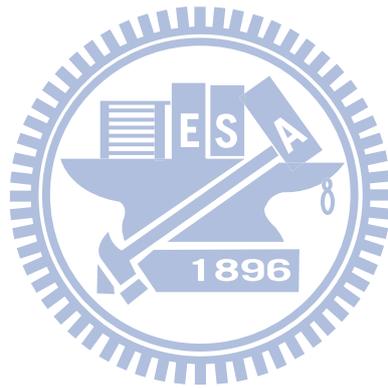
Judith C. Brown[26]利用 DTW(dynamic time warping)的技術，以及基頻追蹤(Fundamental frequency tracking)，並配合殺人鯨的聲音語料庫，加以運算之後，來辨認殺人鯨的情緒分類狀況。

針對以上的技術發展，可以應用在其他各種層面。例如：轉換到其他各類動物及昆蟲上，建置個別的聲音語料庫，以便讓人類更加了解動物的情緒，協助動物照顧人員與研究專家認知動物的需求，拉近人類與動物之間的距離。

現今除了傳統的醫學治療之外，又有著音樂治療此種新型態的治療方式。某些原始民族的巫醫會在宗教儀式當中使用音樂吟唱、或是咒語來治療，姑且不討論其有無科學依據，但可以見得音樂能對人類的心理造成一定程度的療效。根據腦神經學者研究發現，人類的大腦裡原本就存在著音樂元素，許多臨床研究證明了音樂用來成功治療的案例。在十九世紀初期，許多病人對種種刺激都沒有任何反應，但唯獨對音樂有感受力。

因此、搭配音樂情緒的研究，再配合音樂治療的方法，可以利用不同風格的音樂、產生不同類型的情緒，針對不同的病症實施相對的療程，將是未來新一波的治療方式。

傳統上中醫師除了「望診、聞診、問診、切診」另外也可配合中醫脈搏偵測，利用極細微的聲音接收裝置，搭配上上述資料庫的應用，可以基本的建置出寸、關、尺，各部再有浮、中、沉，所以 3\*3 等所謂 3 部 9 候，共約 28 種的脈象，進一步的幫助中醫師判斷病情。



## 五、結論與延伸應用

針對以上的技術發展，可以應用在其他各種層面。例如：轉換到其他各類動物及昆蟲上，建置個別的聲音語料庫，以便讓人類更加了解動物的情緒，協助動物照顧人員與研究專家認知動物的需求，拉近人類與動物之間的距離。

而美國研究發現，聽自己喜歡聽的歌可以擴張血管，對心臟有相當助益。本研究也能針對使用者的習性，提供適合的音樂，使之心情愉悅，積極進取。

音樂所隱含的情緒包羅萬象，往往一首歌曲之中夾雜了許多情緒的成分，目前很難以用四分法去精準的區分出一首歌真正的情緒。在加上聽者會隨著人、事、時、地、物等等各因素層面的影響，導致各聽眾對同一首歌的情緒判別不一致，使得出來的結果客觀性十分強烈。

本研究提供一項做為分析音樂情緒的參考方法，提出以Bass辨認和絃行進，並參照幾種常用的和絃進行，給出數值變化，建構一分析音樂情緒的系統介面。只是使用上仍有些許限制，例如只能針對以Bass Line為伴奏的音樂，無法針對以Bass為主奏樂器分析，另外低音聲部的特徵萃取、系統中的樣板可以改用HMM(Hidden Markov Model)或GMM(Gaussian Mixture Model)等方式訓練，及和絃的辨識率也有改善的空間。希望將來能加強辨識率，提高分析結果，廣泛應用在其他層面，如音樂治療等。

## 六、參考文獻

- [1]方銘健 (1997) 。藝術、音樂情感與意義。台北：全音。
- [2]Reimer, B. (2005). “New brain research on emotion and feeling: Dramatic implications for music education.” *The International Journal of Arts Education*, 3(1), 8-36.
- [3]Radocy, R. E., & Boyle, J. D. (2003). “Psychological foundations of musical behavior(4th ed.)”. Springfield, IL: Charles C. Thomas.
- [4]Thayer, R. E., “The Biopsychology of Mood and Arousal”, Oxford University Press, (1989).
- [5]Lie Lu, Member, IEEE, Dan Liu, and Hong-Jiang Zhang, Fellow, IEEE, “Automatic Mood Detection and Tracking of Music Audio Signals”, *IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING*, VOL. 14, NO. 1, JANUARY (2006).
- [6]Juo-Han Chen, “Content Base Music Emotion Analysis and Recognition”, National Tsing Hua University,(2006).
- [7]黃靜芳, 吳舜文, ” 大學生音樂選曲與情緒反應之相關研究”, 國際藝術教育學刊,(2007)
- [8]Tzanetakis, G. Cook, P. “Musical genre classification of audio signals” , *IEEE Transactions on Speech and Audio Processing*, Jul (2002).
- [9]Yusuke Tsuchihashi,Tetsuro Kitahara, Haruhiro Katayose. “USING BASS-LINE FEATURES FOR CONTENT-BASED MIR”, Kwansai Gakuin University,Japan,CrestMuse Project,CREST,JST,Japan, ISMIR (2008) – Session 5b – Feature Representation
- [10] Kouhei Sumi,Katsutoshi Itoyama,Kazuyoshi Yoshii,Kazunori Komatani,Tetsuya Ogata,and Hiroshi G. Okuno, “AUTOMATIC CHORD RECOGNITION BASED ON PROBABILISTIC INTEGRATION OF CHORD TRANSITION AND BASS PITCH

ESTIMATION”, Dept. of Intelligence Science and Technology Graduate School of Informatics, Kyoto University, National Institute of Advanced Industrial Science and Technology (AIST), ISMIR (2008) – Session 1a – Harmony

- [11] Masataka Goto. “A real-time music-scene-description system: predominant-F0 estimation for detecting melody and bass lines in real-world audio signals”, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan Received 22 May 2002; received in revised form 9 May 2003; accepted 13 March 2004
- [12] Jakob Abeßer, Hanna Lukashevich, Christian Dittmar, Gerald Schuller, “GENRE CLASSIFICATION USING BASS-RELATED HIGH-LEVEL FEATURES AND PLAYING STYLES”, Jakob Abeßer, Hanna Lukashevich, Christian Dittmar, Gerald Schuller, 10th International Society for Music Information Retrieval Conference (ISMIR 2009)
- [13] Emiru Tsunooy, Geroge Tzanetakisyy, Nobutaka Onoy, Shigeki Sagayama, “AUDIO GENRE CLASSIFICATION USING RHYTHM AND BASS-LINE PATTERN INFORMATION”, Graduate School of Information Science and Technology University of Tokyo, Japan, Computer Science Department University of Victoria, Canada(2009)
- [14] Emiru Tsunoo, Nobutaka Ono, Shigeki Sagayama. “MUSICAL BASS-LINE PATTERN CLUSTERING AND ITS APPLICATION TO AUDIO GENRE CLASSIFICATION” Graduate School of Information Science and Technology The University of Tokyo, Japan, 10th International Society for Music Information Retrieval Conference (ISMIR 2009)
- [15] Cyril Laurier and Perfecto Herrera, “Mood Cloud : A Real-Time Music Mood Visualization Tool”, Music Technology Group Universitat Pompeu Fabra, Ocata 1, 08003 Barcelona, Spain,(2007)
- [16] Yi-Hsuan Yang, Yu-Ching Lin, Heng-Tze Cheng, and Homer Chen. “Mr. Emo: Music Retrieval in the Emotion Plane”, National Taiwan University, International Multimedia

- Conference, Proceeding of the 16th ACM international conference on Multimedia,(2008)
- [17] [http://blogs.msdn.com/blogfiles/audiofool/WindowsLiveWriter/LouderSoundsBetter\\_12855/FletcherMunson\\_EqualLoudness2.jpg](http://blogs.msdn.com/blogfiles/audiofool/WindowsLiveWriter/LouderSoundsBetter_12855/FletcherMunson_EqualLoudness2.jpg)
- [18] Jouni Paulus, Anssi Klapuri, "Measuring the Similarity of Rhythmic Patterns", Tampere University of Technology Institute of Signal Processing P.O. Box 553, FIN-33101 Tampere, Finland,(2002)
- [19] Perfecto Herrera-Boyer, Geoffroy Peeters and Shlomo Dubnov, "Automatic Classification of Musical Instrument Sounds", IMTG-IUA, Universitat Pompeu Fabra, Barcelona, Spain, IRCAM, Paris, France and The Hebrew University, Jerusalem, Israel, Journal of New Music Research, 2003
- [20] J. Foote, "Visualizing Music and Audio using Self-Similarity", ACM Multimedia Conference,(1999).
- [21] [http://neural.cs.nthu.edu.tw/jang/books/audioSignalProcessing/dpDtw\\_chinese.asp](http://neural.cs.nthu.edu.tw/jang/books/audioSignalProcessing/dpDtw_chinese.asp)
- [22] Donald J. Bemdt, JameCs lifford, "Using Dynamic Time Warping to Find Patterns in Time Series", Information Systems Department, Stern School of Business, New York University, 44 West 4th Street, New York, New York 10012-1126, AAAI Technical Report WS-94-03. Compilation copyright © (1994)
- [23] Hagen Kaprykowsky, Xavier Rodet, "GLOBALLY OPTIMAL SHORT-TIME DYNAMIC TIME WARPING APPLICATION TO SCORE TO AUDIO ALIGNMENT", Ircam – Centre Pompidou – Analyse Synthèse
- [24] 浦田泰宏著，(2006)，和弦進行活用與演奏秘笈，孫宇虹、蕭良悌譯，典絃音樂文化國際事業有限公司
- [25] <http://allmusic.com/>
- [26] Judith C. Brown, "Automatic classification of killer whale vocalizations using dynamic time warping", Physics Department, Wellesley College, Wellesley, Media Lab, Massachusetts, Institute of Technology, (2006) Acoustical Society of America