

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

以情緒為基礎之自動化遊戲場景配樂產生系統
Emotion-based Automated Music Generation for
Game Scenario

研究生：顏志豪

指導教授：黃志方 教授

曾毓忠 教授

中華民國一百零一年七月

以情緒為基礎之自動化遊戲場景配樂產生系統
Emotion-based Automated Music Generation for
Game Scenario

研究生：顏志豪

Student : Chih-Hao Yen

指導教授：黃志方

Advisor : Chih-Fang Huang

曾毓忠

Yu-Chung Tseng



聲音與音樂創意科技碩士學位學程

碩士論文

A Thesis

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 2012

Hsinchu, Taiwan, Republic of China

中華民國一百零一年七月

以情緒為基礎之自動化遊戲場景配樂產生系統

學生：顏志豪

指導教授：黃志方

曾毓忠

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

本論文嘗試設計一個建構於情緒平面的畫線系統，讓遊戲設計者可以依照想帶給玩家的情緒感知，在平面上畫出想產生的情緒音樂線段。系統將根據線段，依照對應之相關情緒的音樂特徵值，及加入一些樂理或經驗，自動產生音樂。另外本論文建構數個隱含情緒的遊戲場景情節，並於遊戲內加入依照情緒音樂線段轉化的場景顏色、以及前述所產生的情緒音樂，測試玩家的情緒反應。本論文的實驗結果，受測者單純聆聽系統自動產生的情緒音樂，於情緒平面上，其四個代表情緒感知程度的象限，在代表開心興奮的第一象限，辨識率為 78%，代表生氣緊張的第二象限為 33%，代表傷心難過的第三象限為 63%，平靜沉穩的第四象限為 64%，平均為 60%，而只加入遊戲場景不含顏色的階段，結果為 70%、26%、66%、86%，平均為 62%，再加入情緒線段轉換的合適顏色做為背景，結果為 73%、40%、83%、66%，平均為 65.5%。首先顯示受測者只聽系統產生之情緒音樂能有及格以上的辨識率，而加入場景和顏色後，其情緒感知辨識率均提升。本系統可以做為提供遊戲設計者的一個配樂產生系統，以及場景顏色使用的參考。

Emotion-based Automated Music Generation for Game Scenario

Student : Chih-Hao Yen

Advisor : Chih-Fang Huang

Yu-Chung Tseng

Master Program of Sound and Music Innovative Technologies

College of Engineering

National Chiao Tung University

ABSTRACT

This thesis is to design an emotion-based automated music generation system by drawing a curve on the emotion plane. System generates music according to some music theory and music features which related to the curve position. This thesis also creates some game scenes which can evoke emotion perception, and adds background by using suited colors converted by the curve. The experiments test subject's emotional perception in four quadrants of the emotion plane by using the emotional music, colors and emotion evoking game. The results showed that subjects who were only listening to music, the emotion recognition rate is 60%, then add the game scenes without the background color, the emotion recognition rate is 62%, and finally add the background colors, the emotion recognition rate increased to 65.5%. This system can provide game designers a soundtrack generation, and the color reference of game scenes.

誌謝

很開心能夠進入這個很特別且全台獨一無二的有趣研究所，讓我學習到很多電腦與音樂聯結的各式各樣的應用。當初剛進這研究所時就曾經想過，是不是最後的論文可以做有關遊戲和音樂的結合，而最後竟也真的很有關聯性。

首先感謝指導教授黃志方老師，除了課業和論文方向上的指導，研究所期間引薦我和實驗室的同學們到遊戲公司合作，那段時間不但讓我的程式能力進步、也學習到了與公司的同仁團隊合作方面的事物，而最後作出的產品令我們都很有成就感。還要感謝曾毓忠教授的教導，聲音設計方面和數位樂團的課程都讓我學習的相當開心，也讓我培養了很實用的技能。

感謝我的實驗室同學們!! Webber、科科、志翰、朱朱、明瑾、大家一起共度了很多事情，會成為難忘的回憶，大家也要繼續連絡下去喔。感謝我的母親辛苦培育我到一直到碩班，還總是擔心我，真的謝謝您。

目錄

摘要.....	i
ABSTRACT.....	ii
誌謝.....	iii
目錄.....	iv
表目錄.....	vi
圖目錄.....	vii
第一章、 概論.....	1
1.1 研究動機.....	1
1.2 研究目標與方法.....	2
1.3 論文架構.....	3
第二章、 相關研究.....	4
2.1 基本樂理介紹.....	4
2.2 MIDI.....	5
2.3 情緒平面.....	6
2.3.1 Thayer' s emotion plane.....	6
2.3.2 表情情緒分類的實驗.....	7
2.4 音樂與情緒.....	8
2.5 顏色與情緒.....	14
2.5.1 顏色系統 - Munsell Color System (孟塞爾顏色系統).....	14
2.5.2 顏色系統 - 電腦的 RGB model.....	16
2.5.3 顏色與情緒的關係.....	16
2.6 遊戲場景與情緒：EVG.....	20
2.7 節奏複雜度 - LHL.....	21
第三章、 實作方法.....	22
3.1 系統架構.....	22
3.2 與情緒平面連結.....	24
3.3 節奏產生.....	24

3.4 旋律產生.....	26
3.5 伴奏、音量強度、速度.....	28
3.6 顏色轉換演算法.....	29
3.7 XML 與 PROCESSING.....	32
3.8 遊戲場景設計.....	34
第四章、 系統展示.....	35
4.1 自動作曲系統基本介面介紹.....	35
4.2 加入分段 TAG.....	37
4.3 遊戲場景畫面展示.....	39
第五章、 實驗.....	40
5.1 實驗目的及步驟.....	40
5.2 實驗一：12 首情緒音樂的問卷結果分析.....	42
5.3 第一次遊戲實驗結果.....	47
5.4 第二次遊戲實驗結果.....	50
5.5 實驗結果.....	52
第六章、 總結.....	53
6.1 結論.....	53
6.2 未來研究.....	53
參考文獻.....	54
附錄 A.....	56
附錄 B.....	57
附錄 C.....	58
附錄 D.....	59

表目錄

表 2.1	Pitch Class 表.....	6
表 2.2	八區域內的音樂特徵.....	9
表 2.3	音樂特徵與 Arousal 和 Valence 的關係	10
表 2.4a	音樂專家對 16 首歌的音樂特徵評分	11
表 2.4b	音樂特徵的評分高低意義.....	12
表 2.5	Valence、Arousal 轉換至音樂特徵評分的公式表	12
表 2.6	Naz Kaya 使用的受測顏色一覽.....	18
表 2.7	情緒與顏色關係的統計數據.....	18
表 3.1	四種音樂特徵的最高最低評分值及其位置.....	24
表 3.2	大調與小調之調性權重對應表.....	26
表 3.3	各顏色 RGB 值與其最具代表情緒	29
表 3.4	各情緒於情緒平面上的位置與其範圍.....	30
表 5.1	受測者對屬於象限 1 (V+ A+) 三首音樂的情緒反應統計	43
表 5.2	受測者對屬於象限 2 (V- A+) 三首音樂的情緒反應統計	43
表 5.3	受測者對屬於象限 3 (V- A-) 三首音樂的情緒反應統計	44
表 5.4	受測者對屬於象限 4 (V+ A-) 三首音樂的情緒反應統計.....	44
表 5.5a	各象限音樂總體平均的全體受測者回答正確率.....	46
表 5.5b	各象限音樂總體平均的未受過音樂訓練受測者回答正確率表.....	46
表 5.6a	第一次遊戲的情緒感知實驗結果 - 全體受測者.....	48
表 5.6b	第一次遊戲的情緒感知實驗結果 - 僅未受音樂訓練的受測者	48
表 5.7	第一次遊戲場景 3(情緒象限 2)的情緒感知問卷數據	49
表 5.8a	第二次遊戲的情緒感知實驗結果 - 全體受測者.....	50
表 5.8b	第二次遊戲的情緒感知實驗結果 - 僅未受音樂訓練的受測者.....	50
表 5.9	第二次遊戲場景 3(情緒象限 2)的情緒感知問卷數據	51

圖目錄

圖 2.1	七種通用音符及七種通用休止符.....	4
圖 2.2	C 大調自然音階與 C 小調和聲音階.....	5
圖 2.3	Thayer's arousal-valence emotion plane.....	7
圖 2.4	學齡前及成人的表情情緒分類實驗結果圖.....	8
圖 2.5	Dimensional Emotion Space in Octal Form.....	9
圖 2.6	A-V 情緒平面與音樂特徵的關係圖.....	13
圖 2.7	Munsell Color System.....	14
圖 2.8	著色完整的孟塞爾顏色系統.....	15
圖 2.9	Digital color 實驗中的受測顏色.....	16
圖 2.10	Digital color 實驗分布於情緒 Arousal-Valence 平面上之結果.....	17
圖 2.11	LHL 節奏複雜度之樹狀圖及分析例子.....	21
圖 3.1	情緒自動作曲系統流程圖.....	23
圖 3.2	LHL 逆向法示意圖.....	25
圖 3.3	伴奏範例.....	28
圖 3.4	本系統之顏色轉換演算法，所有點的顏色轉換結果.....	30
圖 3.5	以 XML 格式存取系統建立出來的音樂時間、顏色、速度等資訊.....	32
圖 3.6	中介資訊檔產生流程.....	32
圖 3.7a	processing 簡易程式範例.....	33
圖 3.7b	執行之結果.....	33
圖 4.1	系統操作流程.....	35
圖 4.2	系統介面圖.....	36
圖 4.3	按下 setup 鍵後的視窗.....	36
圖 4.4	設定小節數後的線段圖.....	37
圖 4.5	加入兩個新 tag 的線段.....	37

圖 4.6	加入兩個新 tag 後的 setup 視窗.....	37
圖 4.7	設定好每個分段小節使用點數.....	38
圖 4.8	遊戲場景使用之配樂的原始繪製線段，start、2、3、4 共四場景.....	38
圖 4.9	所設計之四個遊戲場景畫面.....	39
圖 5.1	受測者所受音樂訓練時間的比例圖.....	40
圖 5.2	實驗用 12 首歌採用的 12 個線段.....	42
圖 5.3	全體受測者與未受音樂訓練者的各象限正確辨識率比較.....	47



第一章、 概論

1.1 研究動機

音樂自古以來都在人們的生活中佔有重要的一環，不論是否為不同的國家文化、社會階層、民俗風情，均能發展出各式各樣的音樂風格。而許多作曲家也會將所感受到的情緒放入音樂裡，使聽者能夠試著體會作曲家想要表達的心情。

隨著電腦科技的進步，近年電玩遊戲更是蓬勃發展，其中含有劇情導向的遊戲也占了相當大的一部分。遊戲設計者在設計遊戲時一定也希望帶給玩家某種的心情感受，而音樂可以是很好的傳達方式；但是大部分的遊戲設計者並不同時會作曲，而目前的遊戲配樂方式常採用現有的編曲作品，也並不一定能符合遊戲設計師想傳達的情緒感受。

此論文旨在建立一個讓任何人能夠簡易上手的系統 --- 藉由畫線在一個代表情緒的平面上，即可自動產生令聽者具有情緒感知的音樂。因此相當的適合電玩設計者使用，設計者只需依照遊戲情節想帶給玩家的情緒感知，畫上對應的情緒曲線即可產生符合遊戲情節情緒的配樂。

而在既有的研究中，顏色也能帶給人有不同的情緒感知，本論文將上述系統繪製好的曲線轉換成符合情緒的顏色，並加入遊戲場景的色調中。如此可探討除了情緒音樂外，玩家的情緒在遊戲中所受的遊戲場景顏色影響為何，並藉此能給遊戲設計者在設計具有情緒情節的遊戲時，對於場景顏色設計的參考。

1.2 研究目標與方法

為了達成研究動機的目的，本論文建立之系統應具有以下目標：

- a. 簡易的劃線、線段設定 (使其可配合遊戲場景的分段產生音樂)。
- b. 使產生之音樂具有劃線者欲帶給聽者的情緒感知。
- c. 具有將曲線轉換成相關情緒之顏色的功能。

而要達成以上研究目標，嘗試的方法依序如下：

- a. 將可以形容、分類情緒的平面直接設計成一簡易 GUI (Graphical User Interface)，供使用者直覺的在平面上依照想產生的情緒感知繪製線條。
- b. 目前已經有一些有關音樂特徵對應至情緒感知的研究，本研究要試著將這些與情緒變化有關聯的音樂特徵，如節奏複雜度、樂曲速度、調性、音高範圍等，建構成自動作曲的各個演算法，讓系統能夠產生符合情緒的音樂。
- c. 另外，顏色與情緒也有一定的相關性，但須要將顏色放入情緒平面中適當的位置，且必須建立一個關於顏色變化的演算法，使情緒平面上每一點都能轉換成適合的顏色。
- d. 在完成以上步驟後，本研究會再建置數個簡易的遊戲場景，使上述方法產生之音樂與顏色能夠適當的加入遊戲場景，便於受測者實驗。
- e. 最後將遊戲場景、遊戲配樂、遊戲色調經由受測者直接試玩來取得相關的實驗數據，如此能夠確立本研究之自動作曲系統產生之音樂及顏色、遊戲場景三者，對於情緒之正確對應性及程度分析。

1.3 論文架構

接續本章概論，本論文結構安排如下：

第二章：相關研究。

先介紹基本樂理如音符、音階、及電腦使用的一種記錄音樂的方式：MIDI。再來說明情緒平面模型、音樂對情緒、顏色對情緒的相關研究或應用、具有情緒的遊戲場景設計；有助於理解本論文建構之系統的組成細節原理。

第三章：實作方法。

詳細說明系統流程，以及系統每個步驟所使用的演算方法。

第四章：系統展示。

將以清楚的圖文展示本論文開發之系統及其細部功能。

第五章：實驗。

實驗流程說明、結果歸納。

第六章：總結。

總結成果，以及未來發展方向。

第二章、 相關研究

2.1 基本樂理介紹

樂譜是一種以印刷或手寫製作，用符號來記錄音樂的方法[1]，例如可以記錄速度、音量、演奏手法等資訊。而音符在樂譜中是用來表示相對長度的固定音高單位[2]，各種長度和音高的音符，建構成小節，小節再組合成一首完整的樂曲。音符彼此為相對的倍數或是半數的關係，例如全音符為二分音符的兩倍、二分音符為四分音符的兩倍，以此類推。而標記方法，全音符為無符幹、無符尾，僅有一白橢圓的符頭；二分音符為一符幹及白橢圓符頭；四分音符為一符幹及黑橢圓符頭；八分音符開始則是在四分音符符幹上加一符尾，十六分音符加二符尾，依此類推。通用的音符有七種，從全音符至六十四分音符[3]。另外，休止符則是表示樂音停頓或靜止的符號，通常的休止符也有七種，而長度則跟上述的七種音符相當，七種音符及七種休止符如圖 2.1 所示。

全音符	二分音符	四分音符	八分音符
 ← 符頭	 ← 符幹		 ← 符尾
十六分音符	三十二分音符	六十四分音符	
			
全休止符	二分休止符	四分休止符	八分休止符
			
十六分休止符	三十二分休止符	六十四分休止符	
			

圖 2.1 七種通用音符及七種通用休止符

以一音為基礎，在其之上按照高低排列若干音的階級，至其高八度音為止，形成音階。兩音高度間的距離稱為音程，最小音程為半音。近代最通用的音階，有大音階和小音階兩種，兩者皆屬自然音階，且音階內各個音之間的音程排列不太相同，由五個全音和兩個半音組成。而適合於和聲使用的合聲小音階則是將小音階中第七音再升半音。圖 2.2 為 C 大調及和聲 C 小調音階的差異；通常來說，大音階多表示快樂而歡騰的情緒，小音階多表示悲傷而沉悶的情緒。

C 大調自然音階



C 小調和聲音階

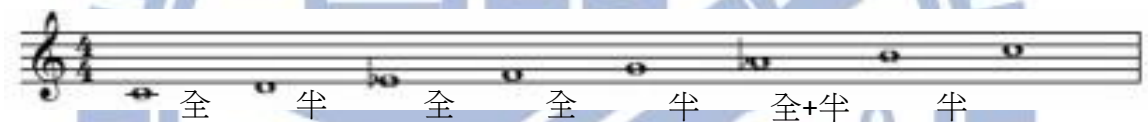


圖 2.2 C 大調自然音階與 C 小調和聲音階

本論文建構之自動作曲系統，將會使用最小至十六分之一音符(或休止符)，及大調自然音階、小調和聲音階內的音作為基礎，來產生情緒音樂。

2.2 MIDI

MIDI[4]為 Musical Instrument Digital Interface 的縮寫，是一種標準的電子通訊協定，可在電腦、電子樂器等之間交換演奏資料。MIDI 不傳送聲音，而是傳送音樂的數據、事件，例如音調、音量、演奏手法等的資料，再由個別電腦輸出，因此不同的電腦上，輸出的聲音也會因為音源器而效果不同。

在 MIDI 中，以 0~127 的數字代表音高資訊，而中央 C，通常又稱為 C4，音高資訊為 60。0~127 代表的音高是什麼，可以利用 Pitch Class[7] 的概念，將一個八度音程分為 0~11 的組別，每個組別包含同音的所有 Pitch，如表 2.1。因此將 MIDI 音高資訊的 $0 \sim 127 \bmod 12$ 後得到 0~11，即可知道音名，而實際音高也就是每隔 12，相差一個八度，例如 MIDI 音高 74，將 $74 \bmod 12 = 2$ ，為音名 D，又因為 74 位在中央 C 所屬八度音程(60~71)的再高八度音程內，因此 74 為高音 D5。

音名	C	C [#] D ^b	D	D [#] E ^b	E	F	F [#] G ^b	G	G [#] A ^b	A	A [#] B ^b	B
Pitch Class	0	1	2	3	4	5	6	7	8	9	10	11

表 2.1 Pitch Class 表

2.3 情緒平面

2.3.1 Thayer's emotion plane

由於情緒很難用普通的方法去量化形容，但可以大略的將各種情緒去做分類。最典型的為 Thayer's emotion plane[5][6]，使用二維平面將常見的情緒分類。二維分別是 Arousal 及 Valence，Arousal 指的是激發程度，表示某種情緒是屬於較激動或是屬於較平靜的；Valence 則是代表情緒的正向或是負向，如圖 2.3 所表示。舉例來說，Excited(興奮)這個情緒分類在正向情緒且激發程度較高，則在第一象限；而 Sad(悲傷)這個情緒分類在負向情緒且激發程度較低，則在第三象限。Thayer's emotion plane 就是用二維平面將常見的情緒分成四大類，使情緒可以用較統一的標準去形容。



圖 2.3 Thayer's arousal-valence emotion plane

2.3.2 表情情緒分類的實驗

James A. Russell 及 Merry Bullock 在他們所著作的 *On the Dimensions Preschoolers Use to Interpret Facial Expressions of Emotion* [12] 論文中，首先請成人受測者，對每個情緒於情緒平面上的 arousal、valence 位置定位，將十種情緒較精確的分類在二維 Arousal-Valence 平面上；再來對 2~4 歲的學齡前孩童使用表情及圖片的比對分類方式，嘗試驗證即使孩童對情緒的感受是否類似於成人，而如圖 2.4 所示，學齡前孩童及成年人的實驗結果相當類似，情緒在分類平面的位置上很接近，差別在於其位置對於中心點的旋轉位移(順或逆)。

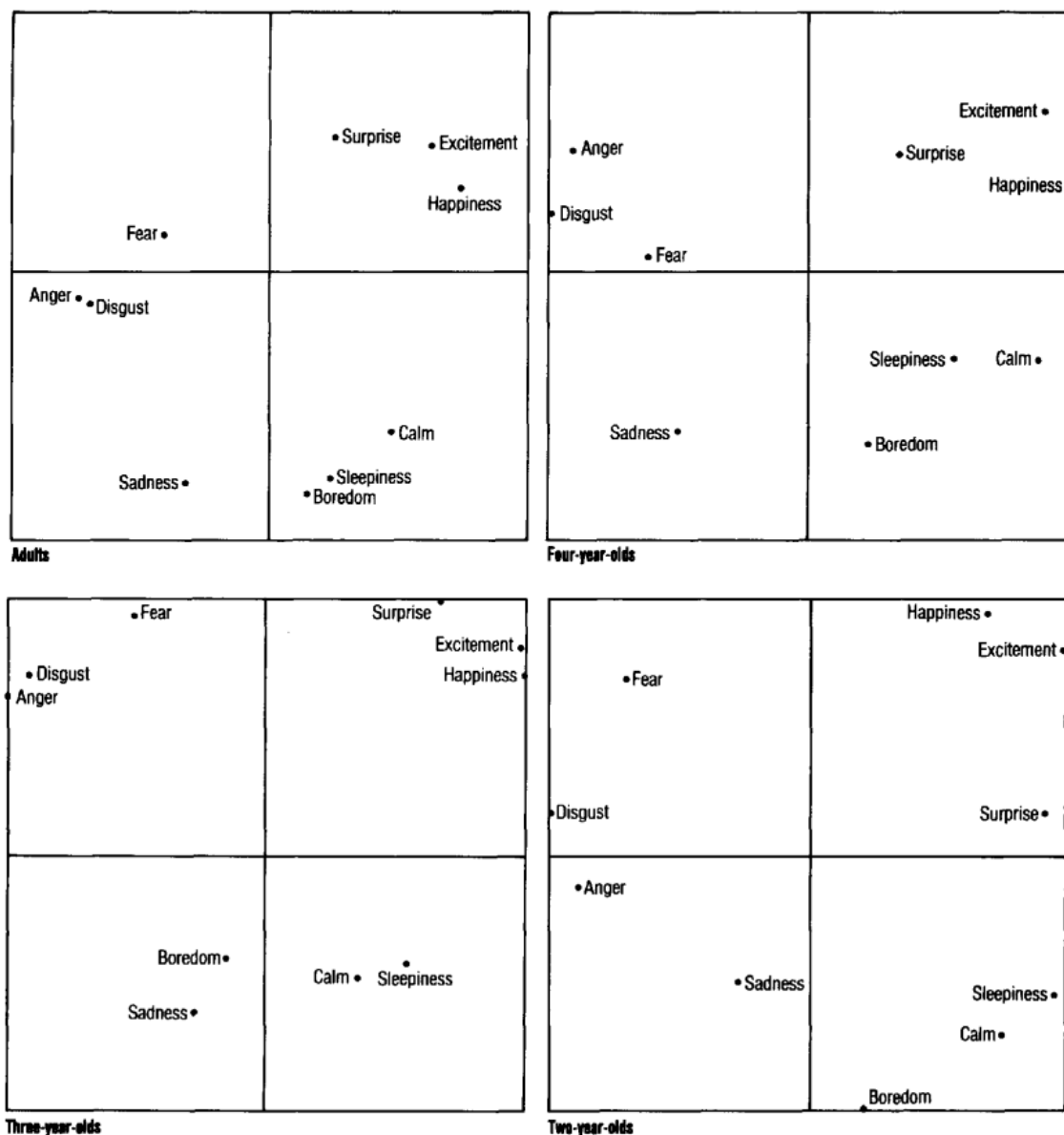


圖 2.4 學齡前及成人的表情情緒分類實驗結果圖

2.4 音樂與情緒

許多研究者將音樂歸納到不同的情緒，並且以前述之情緒平面的方式來表示。Steven R. Livingstone 和 Andrew R. Brown [9]將 A-V 情緒平面的四個象限更細分成八個區域(圖 2.5)，並提出每個區域內較符合的音樂特徵(表 2.2)，其中以粗黑體標示的音樂特徵表示在該區域中較具代表性。

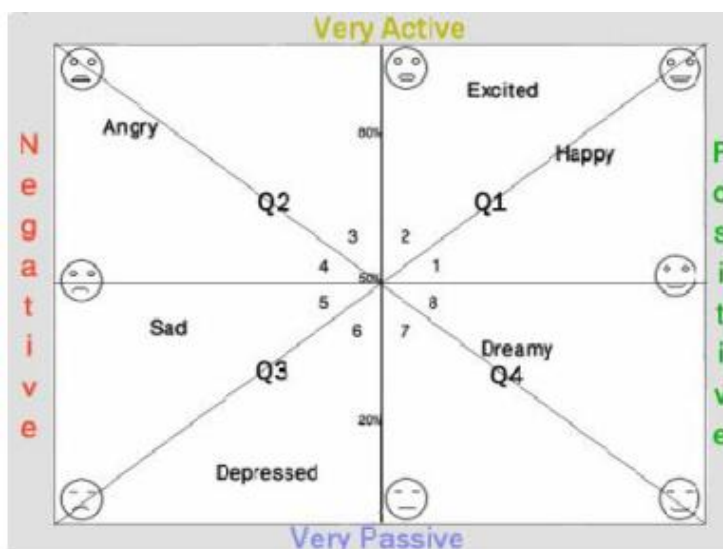


圖 2.5 Dimensional Emotion Space in Octal Form

Octant	2DES Octant Rules
1	<i>Mode Major(19), Tempo Fast(16), Harmony Simple(8), Loudness Loud(7), Articulation Staccato(5), Pitch High(3), Rhythm Flowing(3), Pitch Range High(2), Pitch Variation Large(2), Pitch Contour Up(2), Note Onset Rapid(2), Rhythm Smooth(2), Rhythm Activity(2), Loudness Medium(1), Loudness Soft(1), Loudness Variation Small(1), Loudness Variation Rapid(1), Loudness Variation Few(1), Pitch Low(1), Pitch Range Low(1), Pitch Contour Down(1), Timbre Few(1), Rhythm Rough(1)</i>
2	<i>Tempo Fast(20), Loudness Loud(10), Mode Major(8), Pitch High(4), Pitch Variation Large(4), Harmony Simple(4), Note Onset Rapid(4), Pitch Range High(3), Pitch Contour Up(3), Articulation Staccato(3), Articulation non-legato(2), Harmony Complex(2), Rhythm Flowing(2), Rhythm Activity(2), Rhythm Smooth(2), Loudness Variation Small(1), Loudness Variation Few(1), Loudness Variation Rapid(1), Pitch Low(1), Pitch Range Low(1), Pitch Variation Small(1), Pitch Contour Down(1), Timbre Few(1), Timbre Many(1), Tempo Slow(1), Vibrato fast(1), Rhythm Complex(1), Rhythm Firm(1), Metre Triple(1), Tonality Tonal(1)</i>
3	<i>Mode Minor(14), Loudness Loud(9), Tempo Fast(9), Harmony Complex(8), Note Onset Rapid(5), Pitch Contour Up(5), Pitch High(4), Pitch Range High(3), Pitch Variation Large(3), Loudness Soft(2), Rhythm Complex(2), Loudness Variation Large(2), Timbre Sharp(2), Articulation Non-legato(2), Pitch Variation Small(2), Articulation Staccato(2), Note Onset Slow(2), Timbre Many(1), Vibrato Fast(1), Rhythm Rough(1), Metre Triple(1), Tonality Tonal(1), Tonality Atonal(1), Tonality Chromatic(1), Loudness Variation Rapid(1), Pitch Low(1)</i>
4	<i>Mode Minor(12), Harmony Complex(6), Articulation Legato(3), Pitch Variation Small(3), Tempo Fast(3), Loudness Loud(2), Loudness Soft(2), Loudness Variation Large(2), Note Onset Rapid(2), Note Onset Sharp(2), Note Onset Slow(2), Timbre Sharp(2), Loudness Variation Rapid(1), Pitch High(1), Pitch Low(1), Pitch Range High(1), Pitch Variation Large(1), Pitch Contour Up(1), Pitch Contour Down(1), Timbre Many(1), Harmony Melodic(1), Tempo Slow(1), Articulation Staccato(1), Rhythm Complex(1), Tonality Atonal(1), Tonality Chromatic(1)</i>
5	<i>Tempo Slow(15), Articulation Legato(6), Mode Minor(7), Harmony Complex(7), Loudness Soft(3), Harmony Simple(3), Pitch Low(3), Note Onset Slow(3), Pitch Range Low(2), Pitch Contour Down(2), Rhythm Firm(2), Loudness Loud(1), Loudness Variation Small(1), Loudness Variation Few(1), Pitch Variation Small(1), Pitch Contour Up(1), Mode Major(1), Timbre Few(1), Timbre Soft(1), Harmony Melodic(1), Note Onset Rapid(1), Vibrato Deep(1), Rhythm Smooth(1), Tonality Chromatic(1)</i>
6	<i>Loudness Soft(5), Tempo Slow(5), Pitch Variation Small(3), Articulation Legato(3), Note Onset Slow(3), Pitch Low(3), Pitch Range Low(2), Loudness Variation Rapid(1), Pitch High(1), Pitch Contour Down(1), Mode Minor(1), Timbre Few(1), Harmony Complex(1), Vibrato Deep(1), Metre Duple(1), Tonality Tonal(1)</i>
7	<i>Tempo Slow(10), Loudness Soft(9), Articulation Legato(5), Note Onset Slow(3), Pitch Low(2), Pitch Range Low(2), Pitch Variation Small(2), Timbre Soft(2), Harmony Simple(2), Mode Minor(1), Loudness Variation Rapid(1), Loudness Variation Few(1), Pitch High(1), Note Onset Rapid(1), Vibrato Intense(1), Rhythm Smooth(1), Rhythm Flowing(1), Rhythm Firm(1), Metre Duple(1)</i>
8	<i>Mode Major(12), Harmony Simple(5), Tempo Slow(3), Articulation Staccato(3), Loudness Soft(2), Pitch Contour Up(2), Loudness Variation Few(1), Pitch High(2), Pitch Range High(1), Pitch Range Low(1), Pitch Low(1), Mode Minor(1), Timbre Soft(1), Vibrato Intense(1), Rhythm Smooth(1), Rhythm Flowing(1), Tonality Tonal(1)</i>

表 2.2 八區域內的音樂特徵

另外，Isaac Wallis, Todd Ingalls, 及 Ellen Campana 提出一套有關產生情緒音樂的演算法系統[8]，較明確地將一些音樂特徵的變化與 Arousal/Valence 關聯，例如 arousal 越高，則音樂速度較快；valence 較低，則和聲複雜度較高。其他如表 2.3 所示。

← low -----Arousal----- high →
← Smooth-----rhythmic roughness-----rough→
← slow -----tempo----- fast →
← legato -----articulation-----staccato →
←low-----Valence-----high→
←darkest ----- mode brightness-----Brightest →
←-----harmony complexity (inverse)----- →
← low -----pitch register----- high →

表 2.3 音樂特徵與 Arousal 和 Valence 的關係

其中 mode brightness 依照作者的研究，從 Brightest 至 darkest 為以下順序：Lydian, Ionian, Mixolydian, Dorian, Aeolian, Phrygian and Locrian。(註：以上為希臘調式[3])。而 harmony complexity(和聲複雜度)則是與 Valence 有著反比的關係，Valence 越低，則 harmony complexity 越高，漸漸的會有七和弦的出現。

由以上兩論文發現，有許多音樂特徵在 A/V 情緒平面的定義上，獲得較一致的分類，這對本論文使用情緒圖為基礎來開發情緒自動音樂作曲相當有幫助。以下將接著介紹本論文所架構之系統所主要依據的相關研究理論。

Patrick Gomez 及 Brigitta Danuser [10][11] 做了有關音樂 11 個特徵與生理反應的相關研究，其中他們將音樂特徵量化，並將其對應到 Arousal-Valence 平面上的任一點。方法如下：

- a. 先請三位受過音樂訓練的專家對 16 首音樂片段(作品名稱見附錄 A)的 10 個音樂特徵做 1~7 的評分(Sound intensity 為實際測量值，單位為 Db)，並經由討論調整後得到表 2.4a 的平均數據，表 2.4b 為評分高低的意義。
- b. 受測者聽這 16 首音樂片段後給出 Arousal 和 Valence，1~9 分的評分
- c. 將其中受測者最易辨識的 6 個音樂特徵的 Arousal-Valence 分數與專家的評分數據，產生個別的對照回歸轉換公式(見表 2.5)，使情緒平面上的任一點都能轉換成某個確切的音樂特徵分數，如圖 2.6。

Musical	Sound	Tempo	Rhythm	Pitch	Mode	Consonance
1	67.6	6.3	1.7	3.0	6.3	6.3
2	69.7	6.3	1.3	5.7	7.0	6.3
3	62.2	4.3	2.7	2.0	6.3	5.3
4	66.0	4.7	2.7	5.3	6.3	5.3
5	61.0	2.7	2.0	6.0	6.0	6.3
6	66.7	2.0	2.7	2.3	5.3	5.3
7	71.8	5.0	6.0	4.7	1.3	3.0
8	65.2	3.3	2.7	3.0	1.7	2.7
9	68.6	2.0	2.3	4.0	6.0	6.3
10	61.9	3.7	3.0	4.3	5.0	5.3
11	65.1	2.7	5.3	4.0	1.3	5.7
12	77.5	7.0	5.7	5.3	a	3.0
13	73.1	2.3	3.7	5.7	3.0	4.3
14	67.7	2.0	1.7	3.7	2.7	4.7
15	62.3	1.7	5.0	3.3	5.7	5.0
16	58.9	2.3	2.0	2.3	6.0	6.0

表 2.4a 音樂專家對 16 首歌的音樂特徵評分，僅列 6 個主要特徵，完整見附錄 B

^aSelection 12 為無調性吉他獨奏

音樂特徵	評分高低意義
Sound intensity	為實際測量平均值
Tempo	1 = slow , 7 = fast
Rhythm	1 = outstanding , 7 = vague
Pitch range	1 = narrow , 7 = wide
Mode	1 = minor , 7 = major
Consonance	1 = dissonant , 7 = consonant

表 2.4b 音樂特徵的評分高低意義

音樂特徵評分 轉換結果(1~7 分)	回歸公式， Valence、Arousal 為 1~9 間的值
Sound intensity(dB)	$= 58.4 + 2.14 \times \text{Arousal} - 0.18 \times \text{Valence} \times \text{Arousal}$
Tempo	$= 1.2 + 0.1 \times \text{Valence} + 0.43 \times \text{Arousal}$
Rhythm	$= 2.23 + 0.16 \times \text{Valence} + 0.35 \times \text{Arousal}$ $- 0.05 \times \text{Valence} \times \text{Arousal}$
Pitch range	$= 2.17 + 0.22 \times \text{Valence} + 0.37 \times \text{Arousal}$ $- 0.05 \times \text{Valence} \times \text{Arousal}$
Mode	$= 6.4 - 0.27 \times \text{Valence} - 0.71 \times \text{Arousal}$ $+ 0.11 \times \text{Valence} \times \text{Arousal}$
Consonance	$= 5.8 - 0.37 \times \text{Arousal} + 0.04 \times \text{Valence} \times \text{Arousal}$

表 2.5 Valence、Arousal 轉換至音樂特徵評分的公式表

舉例來說，若在情緒平面上 Valence = 7，Arousal = 5 的位置，則對應出來的 Sound intensity 約為 62dB，Tempo 值約為 4。

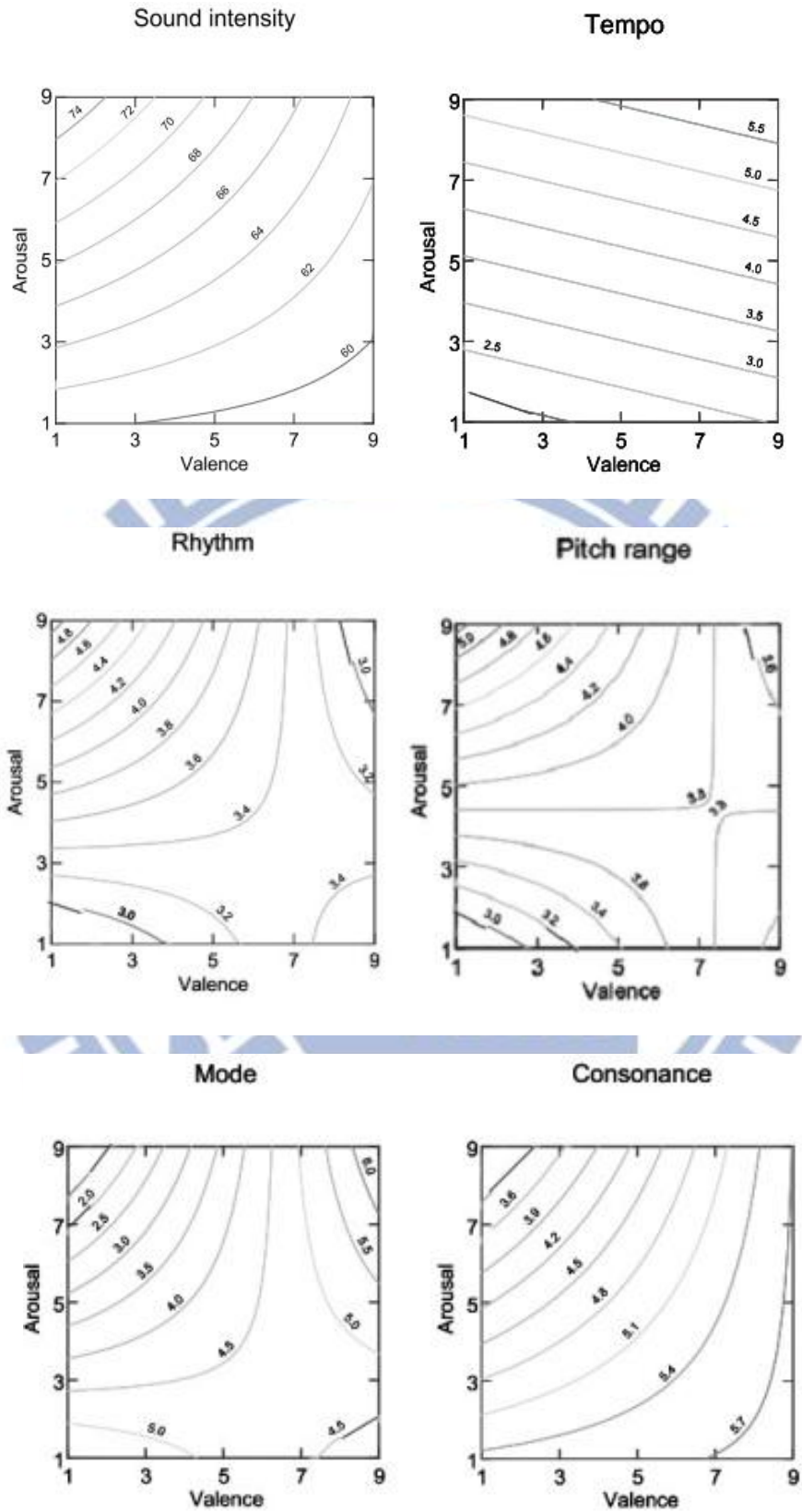


圖 2.6 A-V 情緒平面與音樂特徵的關係圖[11]。

2.5 顏色與情緒

2.5.1 顏色系統 - Munsell Color System (孟塞爾顏色系統)

Munsell Color System (孟塞爾顏色系統)是目前通用的一個比較色法標準，由美國藝術家 Albert H. Munsell 在 1989 年所創製，孟塞爾顏色系統空間看起來像是一個圓柱形，如圖 2.7 所示。

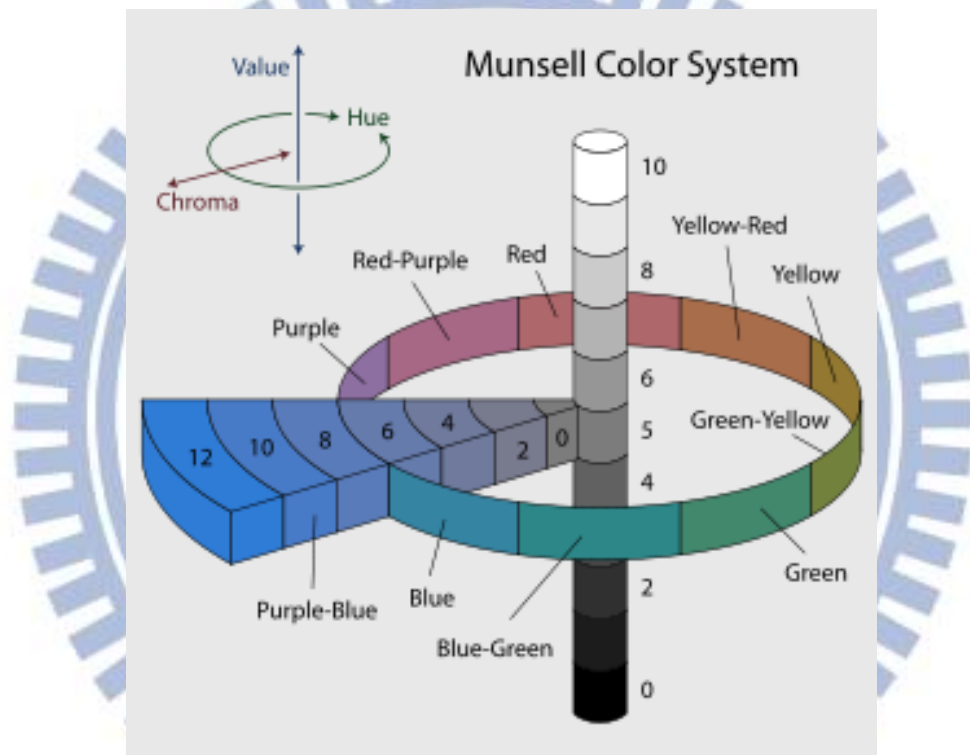


圖 2.7 Munsell Color System

(資料來源：<http://zh.wikipedia.org/wiki/File:Munsell-system.svg>)

顏色系統結構如下(http://en.wikipedia.org/wiki/Munsell_color_system)：

- 南北軸為明度(value)：共 10 等級，由全黑 0~全白 10。
- 色相(hue)：將一周分為五種主色調和五種中間色：紅(R)、紅黃(YR)、黃(Y)、黃綠(GY)、綠(G)、綠藍(BG)、藍(B)、藍紫(PB)、紫(P)、紫紅(RP)，相鄰兩位置再分成 10 份，共 100 份。

- c. 與軸的距離為色度(chroma)：表示色彩的純度，數值由中間 0 向外隨色調純度增加，沒有理論上的上限(普通的顏色實際上限為 10 左右，反光、熒光等材料可高達 30)。

對顏色的描述方式為：色相 + 明度 + 色度。例如 5B 5/10 代表中等明度 5，高色度 10 的正藍色 5B。因此，整個孟塞爾顏色系統的圓柱體會像是圖 2.8 所示。

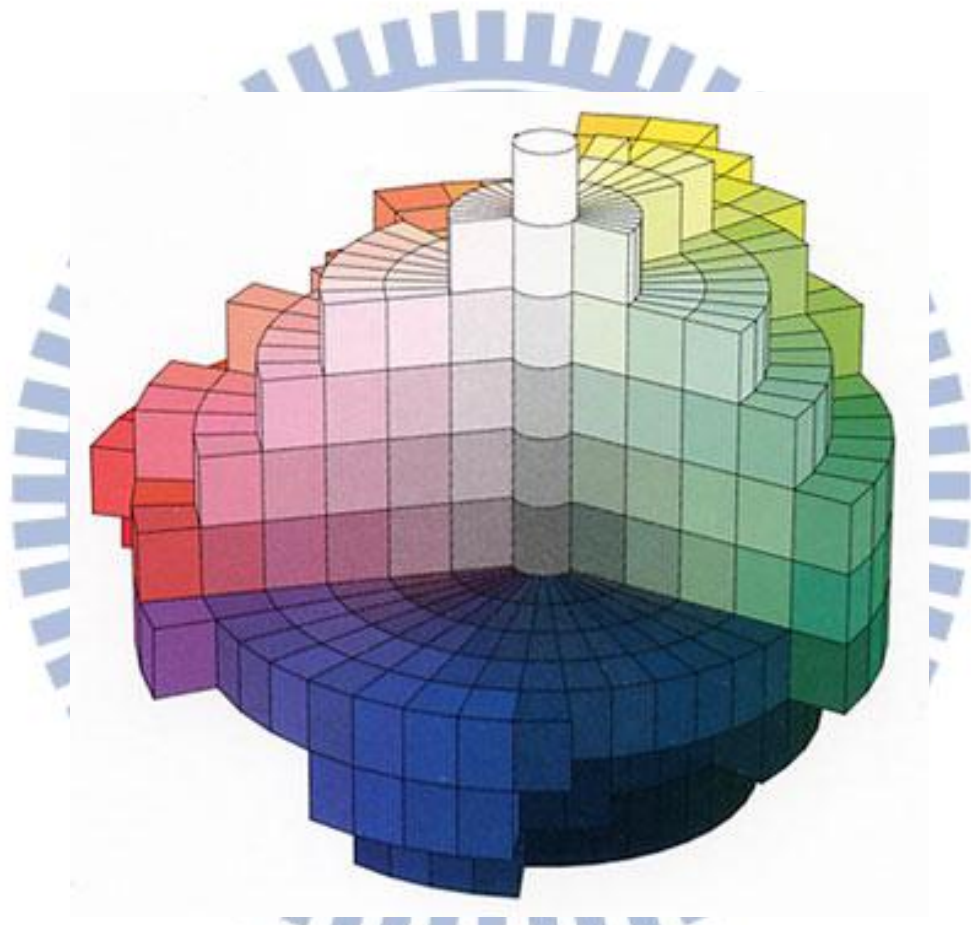


圖 2.8 著色完整的孟塞爾顏色系統

完整的 Munsell Color System 色表詳見連結：

[http://oceandatacenter.ucsc.edu/PhytoBlog/Images/Munsell_A0\[1\].pdf](http://oceandatacenter.ucsc.edu/PhytoBlog/Images/Munsell_A0[1].pdf)

2.5.2 顏色系統 – 電腦的 RGB model

現在最通用的電腦顏色顯示模式為 24bit 的 RGB model，表示方法為(Red, Green, Blue)，而 Red、Green、Blue 三原色各分到 8bit，也就是有 2 的 8 次方 = 256 個數值可以記錄。用 0~255 的數值代表各原色的強度，由強度小(0)至大(255)，這樣的方式可以產生 256 的 3 次方，1 千 6 百多萬種顏色組合，對人眼來說其中很多已分辨不開。

2.5.3 顏色與情緒的關係

在 Color and Emotion[13] 這本著作中，對情緒與顏色的關係做了很多的介紹和相關實驗；例如讓受測者對物品表面顏色(surface color)、對電腦產生的數位顏色(digital color)的評分實驗，圖 2.9 為 digital color 的實驗所使用的受測顏色，使用 Lch 表示法，Lch 也是一種表達顏色的方式，有點類似於孟塞爾顏色系統，其中 L = Lightness，c = chroma，h = hue angle，實驗結果顏色於情緒平面上的分布如圖 2.10，詳盡的 Valence-Arousal 數據請見附錄 C 表格中的 exp.III。

hue (°)	tone categories				
	dark	deep	vivid	brilliant	light
30 (red)	● L: 30, C: 30	● L: 30, C: 45	● L: 40, C: 60	● L: 50, C: 40	● L: 70, C: 30
80 (yellow)	● L: 60, C: 40	● L: 60, C: 70	● L: 80, C: 90	● L: 80, C: 60	● L: 80, C: 40
160 (green)	● L: 30, C: 30	● L: 40, C: 45	● L: 50, C: 60	● L: 40, C: 40	● L: 70, C: 20
260 (blue)	● L: 30, C: 20	● L: 40, C: 30	● L: 40, C: 45	● L: 60, C: 35	● L: 70, C: 25
320 (violet)	● L: 20, C: 25	● L: 30, C: 35	● L: 40, C: 40	● L: 50, C: 30	● L: 70, C: 20

	tone categories				
	black	dark	medium	light	white
achromatic color stimuli	● L:0, C:0	● L: 30, C: 0	● L: 50, C: 0	● L: 70, C: 0	● L:100, C:0
warm grays hue (80°)		● L: 30, C: 10	● L: 50, C: 10	● L: 70, C: 10	
cool grays hue (260°)		● L: 30, C: 10	● L: 50, C: 10	● L: 70, C: 10	

圖 2.9 Digital color 實驗中的受測顏色[13]

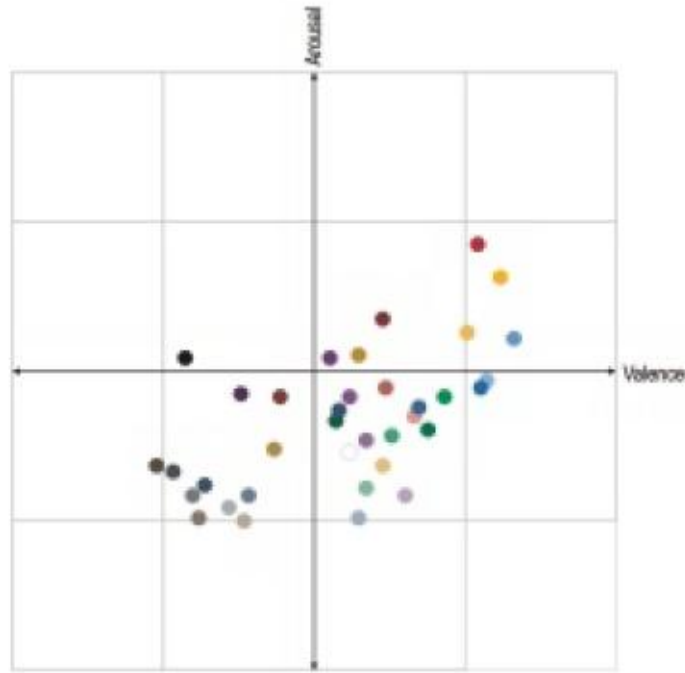


圖 2.10 Digital color 實驗分布於情緒 Arousal-Valence 平面上之結果[13]

以上實驗為對受測顏色在 Arousal-Valence 上評分，因此可以看到分布於情緒平面上面的顏色並無明顯的規律關係，受到受測者對顏色的喜好影響較大。

Naz Kaya 等人對 98 個大學生做了顏色與情緒關聯的研究[14]，他使用了 5 個基本色+5 個中間色+3 個灰階色為受測顏色，如表 2.6 所示，使用孟塞爾顏色系統。接著 Naz Kaya 舉了 22 個常見情緒描述，統計這些大學生受測者的數據。表 2.7 只展示有與隨後章節，本論文系統顏色轉換部分，所採用的相同情緒描述的統計結果表，詳盡 22 個情緒描述的統計表請見附錄 D。例如以紅色來說，最多數為 28 個受測者認為屬於生氣的情緒；因此在隨後章節，有關於本論文製作之系統的顏色轉換演算部分，就會將情緒平面上情緒為生氣的位置，對應至紅色，再繼續進行其附近位置的顏色變化運算。

Color	Hue	Value/Chroma
Red (R)	5R	5/14
Yellow (Y)	7.5Y	9/10
Green (G)	2.5G	5/10
Blue (B)	10B	6/10
Purple (P)	5P	5/10
Yellow-red (YR)	5YR	7/12
Green-yellow (GY)	2.5GY	8/10
Blue-green (BG)	5BG	7/8
Purple-blue (PB)	7.5PB	5/12
Red-purple (RP)	10RP	4/12
White (W)	N/9	
Gray (Gray)	N/5	
Black (Blk)	N/1	

表 2.6 Naz Kaya 使用的受測顏色一覽[14]

emotion	R	Y	G	B	P	YR	GY	BG	PB	RP	W	Gray	Blk
Angry	28	0	0	0	0	0	0	0	0	0	0	3	7
Annoyed	0	0	0	0	0	5	8	7	0	2	0	0	0
Bored	0	0	0	0	5	4	2	0	0	2	5	14	0
Calm	4	0	29	60	28	0	0	16	38	13	8	5	0
Disgusted	0	0	0	0	0	9	26	2	0	3	0	0	0
Excited	18	8	2	0	4	25	6	11	0	12	0	0	0
Fearful	0	0	0	0	5	0	0	0	0	0	0	3	17
Happy	21	74	28	10	21	31	11	36	13	26	0	0	0
Peaceful	0	0	12	4	0	0	0	0	8	0	13	0	0
Sad	4	0	0	8	13	0	0	0	10	0	0	30	24

表 2.7 情緒與顏色關係的統計數據，有網底數值為最多數[14]

在 Mapping emotion to color [15] 這篇論文中，作者 Niels A. Nijdam 將前述 Naz Kaya[14]與章節 2.3.2 James A. Russell[12]等人之研究做一整合，並且利用他所提出的顏色變化演算法，嘗試將情緒平面上的任一點都能轉化成帶有合適情緒的顏色，Niels A. Nijdam 採用的情緒位置為章節 2.3.2，圖 2.4 的 adult 實驗結果位置。

Niels A. Nijdam 的顏色演算法如下：

- a. 首先，E 為所有情緒的有限集合， $E = \{e, e', \dots\}$ ，建立一空集合 R^e 。
- b. e 為一情緒，定義 $e = \langle \text{coordinate}, \text{radius} \rangle$ ， $\text{coordinate} = \langle x, y \rangle$ 為情緒平面上的一點座標。
- c. 設 p 為平面上任一座標點，欲得到 p 的顏色值。
- d. 將 p 與 E 集合內所有情緒 e 比對，設 distance 為 p 和 e 的距離，若 $\text{distance} < \text{radius}_e$ 則將此 e 加入集合 R^e 。
- e. 對所有 R^e 內的 e(情緒)做以下步驟。
- f. 定義 cfinal，初始值為中立色。
- g. 定義 cdif 為 cfinal 和 e 的 RGB 值差值。例如一個情緒 e 的 RGB 值為 $\langle 157, 255, 128 \rangle$ 且 cfinal 為 $\langle 128, 128, 128 \rangle$ ，則 cdif 為 $\langle 157, 255, 128 \rangle - \langle 128, 128, 128 \rangle = \langle 29, 127, 0 \rangle$ 。
- h. 定義 $\text{intensity} = 1 - \text{distance}/\text{radius}$ ，distance 為 p 和 e 的距離，radius 為該 emotion 的 radius。
- i. p 最後的顏色值為：

$$\text{cfinal}(R) = \text{cfinal}(R) + \text{difference}(R) \times \text{intensity}$$

$$\text{cfinal}(G) = \text{cfinal}(G) + \text{difference}(G) \times \text{intensity}$$

$$\text{cfinal}(B) = \text{cfinal}(B) + \text{difference}(B) \times \text{intensity}$$

2.6 遊戲場景與情緒：EVG

EVG[16]，為 Emotion Evoking Game 的縮寫，有就是含有情緒感知的遊戲。Ning Wang 等人將一款 open source 的遊戲 Egoboo[17]，修改成數個 Ning Wang 想帶給玩家某些情緒感知的遊戲場景，方式是讓受測者遊玩後的問卷及遊玩時的臉像側拍，觀察玩家在每個遊戲場景的情緒感受為何。其場景設計如下，遊戲目標是要蒐集滿 2000 金幣。

Stage 1：玩家扮演一個劍士角色，身旁跟著一電腦設定的隊友，有數個小房間，小房間彼此之間只有單向迴廊相連，房間內有寶箱，玩家穿梭在這幾個小房間內，最後總共可以蒐集到 1750 金幣，期間沒有任何敵人出現。此階段的情緒感知設定為 boredom。

Stage 2：玩家走到最後一個小房間，房間內忽然出現魔王及小兵。此階段的情緒感知設定為 surprise。

Stage 3：玩家打贏魔王後掉落 1000 金幣。此階段的情緒感知設定為 joy。

Stage 4：正當玩家想去拿這 1000 金幣以達成目標，此時隊友忽然攻擊自己。此階段的情緒感知設定為 anger。

Stage 5：隊友殺掉玩家，並拿走所有金幣，結果隊友勝利，玩家宣告失敗，遊戲結束。此階段的情緒感知設定為 disappointment。

其實驗結果，多數玩家在最後的問卷及遊戲中的臉像側拍，表現出與設計者想帶給玩家的情緒感知接近，本論文的遊戲場景部分將參考 EVG 的方式建構。

2.7 節奏複雜度 - LHL

節奏的複雜度是用來測量人類對於感知節奏程度的方法，而其中 LHL 為計算節奏的切分音的感知程度，複雜度越高，切分音越多，則聽起來越複雜。

LHL 節奏分析法是由 Longuet-Higgins 和 Lee 提出[18]。LHL 使用樹狀層次來表示各個音符的權重，首先將一小節分成最小單位的數量等分，例如 4/4 的節奏用 16 分音符當最小單位，則分成 16 等分，將 16 質因數分解，得到 $2*2*2*2$ ，利用此質因數分解數列建構成樹狀圖；接著在各最小等分處標出權重，如圖 2.11 所示，0 處為權重最大，-4 為最小。

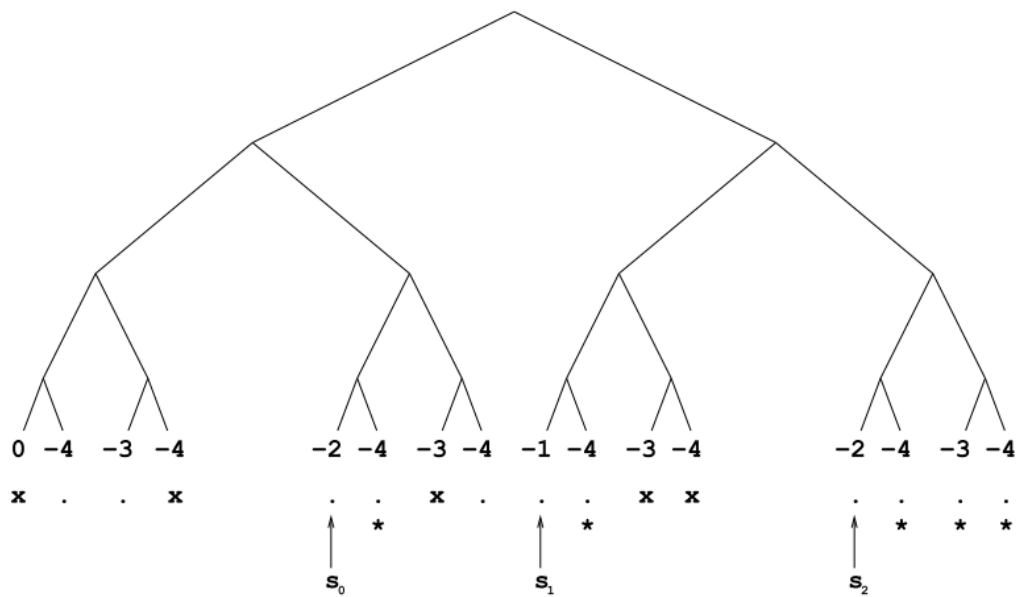


圖 2.11 LHL 節奏複雜度之樹狀圖及分析例子

權重標示完後找出 Onset(x)和 Rest(.)的位置，Onset 代表音符出現的地方，Rest 代表音符的延音部分或是休止部分。計算 Rest(.)減去其前面的 Onset(x)的數值，若大於 0 代表該位置為切分音的地方。舉例來說，圖 2.11 的 s_0 位置為 $-2 - (-4) = 2 > 0$ ， s_1 位置為 $-1 - (-3) = 2 > 0$ ， s_2 位置為 $-2 - (-4) = 2 > 0$ ，代表這三個地方均為切分音，此小節複雜度大小為三個位置的值相加， $2 + 2 + 2 = 6$ 。

第三章、實作方法

本論文之系統實作主要分為兩大部分，第一部分為創建一個自動作曲系統，可在情緒平面上畫線後，產生一首隨著線段行進及其在情緒平面上的位置，自動產生出一首含有該位置情緒感知的音樂。

第二部分為設計數個 Emotion Evoking Game 的遊戲場景，讓受測者玩此遊戲時能夠藉由場景，獲得遊戲設計者想激發之情緒感知。而遊戲場景的背景色，由顏色轉換系統獲取第一部分之音樂線段在情緒平面上的位置點轉換而來，同樣的，顏色也能夠激發受測者的情緒感知。

3.1 系統架構

圖 3.1 為本論文的自動作曲系統流程圖，流程說明如下：

Step 1 :

為獲取作曲資訊階段。系統將提供一個情緒平面介面供使用者畫線，並可設定該線段使用之點數個數(系統程式內一個線段是由許多點數所連結畫成)，一點代表一小節，使用的點數將平均的從線段上取得；也就是說當使用者在系統介面上畫一線段，若設定點數較多，則作曲取用點數資訊較多，小節較多、音樂長度較長，因為點數距離近，樂曲內的音樂特徵則變化和緩；若設定點數很少，則取用點數平均分散於線段上，小節少、音樂長度較短、點數距離遠，音樂特徵變化較劇烈。

Step 2 :

為自動作曲進行階段。將以 Patrick Gomez 及 Brigitta Danuser [10][11]的研究為基礎，把各種音樂特徵如節奏、速度、pitch range 等建構其與情緒平面的關聯方程式，產生情緒音樂，另外本系統均使用調性音樂，以一次建立一小節進行。

Step 3 :

另外依據線段的點數資訊，經由本系統設計之顏色演算法，產生一個顏色中繼資訊檔案，使用 XML[20]格式儲存，方便遊戲部分的場景顏色使用。

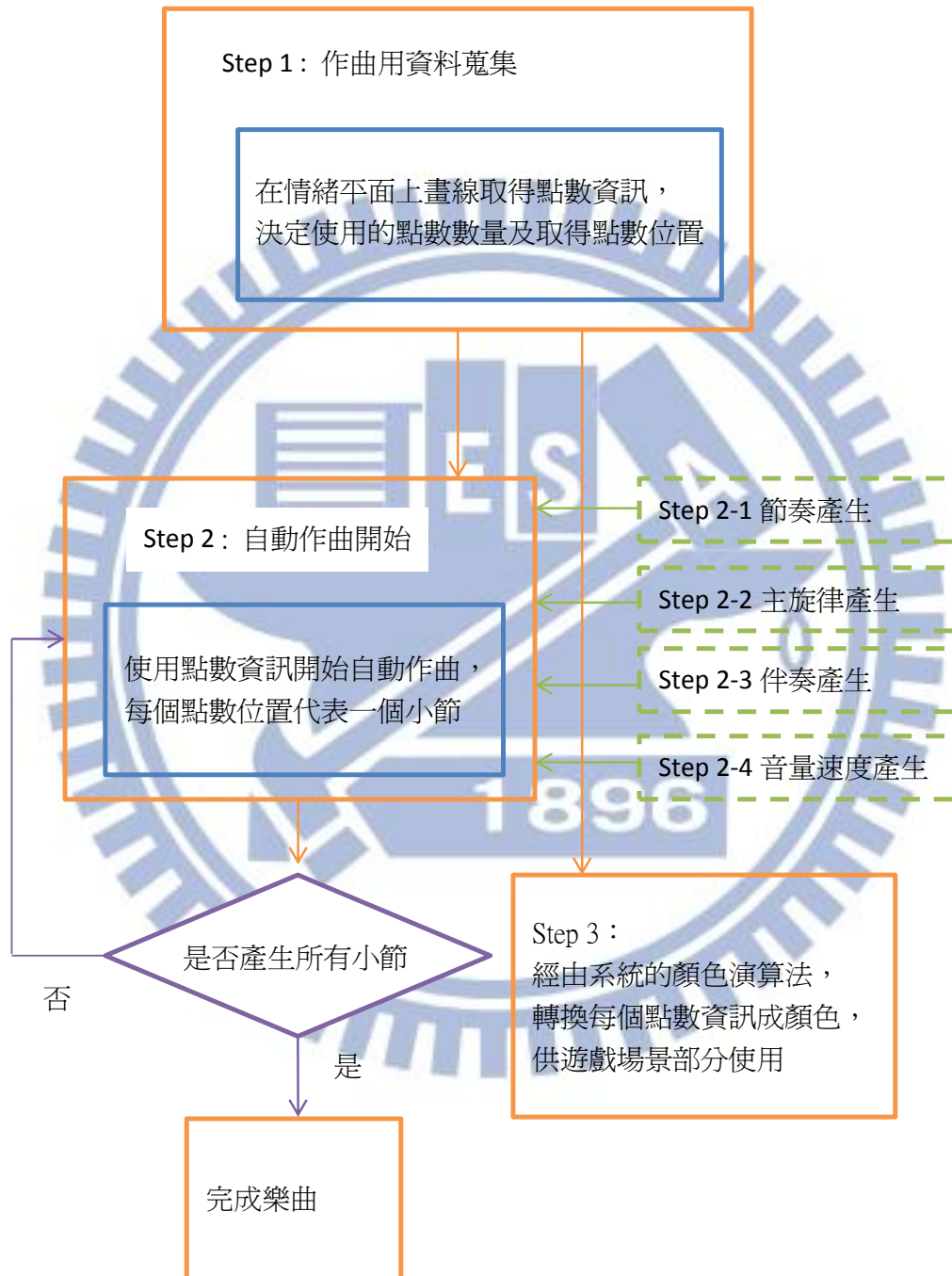


圖 3.1 情緒自動作曲系統流程圖

3.2 與情緒平面連結

本論文系統使用 Patrick Gomez 及 Brigitta Danuser [10][11]的研究中，其中四個音樂特徵與情緒平面間的回歸公式為基礎來產生情緒音樂，使用的四個情緒特徵為 Sound intensity、Tempo、Rhythm、Pitch range(見圖 2.7)。表 3.1 為該四個情緒特徵於 Valence - Arousal 情緒平面上的哪點為最高/最低數值。

音樂特徵	最高值(位置：(V,A))	最低值(位置：(V,A))
Sound intensity	76dB (1,9)	58.9dB (9,1)
Tempo	5.959 分 (9,9)	1.73 分 (1,1)
Rhythm	5.084 分 (1,9)	2.69 分 (1,1)
Pitch range	5.263 分 (1,9)	2.71 分 (1,1)

表 3.1 四種音樂特徵的最高最低評分值及其位置

3.3 節奏產生

本系統採用 4/4 拍的節奏，每小節最小的單位為十六分之一音符，並以 LHL 節奏分析法的逆向推算法，來產生音樂的節奏，共能產生 0~11 的節奏複雜度。首先把將要進行節奏產生的小節資訊點(即繪製在情緒平面上的線段的某一點)，利用節奏與情緒平面的轉化公式，取得該點的節奏分數(表 2.4)：

$$\text{節奏分數} = 2.23 + 0.16 \times \text{Valence} + 0.35 \times \text{Arousal} - 0.05 \times \text{Valence} \times \text{Arousal}$$

取得節奏分數後，將其最小最大值區間調整放大至 0~11，小數點四捨五入取整數值：

$$\text{Complexity} = (\text{節奏分數} - 2.69) \times 11 / (5.084 - 2.69)$$

得到 Complexity 值(0~11)後，進行 LHL 逆向推算法：

見圖 3.2，”x”表示 Onset 的位置，”.”表示 Rest 的位置，可能為休止符或是 x 位置的延音；英文字母表示有複雜度出現的地方，即”.”的權重值減去”x”權重值大於 0 的地方，例如在位置 b 有複雜度的出現，complexity 值為 $-2 - (-4) = 2 > 0$ 。

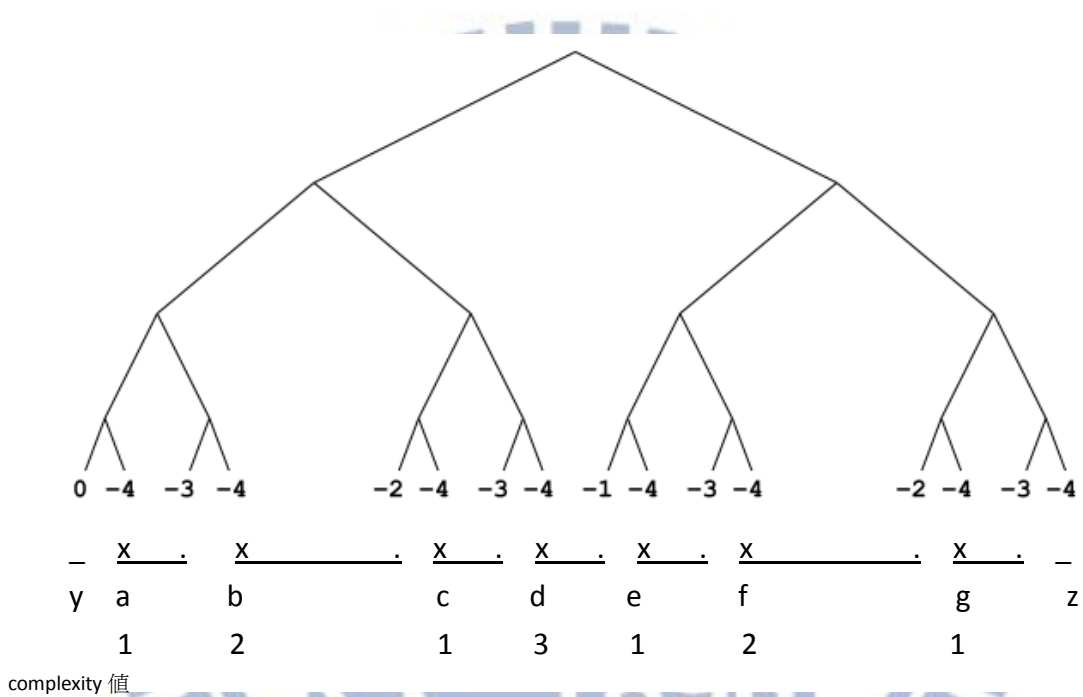


圖 3.2 LHL 逆向法示意圖

定義 a ~ g 七個位置的 complexity 值後，即可依照所需的 complexity 值組合，例如需要產生 complexity 為 5 時，組合的方式有 4 種，若使用 2+3 的方式，則是在 b、f 這兩個 complexity 為 2 的位置選一個，再選 complexity 為 3 的 d。

另外還可加音但不影響複雜度，增加其變化，圖 3.2 首尾 y、z 的位置均可選或不選，共 $2*2 = 4$ 種組合，而若加在 y、z 以外的位置，則必須注意以下條件：

若隨機在 Onset 的地方(即 a ~ g 的位置)，則和它一組的 ”.” 也必須選到，以免增加複雜度；而若隨機在 ”.” 的地方，則先檢查和它一組的 “x” 是否為 Onset，

若為 Onset，則該 ”.” 的位置不可以 Onset，以免減少複雜度，若和它一組的 ”x” 並無 Onset，則可以加入 Onset。最後其他未 Onset 的位置可以當作休止符或是當作其前一個 Onset 的延音。

3.4 旋律產生

得到了一個小節內的節奏後，系統即可開始展開主旋律的創作，本論文建構之情緒作曲系統為有調音樂，旋律使用大調自然音階以及小調和聲音階的音組成，在繪製線段的第一點，若在 Valence 為正向則使用大調創作，12 個大調隨機選擇、負向則用小調、12 個小調隨機選擇，也就是以 Arousal 軸區分大小調，而若繪製之曲線有通過 Arousal 軸，則樂曲調性轉移成先前調號的關係大小調。

旋律產生使用的大小調音階內音，是隨機但依照其權重[19]比例產生，如表 3.2，而若產生音樂的調號非 C 調時則需做移調的動作。例如音樂為 G 大調，則 G 大調內的 G 需要視為權重對應表內的 C 音，權重 6.35。產生方式為將大或小調的所有權重值加起來，例如大調共 41.79，其中 C 音 6.35，占了 15.2%的權重，因此該大調的”C”音在隨機產生時有 15.2%機會選取到。

調	C	C#	D	D#	E	F
大調	6.35	2.23	3.48	2.33	4.38	4.09
小調	6.33	2.68	3.52	5.38	2.6	3.53
調	F#	G	G#	A	A#	B
大調	2.52	5.19	2.39	3.66	2.29	2.88
小調	2.54	4.75	3.98	2.69	3.34	3.17

表 3.2 大調與小調之調性權重對應表[19]

一個大或小調的主音”C”，將其視為 0，則系統的旋律產生採用的音階內音，一個八度大調自然音階為 {0,2,4,5,7,9,11}，小調和聲音階為 {0,2,3,5,7,8,11}，再加上調號位移及本作曲系統的音高從 C4 (midi 以數值 60 代表其音高) 為基礎延伸，例如系統若採用 D 大調音階，則使用了 $60 + 2$ (調號位移) + {0,2,4,5,7,9,11}，為系統使用在 midi 內表示的音高值。

本作曲系統不只使用了一個 8 度內的音階，參考了 Pitch range 這個音樂特徵，在欲產生的音樂中，Pitch range 最窄為該調主音的 -5 ~ 11，也就是低一個八度的 G 音至現在這個八度的 B 音，最寬為 -5 ~ 28，為低一個八度的 G 音至高兩個八度的 E 音，先求得 Pitch range 分數為：

$$2.17 + 0.22 \times \text{Valence} + 0.37 \times \text{Arousal} - 0.05 \times \text{Valence} \times \text{Arousal}$$

擴展調整成所需的 pitch range 範圍公式：

$$\text{Pitch Range 最小值} = -5$$

$$\text{Pitch Range 最大值} = 11 + (\text{Pitch range 分數} - 2.71) \times 17 / (5.263 - 2.71)$$

而一個小節之中，音與音之間的音程關係也參考 pitch range 的分數，越低則前後音不能跳動太大，越高分則前後音的音程較能夠大跳，前後音最低音程幅度的上限從 4 開始，隨著 pitch range 增加，最高能到 12，也就是能跳一個八度音。

跳躍允許音程幅度為：

$$4 + (\text{Pitch range 分數} - 2.71) \times 8 / (5.263 - 2.71)$$

3.5 伴奏、音量強度、速度

本作曲系統的音樂和聲行進方式為 $I \rightarrow IV \rightarrow V \rightarrow I$ ，為正格終止，系統以四小節為單位，依據作曲系統進行到的小節數 $\text{mod } 4$ ，小節數 $\text{mod } 4 = 0$ 或 1 時則為 I 級和弦伴奏， $\text{mod } 4 = 2$ 時則為 IV 級和弦伴奏， $\text{mod } 4 = 3$ 時則為 V 級和弦伴奏。在小節所屬的情緒平面位置，若是位置在 Valence 軸的上方，則使用輕快活潑的伴奏，五度音行進，若在 Valence 軸下方，則為分解和弦，圖 3.3 顯示以 C 大調為例，輕快活潑的伴奏及分解和弦。

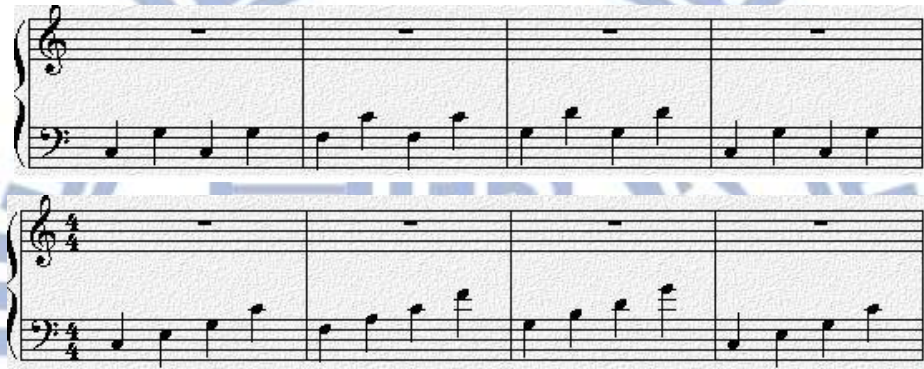


圖 3.3 伴奏範例，上為小節資訊點在 valence 軸之上時使用的伴奏，下為在 valence 軸之下使用的伴奏。

另外，在伴奏音 onset 的拍點上，若主旋律也同時有 onset 的音在，則該主旋律音須為半奏和弦的內音，使聽覺感受不會有不諧和感。

音量強度(dB) 則直接使用表 2.4 公式：

$$\text{Sound intensity} = 58.4 + 2.14 \times \text{Arousal} - 0.18 \times \text{Valence} \times \text{Arousal}$$

速度(Tempo) 則由公式 $1.2 + 0.1 \times \text{Valence} + 0.43 \times \text{Arousal}$ 產生，其分數最高為 5.95，最低為 1.73，而 [10][11]所採用的 16 首曲子，樂曲速度最高約為 130，最慢約為 35，因此將分數用線性方程式 mapping 到 bpm 35~130 的速度：

$$\text{實際速度值(bpm)} = 20.8696 \times \text{Tempo 分數} - 1.4783$$

3.6 顏色轉換演算法

在 2.5.3 章節中，Niels A. Nijdam [15] 是使用在平面上的任一點與某情緒的位置及其影響半徑來決定強度，進而轉換成顏色，這樣的缺點是沒辦法將情緒平面上，某一情緒的程度變化完整的反應在顏色。例如某一點在原點和所屬情緒的中間，另一點在所屬情緒的外側，這兩點和所屬情緒的距離相同，則依據 Niels A. Nijdam 的演算法，此兩點的顏色卻將是雷同的。因此本論文提出另一新的顏色轉換演算法，能夠把情緒平面上任一點，適當且有程度變化的轉換成顏色。

圖 3.4 為情緒平面上所有點轉換後的顏色，各個情緒的位置分布參考[6]、[12] 加以綜合，包含與原點的距離，定位到情緒平面上。而情緒該使用甚麼顏色代表，則參考了 Naz Kaya 的顏色與情緒相關實驗[14]，最具代表該顏色的情緒(投票數最多)如表 3.3。每個情緒所在的位置使用角度、與原點距離來表示，如表 3.4。其中位置角度代表從 Valence 軸線正向開始為 0 度，逆時針計算；離中心點距離為 hA 的倍數，hA 代表 half of Axis，也就是軸線一半長；所佔範圍為該情緒所包含的範圍共佔幾度。

顏色(Munsell Syetem)	最具代表情緒	轉換成 RGB 數值
Red (5R 5/14)	Anger	(222, 64, 79)
Yellow (7.8Y 9/10)	Happy	(255, 230, 85)
Green (2.5G 5/10)	Peaceful	(0, 142, 83)
Blue (10B 6/10)	Calm	(31, 158, 223)
Yellow -red (5YR 7/12)	Excited	(255, 149, 59)
Green-Yellow (2.5GY 8/10)	Annoyed, Disgusted	(204, 210, 63)
Gray (N/5)	Sadness, Bored	(143, 143, 143)
Black (N/1)	Fear	(0, 0, 0)

表 3.3 各顏色 RGB 值與其最具代表情緒[14]

情緒	位置角度	離中心點距離, $1hA = \text{half of axis}$	範圍度數
Happy	22.5	$3/3.5 \times hA$	45
Excited	67.5	$3/3.5 \times hA$	45
Annoyed	105	$3/3.5 \times hA$	30
Anger	135	$2.5/3.5 \times hA$	30
Fear	165	$1.5/3.5 \times hA$	30
Disgusted,	191.25	$2.5/3.5 \times hA$	22.5
Sadness	236.25	hA	67.5
Bored	277.5	hA	15
Calm	307.5	$3/3.5 \times hA$	45
Peaceful	345	$3/3.5 \times hA$	30

表 3.4 各情緒於情緒平面上的位置與其範圍[14]



圖 3.4 本系統之顏色轉換演算法，所有點的顏色轉換結果

本系統之顏色轉換演算法如下：

- a. 若想轉換一點 p 座標的顏色，則先將 p 點之角度算出，為 pAngle。
- b. 設情緒集合 $E = \{e, e', \dots\}$ ，其中 e 為某一情緒，每個 e 有以下資訊：
<位置角度 eAngle，與中心距離 eDis，範圍所占角度 eRangeA，
顏色值(eR, eG, eB)>。

- c. 將 p 對集合 E 內的所有 e 比對，若 pAngle 角度滿足

$$eAngle - \frac{eRangeA}{2} \leq pAngle < eAngle + \frac{eRangeA}{2}$$

，則判斷 p 屬於 e 情緒。

- d. 首先設 pDis 為 p 點跟中心的距離，設 scaleDis = pDis/eDis，而若 e 情緒的(eR, eG, eB)值個別 \times scaleDis 後，R, G, B 中任一值超過 255，則將 scaleDis 重設縮小到使 R, G, B 某一值恰好能到 255 即停止。
- e. 例如 Happy 位置為 22.5 度，範圍 45 度，則 Happy 在情緒平面上的範圍角度值為 $22.5 - 45/2 \sim 22.5 + 45/2$ ，也就是 0~45 度。計算 p 點與 e 位置的角度差，設定 p 若在 e 情緒範圍角度的最兩端，舉例來說，若 p 點落於 happy 範圍，最兩端值即 pAngle 為 0 或 45 度，則設 scaleAngle 最小值 0.6，scaleAngle 值為 0.6~1，算法為：

$$scaleAngle = 1 - \left(\frac{|pAngle - eAngle|}{\frac{eRangeA}{2}} \right) \times 0.4$$

- f. 設 finalScale = scaleDis \times scaleAngle，取得最後的顏色調整值。
- g. 設原點為中立色，RGB = (173, 178, 178)[15]，p 的顏色值(pR, pG, pB)為：

$$pR = 173 + (eR - 173) * finalScale$$

$$pG = 178 + (eG - 178) * finalScale$$

$$pB = 178 + (eB - 178) * finalScale$$

3.7 XML 與 Processing

XML[20]：eXtensible Markup Language，是一種標記語言，XML 設計用來傳送及攜帶數據信息，並定義結構、存儲信息、傳送信息。本論文將顏色轉換系統產生的顏色資訊、樂曲時間資訊用 XML 儲存成一個中介資訊檔，方便遊戲場景使用這些資訊，如圖 3.5，Time 標籤為每小節的時間，RGB 為轉換出來的顏色，Tempo 為速度。圖 3.6 為產生中介資訊檔的流程示意圖。

```
- <SongTimeRGB>
  <BarNum>12</BarNum>
  <TimeRGB Time="0.1" Red="45.0" Green="161.0" Blue="220.0" Tempo="51.12144" />
  <TimeRGB Time="4.6947036" Red="39.0" Green="159.0" Blue="221.0" Tempo="52.064747" />
  <TimeRGB Time="9.304348" Red="32.0" Green="158.0" Blue="222.0" Tempo="53.22927" />
  <TimeRGB Time="13.813146" Red="24.0" Green="156.0" Blue="223.0" Tempo="55.057446" />
  <TimeRGB Time="18.17223" Red="18.0" Green="155.0" Blue="224.0" Tempo="56.622665" />
  <TimeRGB Time="22.410814" Red="21.0" Green="155.0" Blue="224.0" Tempo="58.505104" />
  <TimeRGB Time="26.513021" Red="25.0" Green="156.0" Blue="223.0" Tempo="61.2432" />
  <TimeRGB Time="30.431824" Red="29.0" Green="157.0" Blue="223.0" Tempo="63.95207" />
  <TimeRGB Time="34.184635" Red="34.0" Green="158.0" Blue="222.0" Tempo="67.600075" />
  <TimeRGB Time="37.734924" Red="15.0" Green="146.0" Blue="92.0" Tempo="70.68043" />
  <TimeRGB Time="41.13049" Red="9.0" Green="144.0" Blue="88.0" Tempo="73.67731" />
  <TimeRGB Time="44.38794" Red="0.0" Green="142.0" Blue="83.0" Tempo="77.55905" />
  <EndTime>47.482357</EndTime>
</SongTimeRGB>
```

圖 3.5 以 XML 格式存取系統建立出來的音樂時間、顏色、速度等資訊

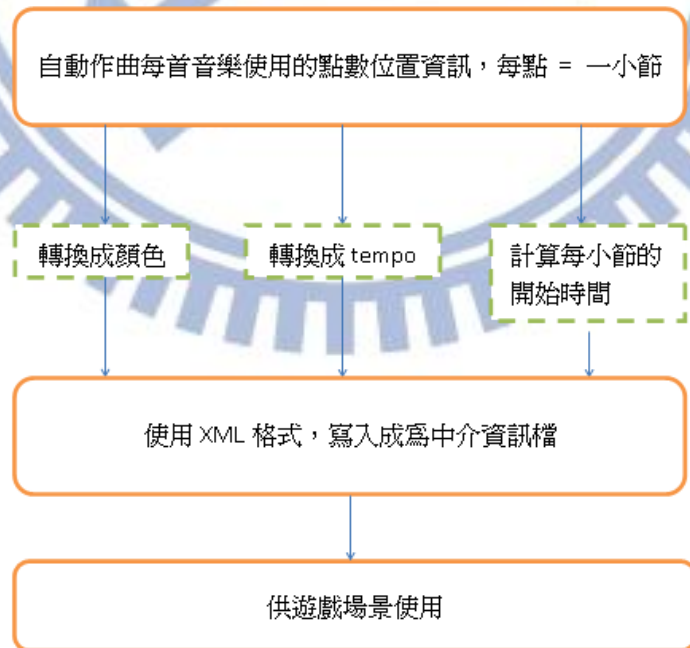
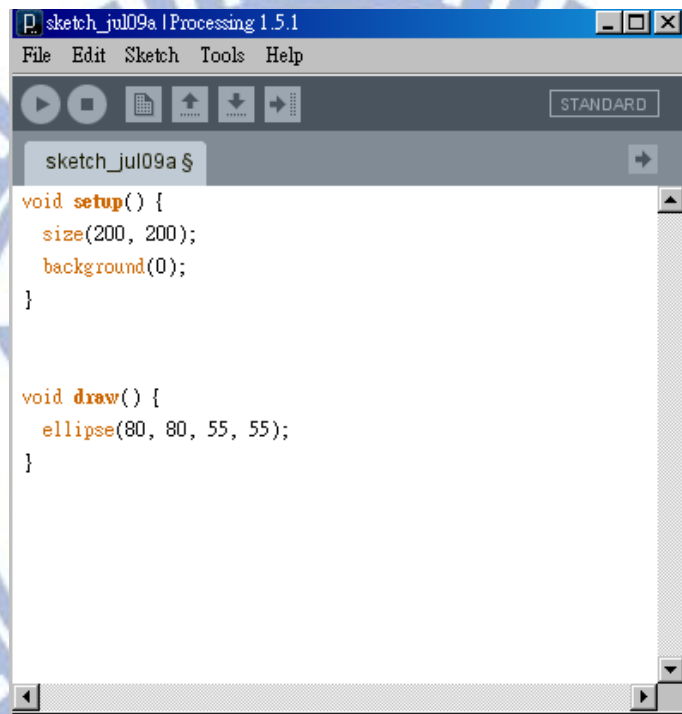


圖 3.6 中介資訊檔產生流程

Processing[21] 為一個能夠簡單撰寫視覺、圖像化程式的軟體，使用 Java 為基礎開發，且其中運作原理為不斷執行一個繪圖用函式，因此很適合用來製作含有流程情節的圖像化遊戲，本論文所製作的遊戲場景將使用 processing 軟體撰寫。一個簡易的程式說明及執行結果如圖 3.7 所示。圖 3.7a 為一簡單的程式，目的為畫出一個圓形，在 `setup()` 函式內為基本設定，內容為設定畫布大小 200*200 (單位為 pixel)，背景色為 0 也就是黑色。程式會一直執行繪圖函式 `draw()`，也就是一直畫圈在畫布上 80,80 的位置，圓的長寬為 55,55。圖 3.7b 為執行之結果。



```
sketch_jul09a $
void setup() {
  size(200, 200);
  background(0);
}

void draw() {
  ellipse(80, 80, 55, 55);
}
```

圖 3.7a processing 簡易程式範例

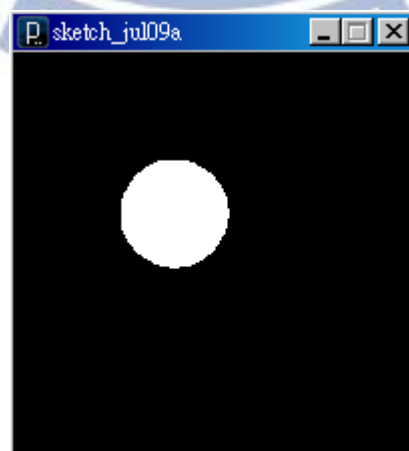


圖 3.7b 執行之結果

3.8 遊戲場景設計

本論文參考 EVG 遊戲場景設計，設計了 4 個場景。遊戲方式，玩家扮演一主角飛機，目標為 1000 金幣。

場景 1：設計帶給玩家在情緒象限 4 的感覺，也就是正向 Valence 和低程度的 Arousal。此場景階段內容沒有任何敵人，隨音樂時間，畫面會隨機出現數個 10 金幣，玩家可移動鍵盤方向鍵去獲取，此階段情緒設定為平靜。

場景 2：設計帶給玩家在情緒象限 1 的感覺，也就是正向 Valence 和高程度的 Arousal。此場景階段會隨機出現少量敵人，玩家可用空白鍵發射子彈，消滅一個敵人後能得到 100 金幣，而在場景後段，將加入一個電腦夥伴說要幫助玩家。此階段情緒設定為：興奮、開心。

場景 3：設計帶給玩家在情緒象限 2 的感覺，也就是負向 Valence 和高程度的 Arousal。此場景出現大量敵人，且較場景 2 難應付，中段出現魔王。在消滅魔王後，夥伴忽然攻擊玩家。此階段情緒設定為：緊張、惱人、生氣。

場景 4：設計帶給玩家在情緒象限 3 的感覺，也就是負向 Valence 和低程度的 Arousal。此場景玩家被攻擊而死亡，夥伴偷走所有的金幣而獲勝，畫面出現墓碑，玩家宣告失敗。此階段情緒設定為：傷心、難過。

第四章、 系統展示

本章節將以圖文說明本論文產生之自動作曲系統，以及遊戲場景部分的圖像展示。圖 4.1 為操作流程圖，隨後將系統的功能個別詳細介紹。



圖 4.1 系統操作流程

4.1 自動作曲系統基本介面介紹

圖 4.2 為作曲系統的介面圖，可以看到左邊可以畫線的區域，與情緒平面的形式展現，使用滑鼠點擊拖曳即可畫出線條，線條的 Start 處為音樂開始位置。當左方已有畫上的線段而想要重新作畫時，則按右邊 clear curve 即可。

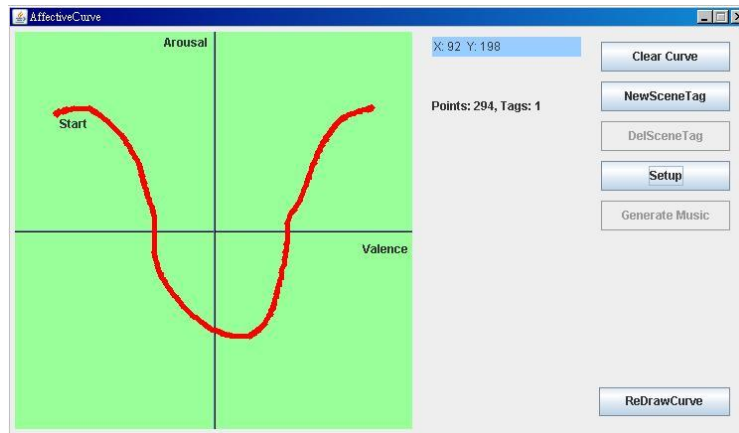


圖 4.2 系統介面圖

再來可以按下右方 setup 進行小節數的設定，按下 setup 後跳出視窗如圖 4.3：

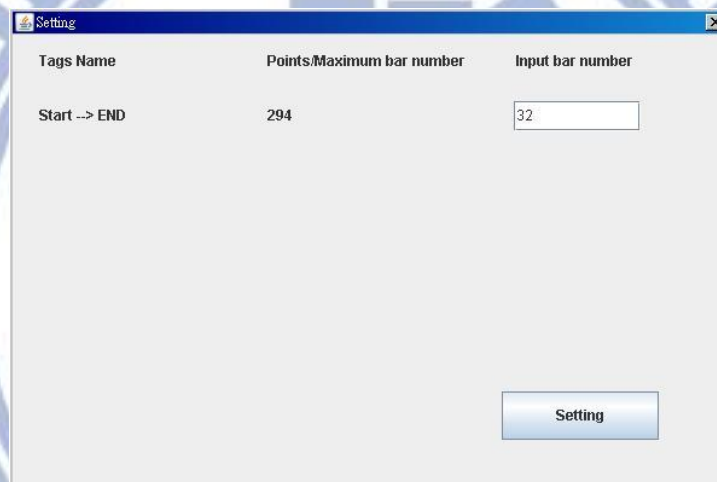


圖 4.3 按下 setup 鍵後的視窗

可以看到 setup 視窗有 points/maximum bar 的資訊，以圖例來說，這條線目前含有 294 個點，最多可以產生 294 個小節的音樂，在 input bar 欄則可設定想要的小節數，假設輸入 32 來看。設定完成後原本的線段上會出現平均分布的紅點，則為使用的資訊點，這些點每一個將用來產生一個小節的音樂，故稱其為小節資訊點(圖 4.3)，最後按 generate music 即可產生音樂 midi 檔和之後遊戲場景將用到的 XML 資訊文件檔。音樂及資訊文件檔的內容資訊均是從 start 開始，從在線段上的紅點依序產生。

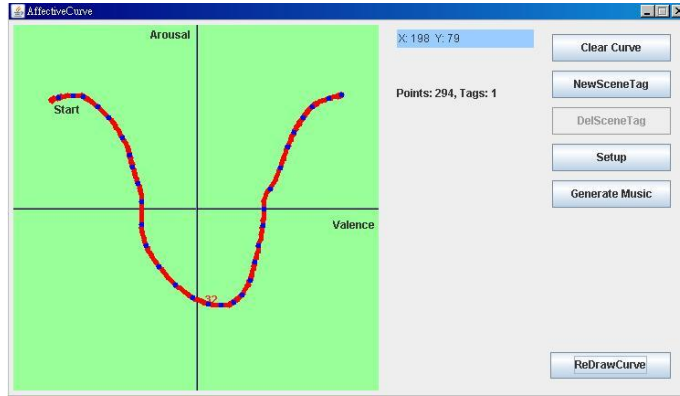


圖 4.4 設定小節數後的線段圖

4.2 加入分段 tag

本系統除了畫線直接產生音樂外，還可以加入線段的分段 tag，也就是說，同一線段可以藉由加入多個 tag，產生好幾首音樂，相當適合遊戲場景的分段設計。點選右方的 newScene tag 即可在線段上增加新的分段 tag。例如圖 4.5 的線段加入 2 個新的 tag，則整個線段被分成三個部分。

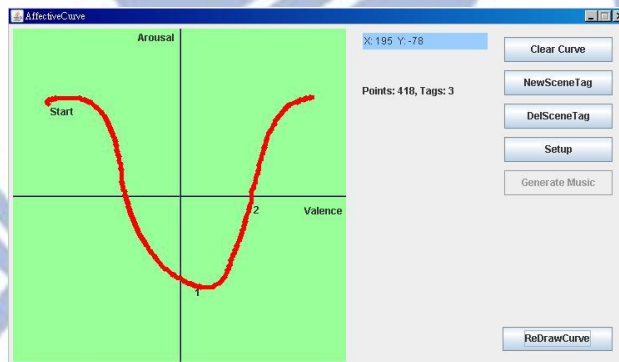


圖 4.5 加入兩個新 tag 的線段

Tags Name	Points:Maximum bar number	Input bar number
Start -> 1	246	24
1 -> 2	78	8
2 -> END	94	12

圖 4.6 加入兩個新 tag 後的 setup 視窗

按下 setup 後如圖 4.6，則可以輸入三首樂曲的小節選擇，輸入 24、8、12 後，如圖 4.7，各線段使用的資訊點分布密度不同，start ~ tag 1 只用了 24 個點，tag1 ~ tag2 用了 8 點，tag2 到線段結束用了 12 點；也就是可以依照喜好產生變化劇烈程度不同的音樂。最後按下 generate music 則可同時產生這三首 midi 音樂及其 XML 資訊檔。

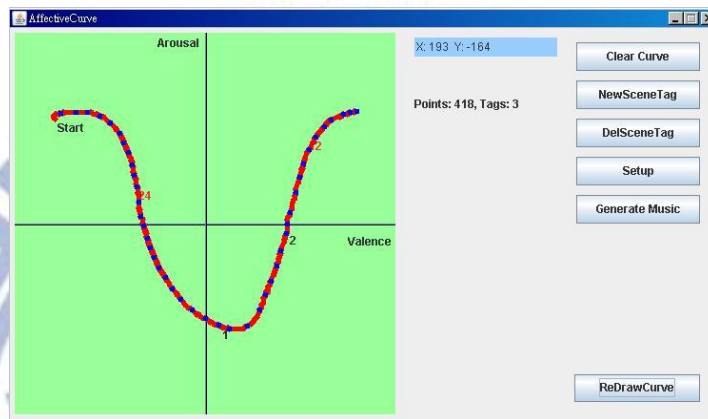


圖 4.7 設定好每個分段小節使用點數

本論文的實驗，遊戲場景的配樂，是用此介面繪製一條類似環繞四個象限的大圈，然後使用 tag 功能，產生四首配樂，如圖 4.8，從 start 到 tag “2”之前為第一首歌使用的線段資訊，tag “2”至 tag “3”為第二首歌所使用，以此類推。

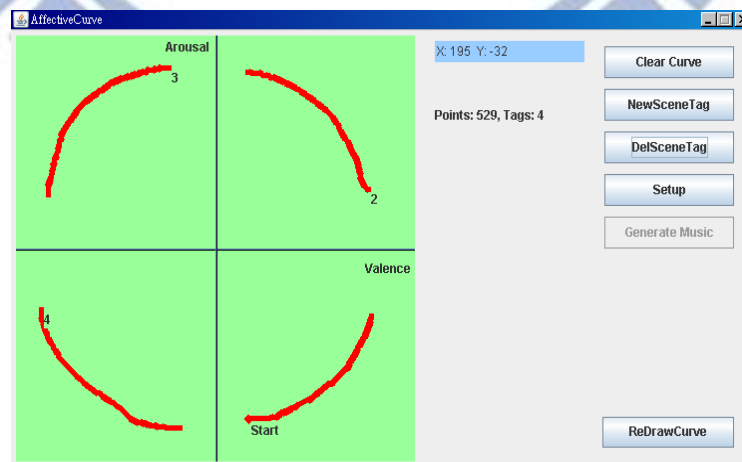


圖 4.8 遊戲場景使用之配樂的原始繪製線段，start、2、3、4 共四場景

4.3 遊戲場景畫面展示

圖 4.9 為設計之四個遊戲場景的片段畫面，遊戲使用 processing 程式製作。

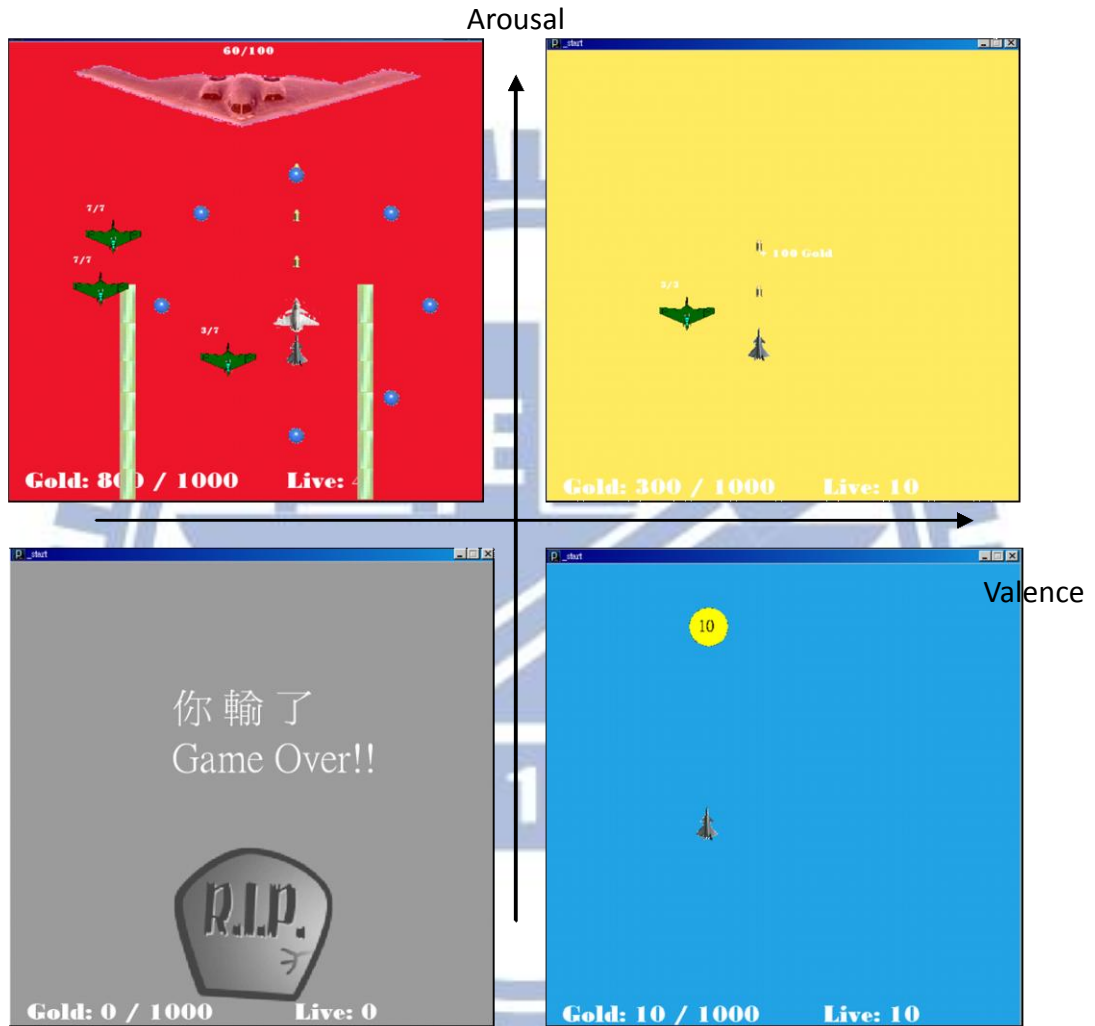


圖 4.9 所設計之四個遊戲場景畫面

遊戲場景 1~4 依序為圖 4.9 的第四象限、第一象限、第二象限、第三象限，也就是各個遊戲場景的目的即為提供所屬象限的情緒感知給受測者。在遊戲介面中，左下方代表目前擁有金幣/達成目標總金幣，右下方為剩餘生命，若生命歸 0 則無法動作 5 秒，也相對減緩得到金幣的速度。

第五章、 實驗

5.1 實驗目的及步驟

本論文之實驗目的將會試著驗證：

1. 建構之情緒自動作曲系統是否能產生出符合對應情緒的音樂。
2. 將樂曲加入以 EVG 方法建構的遊戲場景，試驗受測者情緒反應結果。
3. 再加入符合情緒的顏色至遊戲場景，試驗受測者情緒反應結果。
4. 以上三項實驗結果的相關性。

邀請的受測者總共 30 位，其中有 10 位受測者未受過音樂訓練，實驗會附上未受過音樂訓練的受測者實驗數據，藉以觀察與總體平均差異，驗證有無學習過音樂是否能對本系統之實驗有類似的情緒感受。在受過音樂訓練的 20 位受測者中，學習音樂超過十年以上者有 2 人，五至十年 3 人，三至五年 4 人，一至三年 3 人，一年以內 8 人，如圖 5.2 所示。

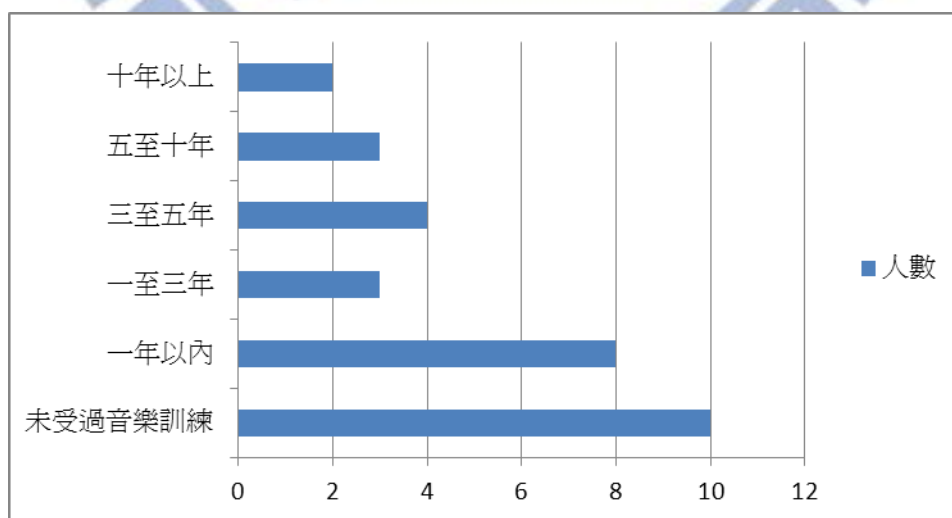


圖 5.1 受測者所受音樂訓練時間的比例圖

問卷將用以下步驟進行：

1. 受測者聆聽系統產生之 12 首情緒音樂(章節 5.2)，音樂排序編號為亂數，受測者聽完每首音樂後將於問卷答覆：
 - a. 感受為哪個象限的情緒。
 - b. 感受程度 (1~5 分，1 分最低，5 分最高)。
2. 休息 10 分鐘，說明接下來進行第一次遊戲的操作方式及遊戲過關目的。
3. 開始進行第一次遊戲(章節 5.3)，遊戲共有四個場景階段，而遊戲配樂使用該場景階段對應至合適情緒象限的音樂，且為該象限之中間線段的新產生音樂。每個場景階段結束時，填寫問卷：
 - a. 感受為哪個象限的情緒。
 - b. 感受程度 (1~5 分，1 分最低，5 分最高)。
 - c. 音樂帶來的情緒感受佔的比重分數 (1~5 分，1 分最低，5 分最高)。
 - d. 場景情節帶來的情緒感受佔的比重分數 (1~5 分，1 分最低，5 分最高)。
4. 第四個場景階段問卷填寫完後，休息 10 分鐘，準備進行第二次遊戲。
5. 開始進行第二次遊戲(章節 5.4)，遊戲共有四個場景階段，遊戲配樂產生方式如同第一次遊戲流程所述，在每個階段結束時，填寫問卷：
 - a. 感受為哪個象限的情緒。
 - b. 感受程度 (1~5 分，1 分最低，5 分最高)。
 - c. 音樂帶來的情緒感受佔的比重分數 (1~5 分，1 分最低，5 分最高)。
 - d. 場景情節帶來的情緒感受佔的比重分數 (1~5 分，1 分最低，5 分最高)。
 - e. 場景顏色帶來的情緒感受佔的比重分數 (1~5 分，1 分最低，5 分最高)。
6. 第四個場景階段問卷填寫完後，結束所有流程。

5.2 實驗一：12 首情緒音樂的問卷結果分析

首先，在作曲系統的情緒平面上繪製每個象限各三首，且三首平均分布於內(inner)、中(middle)、外(outer)；就是與中心原點的距離近、中、遠，總共 12 首曲子，每首曲子為 8~12 小節；換句話說，本階段實驗試著讓受測者不但能感到主要的四個象限情緒音樂差異，更可以在每個象限內不同的三首音樂中感受到類似情緒但程度不同的感覺，藉以驗證本系統的可信性。如圖 5.2 所示，12 個線段即是用來自動產生 12 首情緒音樂。



圖 5.2 實驗用 12 首歌採用的 12 個線段

依照圖 5.2 之線段產生好 12 首情緒音樂後，亂數編 1~12 號，受測者從 1 號開始依序聆聽及填寫問卷。

表 5.1~5.4 是所有受測者對於 12 首歌曲，每象限及該象限三首程度的歌曲的情緒反應統計表及其折線圖，表格字體含有網底代表正確選擇的受測人數：

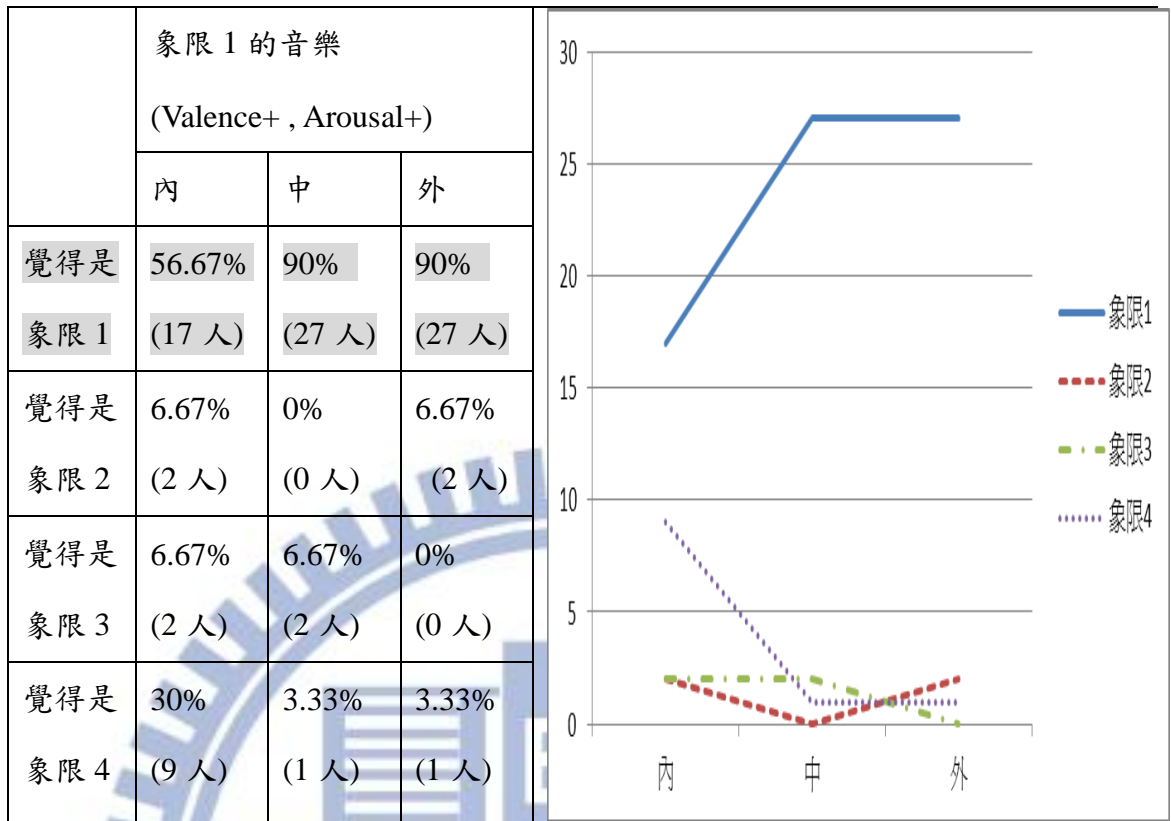


表 5.1 受測者對屬於象限 1 (V+ A+) 三首音樂的情緒反應統計表及其折線圖

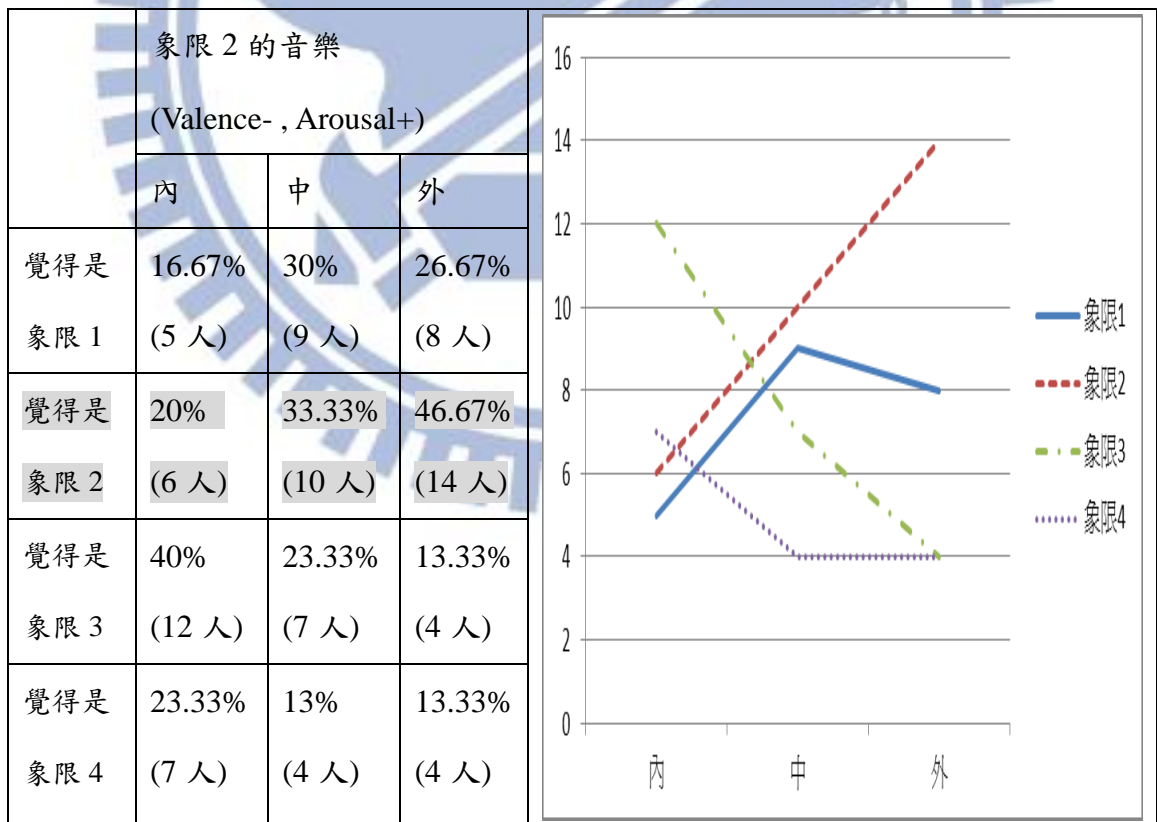


表 5.2 受測者對屬於象限 2 (V- A+) 三首音樂的情緒反應統計表及其折線圖

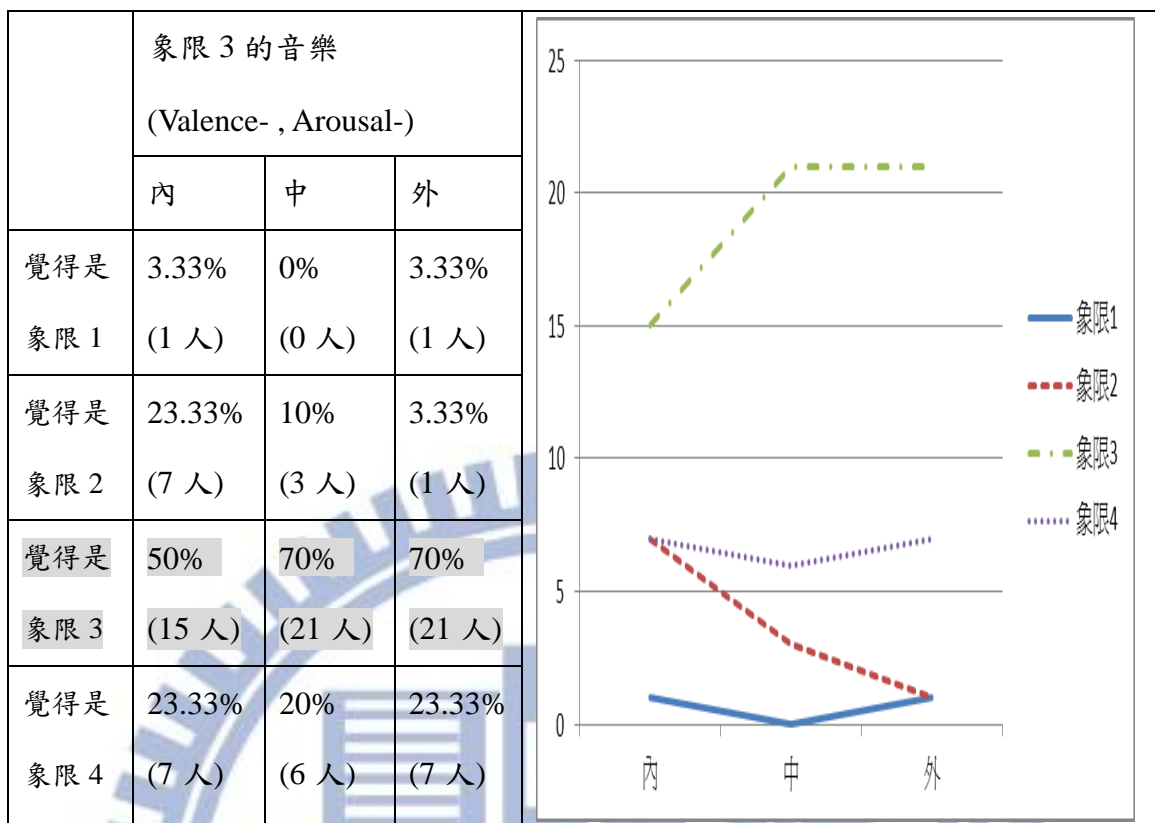


表 5.3 受測者對屬於象限 3 (V-A-) 三首音樂的情緒反應統計表及其折線圖

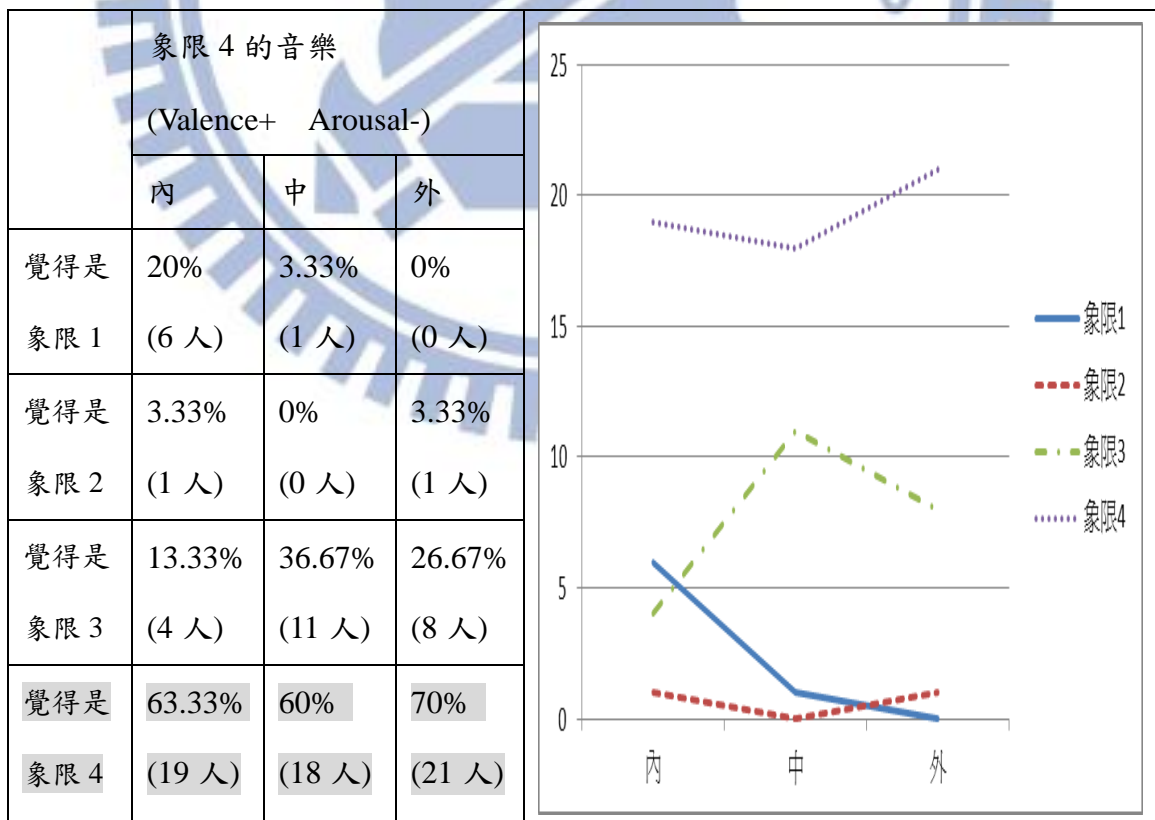


表 5.4 受測者對屬於象限 4 (V+A-) 三首音樂的情緒反應統計表及其折線圖

由以上四組圖表可以看出，象限 2 內圈的音樂與其他象限的內圈音樂令受測者較容易混淆，但隨著越外圍其辨識正確率有明顯改善；而其他 11 首音樂均能夠讓受測者有符合於該音樂應有的情緒感受；另外可以看到，四張折線圖中，各個象限音樂的正確回答人數率，大多隨著內、中、外圈而有漸增的趨勢，顯示即便在同一個象限內的情緒音樂，也能夠隨著離原點距離不同而有不同程度的感受，也就是越外圍則對該象限的情緒感知越清楚。

再來我們可以發現每一象限在內圈時，受測者除了正確答案之外，第二多的選擇為相對於正確象限的上或下象限，例如象限 1、4 內圈的音樂除了正確答案外最會被互相混淆；而象限 2、3 內圈的音樂亦然。

而到了每一象限的外圈音樂時，受測者除了正確答案之外，第二多的選擇變成了相對於正確象限的左或右象限，例如象限 1、2 外圈的音樂除了正確答案外最會被互相混淆；而象限 3、4 外圈的音樂亦然。

由於本系統自動作曲之音樂速度與 Arousal 最為正向線性相關，因此可以歸納出，相同的調性的兩個象限音樂(1、4 象限為大調、2、3 象限為小調) 在速度中等且相近時(內圈音樂)會有少數受測者混淆彼此，聽起來均感覺為正或反向情緒，但激動程度中等的音樂；而速度較極端(很快或很慢，外圈音樂)則有些許受測者可能將大小調混淆。

接著將同一象限的三首歌合併起來統計，較能夠完整代表某一象限能夠帶給受測者怎樣的情緒感受：

全體	象限 1 音樂	象限 2 音樂	象限 3 音樂	象限 4 音樂	平均分數
覺得是 象限 1	78.89% (71 人次)	24.44% (22 人次)	2.22% (2 人次)	7.78% (7 人次)	3.83
覺得是 象限 2	4.44% (4 人次)	33.33% (30 人次)	12.22% (11 人次)	2.22% (2 人次)	3.47
覺得是 象限 3	4.44% (4 人次)	25.56% (23 人次)	63.33% (57 人次)	25.56% (23 人次)	3.51
覺得是 象限 4	12.22% (11 人次)	16.67% (15 人次)	22.22% (20 人次)	64.44% (58 人次)	3.29

表 5.5a 各象限音樂總體平均的全體受測者回答正確率表，
每象限最高有 30 人*3 首歌=90 人次投票數

未受訓練	象限 1 音樂	象限 2 音樂	象限 3 音樂	象限 4 音樂	平均分數
覺得是 象限 1	86.67% (26 人次)	30% (9 人次)	0% (0 人次)	6.67% (2 人次)	3.85
覺得是 象限 2	0% (0 人次)	16.67% (5 人次)	23.33% (7 人次)	3.33% (1 人次)	3.2
覺得是 象限 3	6.67% (2 人次)	30% (9 人次)	46.67% (14 人次)	36.67% (11 人次)	3.57
覺得是 象限 4	6.67% (2 人次)	23.33% (7 人次)	30% (9 人次)	53.33% (16 人次)	3.31

表 5.5b 各象限音樂總體平均的未受過音樂訓練受測者回答正確率表，
每象限最高有 10 人*3 首歌=30 人次投票數

表 5.5a 和 5.5b 分別代表全體受測者及未受過音樂訓練者的情緒感知回答正確率表，可以得知對全體人數而言，受測者能夠正確的感受到該音樂所含有的情緒象限分類，並且分數在平均以上；即便象限 2 在內圈音樂方面與其他象限內圈音樂混淆，最後的總體平均上還是最能夠被分辨出來。最右一行表示選對象限音樂的受測者平均感受分數。而僅看未受過音樂訓練的受測者數據，則可發現象限 2 的分辨率很低，其他則是較全體人數平均低一些，平均分數也低了一點。圖 5.3 顯示，總體而言，受過音樂訓練的受測者在較易混淆的歌曲上能夠用自己的音樂經驗獲得較好的情緒感知；而其他象限的音樂，未受過音樂訓練的受測者雖然低於平均數據(象限 1 時稍高於平均數據)，但在該象限音樂的四個情緒選擇上，大多還是能夠選擇正確的答案。

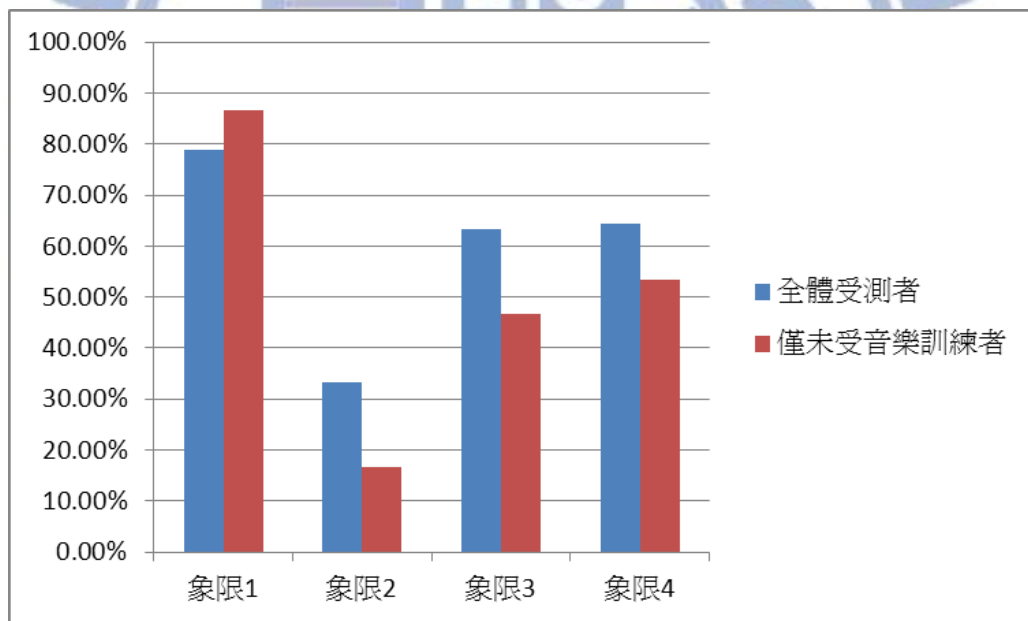


圖 5.3 全體受測者與未受音樂訓練者的各象限正確辨識率比較

5.3 第一次遊戲實驗結果

我們觀察在第一次遊戲後的問卷結果，在第一次遊戲中，總共有四個場景。在第一次遊戲中，只有音樂和場景，並無顏色的加入，實驗結果如表 5.6 所示，

第一行的場景 1(象限 4)代表場景 1 的遊戲場景設計是對應到情緒平面象限 4，且取情緒平面象限 4 的中間線段，來產生音樂當作該場景階段的遊戲配樂，試著給受測者有該象限相同的情緒感受。

	情緒感知正確率	平均分數	音樂比重	場景比重
場景 1(象限 4)	86.67% (26 人)	3.46	3.42	2.62
場景 2(象限 1)	70% (21 人)	3.62	3.62	3
場景 3(象限 2)	26.67% (8 人)	4	3.87	4.37
場景 4(象限 3)	66.67% (20 人)	4	3.75	3.6

表 5.6a 第一次遊戲的情緒感知實驗結果 - 全體受測者

	情緒感知正確率	平均分數	音樂比重	場景比重
場景 1(象限 4)	80% (8 人)	3.13	2.86	2.75
場景 2(象限 1)	80% (8 人)	3.63	3.63	3
場景 3(象限 2)	30% (3 人)	3.67	3.67	4.67
場景 4(象限 3)	80% (8 人)	3.88	3.63	3.25

表 5.6b 第一次遊戲的情緒感知實驗結果 - 僅未受音樂訓練的受測者

表 5.6 的音樂比重和場景比重，代表的是受測者除了感受到的是哪個象限的情緒，並且寫下平均分數後，在音樂和場景兩方面，受測者感覺到這兩個因素占的份量為何，彼此並無直接關係。將表 5.6a 的結果與表 5.5a 對比，可以發現在場景 3、4、1，也就是對應到情緒平面象限 2、3、4，的平均分數分別為 4、4、3.46 分，較表 5.5a 的 12 音樂分析結果相同象限的分數為 3.47、3.51、3.2931 均有些微提升，而情緒感知正確率也都相當高。但可以看到跟表 5.5a 相同的是，在場景 3，也就是對應到情緒平面象限 2 的正確率較低，以下試著探討原因為何。

首先將場景 3(情緒平面象限 2)的問卷情況單獨列出來觀察：

	場景 3(情緒象限是 2 的音樂+場景設計)
覺得是象限 1	43.33% (13 人)
覺得是象限 2	26.67% (8 人)
覺得是象限 3	23.33% (7 人)
覺得是象限 4	6.67% (2 人)

表 5.7 第一次遊戲場景 3(情緒象限 2)的情緒感知問卷數據

可以發現表 5.7，多數人將場景 3(情緒象限 2)的情緒混淆成情緒象限 1 的情緒，比較表 5.2 的第 2 象限情緒音樂數據，在中、外圈的音樂(由於遊戲配樂選用的是中圈的音樂)，混淆成情緒象限 1 的人數恰好是除了正確的選擇(選擇象限 2)外第二高的(中 9 人，外 8 人)。另外，由於遊戲場景 2 的設計內容最為豐富，且有刺激緊張感的劇情，或許會令受測者與第一象限的興奮感覺混淆。

而關於音樂比重和場景比重的部分，可以看到只有場景 3(象限 2)的場景比重大於音樂比重，推測是因為遊戲設計上，場景 3 的內容相當豐富和變化多，使受測者有較大的注意力在遊戲場景上；相反的，場景 1 的音樂比重大於場景比重最多(達 0.8 分)，則是因為該場景的內容設計較簡單、無變化，受測者較專注於音樂上。

由表 5.6b 得知，未受音樂訓練的受測者跟全體受測者的辨識率差異不大，但是在音樂比重的分數較低，例如在場景 1 內容單調沒變化，但未受音樂訓練的受測者卻不像整體數據一樣會較著重於音樂部分；而在內容豐富多變化的場景 3 則場景比重更是比音樂比重多了一分！顯示未受音樂訓練的受測者在辨識率上雖然比起整體數據無太大變化，但是對音樂方面感受程度較淺。

5.4 第二次遊戲實驗結果

接著我們看第二次遊戲的問卷結果，也就是加入場景顏色後的數據：

	情緒感知正確率	平均分數	音樂比重	場景比重	顏色比重
場景 1(象限 4)	66.67% (20 人)	3.45	3.55	2.75	2.75
場景 2(象限 1)	73.33% (22 人)	3.86	3.82	3	3.5
場景 3(象限 2)	40% (12 人)	3.58	3.67	4	3.92
場景 4(象限 3)	83.33% (25 人)	4	4.08	3.92	3.28

表 5.8a 第二次遊戲的情緒感知實驗結果 – 全體受測者

	情緒感知正確率	平均分數	音樂比重	場景比重	顏色比重
場景 1(象限 4)	70% (7 人)	3.29	3.29	2.57	3.29
場景 2(象限 1)	70% (7 人)	3.71	3.71	3	4.14
場景 3(象限 2)	30% (3 人)	3.67	4.33	4	4
場景 4(象限 3)	100% (10 人)	3.9	4	3.5	3.6

表 5.8b 第二次遊戲的情緒感知實驗結果 – 僅未受音樂訓練的受測者

與第一次遊戲的實驗結果比較，在各場景平均分數總體算是持平，但可以看到本來場景 3(象限 2)的正確率為 40%，將第二次遊戲的場景 3 數據單獨拿出來看：

	場景 3(情緒象限是 2 的音樂+場景設計+顏色)
覺得是象限 1	30% (9 人)
覺得是象限 2	40% (12 人)
覺得是象限 3	20% (6 人)
覺得是象限 4	10% (3 人)

表 5.9 第二次遊戲場景 3(情緒象限 2)的情緒感知問卷數據

結果發現在第二次遊戲，原本正確率低的場景 3(情緒象限 2)的正確性從 26.67% 提升到 40%，為最高的辨識度；雖然並不像其他三個場景正確率均超過 65%，但已是很大的增長。

分析其原因可以看到在場景 3 中，場景比重和顏色比重為 4 和 3.92，分數接近，且都高於音樂比重 3.67，可以推論在場景 3 中，除了豐富的遊戲內容外，顯現的顏色也是有幾乎同樣比重的情緒影響力，並且大幅提升了情緒分辨的正確率。

其他三個場景方面，則是與第一次遊戲時類似，都是音樂比重較高，但可以看到在高 Arousal 的場景 2、3(對應到情緒象限 1、2)，顏色的比重跟場景比重比較，能夠接近甚至超過，顯示在高 arousal 情緒象限的場景階段中，顏色能較明顯的影響受測者的情緒。

表 5.8b 顯示未受音樂訓練的受測者數據，在音樂比重方面大多也略低於整體數據，而顏色比重在高 arousal 時較高。

5.5 實驗結果

本論文所建置之情緒自動作曲與遊戲場景、顏色的相關實驗，可以歸納出：

1. 自動作曲系統能夠經由簡易的在情緒平面上畫線產生符合該情緒的音樂，而隨著線段遠離中心點，帶給受測者的感受逐漸更深，辨識正確率獲得提升，可驗證本實驗目的：希望帶給聽者不但符合情緒且有程度差異的音樂。
2. 加入符合情緒的遊戲場景後，各象限情緒感知的分數大多會微幅提升，而若遊戲場景內容豐富，受測者的專注力會在場景上感受到的情緒比重超過音樂，反之亦反。
3. 再加入顏色後，場景 3(象限 2)的情緒感知辨識度獲得改善(33% → 40%)，其他場景總體而言，辨識度也上升；另外，在高 arousal 的遊戲場景，顏色部份佔了情緒感知相當大的比重。
4. 另外在單獨探討未受音樂訓練的受測者數據，在辨識率上與整體數據相當，但是對於音樂的感受程度稍低。

由以上的實驗歸納結果，我們可以說本論文建構之系統相當有效，且能帶給遊戲設計者的幫助為：在設計較平靜不變的的遊戲內容時，本系統自動產生之遊戲配樂能夠提升玩家遊玩時的情緒感知；而在遊戲內容精彩豐富的場景時，參考本系統在情緒平面轉換出來的顏色，也能更加深玩家的情緒感受。

第六章、 總結

6.1 結論

本論文從情緒平面為出發點，做了有關情緒與音樂、情緒與顏色、情緒與遊戲場景的一個整合性研究，並且依據音樂特徵與情緒的關係，加上一些基礎樂理，發展出一套能夠經由簡單繪製線條產生情緒音樂的系統；而自動產生出來的音樂、對應情緒平面轉換出來的顏色、依照 EVG 理念設計的遊戲場景，在經過實驗後的結果，大致上符合設計想給受測者怎樣的情緒感知的期望；唯在高 Arousal 部分，受測者對大小調的鑑別度較接近，尤其在象限 2 的情緒感知若加入複雜的遊戲場景，受測者可能綜合音樂和複雜場景的影響，所測試之結果較為分散不集中。加入顏色後雖然可以提升其鑑別度，但還有可進步的空間。不過此系統的實驗結果已可大致上證明，除了以往個別研究情緒對音樂、對顏色的關係的正確性外，用適當的方式將這些個別的研究，歸納、整合起來造成的結果，甚至能讓受測者的情緒感知更加明顯，符合實驗設計者的期待，也因此可證明情緒平面↔顏色↔音樂特徵這三者，關聯性明確，且理論可以互通無礙。

6.2 未來研究

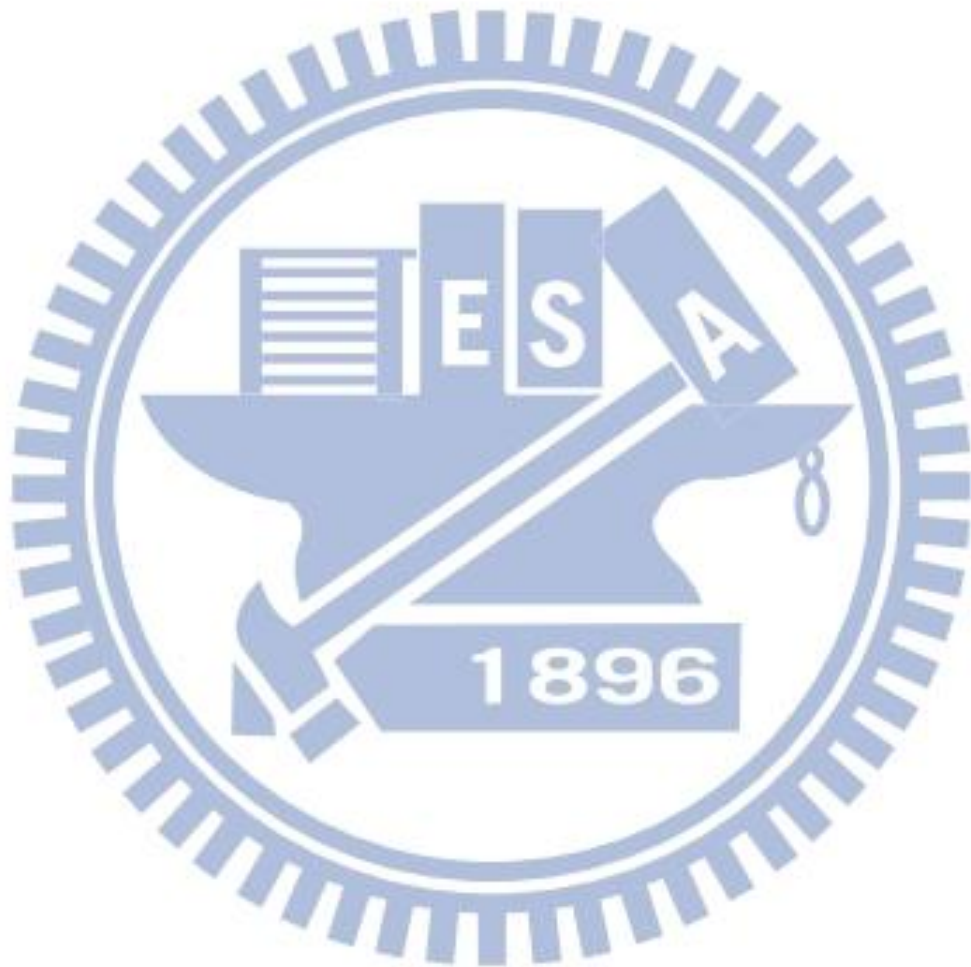
目前的自動作曲系統並無遵循任一特別的風格，若能夠研究風格與情緒的關係並加入作曲系統，可使音樂聽起來更加豐富，音高組成、和聲搭配、節奏變化等都有更多的選擇。

另外在遊戲場景中，只是單向的提供顏色、故事情節來觸發受測者的情緒感知，若能藉由一些互動裝置，例如分析受測者的鍵鼠行為、視訊擷取面部表情等，能夠對受測者的情緒反應掌握更好。

參考文獻

- [1] 音樂記譜法, Wikipedia,
<http://zh.wikipedia.org/zh-tw/%E6%A8%82%E8%AD%9C>
- [2] 音符, Wikipedia, <http://zh.wikipedia.org/zh-tw/%E9%9F%B3%E7%AC%A6>
- [3] 張錦鴻, 基礎樂理, 大陸書店, 1996
- [4] midi, Wikipedia, <http://zh.wikipedia.org/wiki/Midi>
- [5] Robert E Thayer, The biopsychology of mood and arousal, Oxford University Press New York (Oxford University Press, 1989)
- [6] Yi-Hsuan Yang et al., “Music Emotion Classification: A Regression Approach,” Multimedia and Expo 2007 IEEE International Conference on (2007): 208-211
- [7] Straus, J.N., Introduction to Post-Tonal Theory
- [8] I Wallis, T Ingalls, and E Campana, “Computer Generating Emotional Music The Design Of An Affective Music Algorithm,” in DAFx 08, 2008
- [9] Steven R Livingstone and Andrew R Brown, “Dynamic response: real-time adaptation for music emotion,” ACM Transactions on Computer Systems 9, no. 2 (2005): 105-111
- [10] Patrick Gomez and Brigitta Danuser, “Affective and physiological responses to environmental noises and music.,” International Journal of Psychophysiology 53, no. 2 (2004): 91-103
- [11] Patrick Gomez and Brigitta Danuser, “Relationships between musical structure and psychophysiological measures of emotion.,” Emotion Washington Dc 7, no. 2 (2007): 377-387
- [12] J A Russell and M Bullock, “On the dimensions preschoolers use to interpret facial expressions of emotion,” Developmental Psychology 22, no. 1 (1986): 97-102.
- [13] Hyeon-Jeong Suk, “Color and Emotion: A Study on the Affective Judgment of Color Across Media and in Relation to Visual Stimuli,” Context (Universität Mannheim, 2006):
- [14] N Kaya and H H Epps, “Relationship between color and emotion: a study of college students,” College Student Journal (PROJECT INNOVATION INC, 2004)
- [15] Niels A Nijdam, “Mapping emotion to color,” Emotion (2007): 2-9
- [16] Ning Wang and Stacy Marsella, “Introducing EVG: An emotion evoking game,” Intelligent Virtual Agents 4133 (2006): 282–291
- [17] Egoboo. <http://zippy-egoboo.sourceforge.net/>

- [18] H C Longuet-Higgins and C S Lee, “The Rhythmic Interpretation of Monophonic Music,” ed. H Longuet-Higgins, *Music Perception* 1, no. 4 (1984): 424-441
- [19] 董信宗、沈錕坤, 基植於音樂風格的電腦音樂自動伴奏
- [20] XML, Wikipedia, <http://en.wikipedia.org/wiki/XML>
- [21] Processing, <http://processing.org>



附錄 A

16 首歌曲名稱

編號	歌曲名
1	Offenbach-Rosenthal, Cancan
2	A. Ponchielli, La Gioconda —Dance of the Hours
3	E. Serra, Le Grand Bleu — Spaghetti del Mare
4	G. Mahler, Symphony No. 1, Langsam Schleppend
5	E. Elgar, Salut d' Amour
6	G. Holst, The Planets—Venus
7	R. Strauss, Also sprach Zarathustra—Von den Freuden und Leidenschaften
8	E. Serra, Le Grand Bleu—The Monastery of Amorgos
9	A. Dvorak, Symphony No. 9—Largo
10	E. Elgar, Enigma Variations—Romanza
11	Gandalf, From Source to Sea—Refuge Island
12	Manowar, Hail to England—Black Arrows
13	S. Barber, Adagio for strings Op. 11
14	D. Borgir, Puritanical Euphoric Misanthropy—Fear and Wonder
15	G. Mahler, Symphony No.5—Adagietto
16	M. Ravel, Piano Concerto in G major—Adagio assai

Technical Characteristics of the Musical Selections

Musical selection	Sound intensity	Tempo	Rhythm	Accentuation	Rhythmic articulation	Melodic direction	Pitch level	Pitch range	Mode	Complexity	Consonance
1	67.6	6.3	1.7	7.0	2.0	4.0	5.0	3.0	6.3	3.0	6.3
2	69.7	6.3	1.3	6.3	3.0	2.0	5.7	5.7	7.0	2.0	6.3
3	62.2	4.3	2.7	4.0	3.0	3.7	4.0	2.0	6.3	2.0	5.3
4	66.0	4.7	2.7	5.0	4.3	5.0	4.7	5.3	6.3	3.7	5.3
5	61.0	2.7	2.0	3.0	6.7	3.7	5.7	6.0	6.0	2.0	6.3
6	66.7	2.0	2.7	3.0	5.3	4.0	5.0	2.3	5.3	3.3	5.3
7	71.8	5.0	6.0	5.3	5.0	2.7	2.7	4.7	1.3	6.3	3.0
8	65.2	3.3	2.7	6.0	4.0	4.7	3.3	3.0	1.7	3.3	2.7
9	68.6	2.0	2.3	2.7	7.0	2.7	3.3	4.0	6.0	2.3	6.3
10	61.9	3.7	3.0	4.0	5.3	3.7	4.0	4.3	5.0	4.0	5.3
11	65.1	2.7	5.3	2.7	5.3	3.0	5.0	4.0	1.3	2.0	5.7
12	77.5	7.0	5.7	6.7	4.7	3.3	2.0	5.3	— ^a	4.3	3.0
13	73.1	2.3	3.7	5.7	6.7	1.0	6.3	5.7	3.0	4.3	4.3
14	67.7	2.0	1.7	2.3	6.7	3.0	4.0	3.7	2.7	4.7	4.7
15	62.3	1.7	5.0	2.7	6.7	2.0	4.7	3.3	5.7	4.3	5.0
16	58.9	2.3	2.0	2.0	7.0	3.0	4.7	2.3	6.0	2.7	6.0

附錄 B

音樂專家對 16 首樂曲的音樂特徵評分圖表

附錄 C

對 digital color 的 Arousal-Valence 數據表

colors	valence			arousal		
	M exp.II	M exp. III	P	M exp.II	M exp. III	P
dark red	2.778	2.783	(.983)	2.694	2.826	.575
deep red	3.889	3.457	(.035*)	3.139	3.348	.359
vivid red	4.056	4.087	.883	4.083	3.848	.264
brilliant red	3.056	3.478	.020*	2.667	2.891	.323
light red	3.028	3.674	.009**	2.611	2.696	.733
dark yellow	2.472	2.739	.256	2.278	2.478	.341
deep yellow	2.556	3.304	.003*	2.917	3.109	.445
vivid yellow	4.333	4.239	.662	3.694	3.630	.812
brilliant yellow	4.139	4.022	.584	3.472	3.261	.388
light yellow	3.472	3.457	.943	2.389	2.370	.935
dark green	3.806	3.152	.003**	2.667	2.674	.974
deep green	3.833	3.761	.727	3.028	2.609	.089
vivid green	3.583	3.870	.207	3.389	2.826	.023*
brilliant green	3.444	3.522	.733	2.972	2.565	.092
light green	3.111	3.348	.291	2.278	2.217	.798
dark blue	3.472	3.174	.183	2.389	2.739	.127
deep blue	3.639	3.696	.783	2.444	2.761	.172
vivid blue	4.111	4.109	.990	3.444	2.891	.043*
brilliant blue	4.139	4.326	.292	2.917	3.217	.262
light blue	3.778	4.152	.084	2.667	2.935	.300
dark violet	3.417	2.522	.001**	2.889	2.848	.863
deep violet	3.500	3.109	.081	3.472	3.087	.069
vivid violet	3.389	3.239	(.576)	3.389	2.826	.019*
brilliant violet	3.444	3.348	.679	3.167	2.544	.011*
light violet	3.528	3.609	(.739)	3.000	2.174	.000***
dark gray	2.167	2.065	.653	2.500	2.326	.480
medium gray	2.167	2.196	.887	2.361	2.174	.452
light gray	2.444	2.435	.961	2.028	2.087	.787
dark warm gray	2.056	1.957	.640	2.306	2.370	(.773)
medium warm gray	2.139	2.239	.633	2.222	2.022	.314
light warm gray	2.111	2.544	.048*	2.111	2.000	.583
dark cool gray	2.611	2.283	.086	2.417	2.239	.419
medium cool gray	2.694	2.565	.485	2.250	2.174	.714
light cool gray	3.222	3.304	.676	1.944	2.022	.709
white	3.333	3.239	.701	2.611	2.457	.578
black	2.472	2.152	.208	2.611	3.087	.115

附錄 D

96 名大學生對顏色的情緒相似度問卷數據結果

Emotions*	Red	Yellow	Green	Blue	Purple	Yellow-red	Green-yellow	Blue-green	Purple-blue	Red-purple	White	Gray	Black
Angry ^a	28 (28.6)	0	0	0	0	0	0	0	0	0	0	3 (3.1)	7 (7.1)
Annoyed ^a	0	0	0	0	0	5 (5.1)	8 (8.2)	7 (7.1)	0	2 (2.0)	0	0	0
Bored ^a	0	0	0	0	5 (5.1)	4 (4.1)	2 (2.0)	0	0	2 (2.0)	5 (5.1)	14 (14.3)	0
Calm ^b	4 (4.1)	0	29 (29.6)	60 (61.2)	28 (28.6)	0	0	16 (16.3)	38 (38.8)	13 (13.3)	8 (8.2)	5 (5.1)	0
Comfortable ^b	0	0	15 (15.3)	4 (4.1)	3 (3.1)	3 (3.1)	7 (7.1)	7 (7.1)	0	0	0	0	5 (5.1)
Confused ^a	0	0	0	0	0	0	2 (2.0)	6 (6.1)	0	0	0	6 (6.1)	0
Depressed ^a	0	0	0	6 (6.1)	0	0	0	0	12 (12.2)	8 (8.2)	0	23 (23.5)	22 (22.4)
Disgusted ^a	0	0	0	0	0	9 (9.2)	26 (26.5)	2 (2.0)	0	3 (3.1)	0	0	0
Empty/void ^a	0	0	0	0	0	0	0	0	0	0	25 (25.5)	0	0
Energetic ^b	5 (5.1)	10 (10.2)	0	0	0	14 (14.3)	0	10 (10.2)	0	0	0	0	0
Excited ^b	18 (18.4)	8 (8.2)	2 (2.0)	0	4 (4.1)	25 (25.5)	6 (6.1)	11 (11.2)	0	12 (12.2)	0	0	0
Fearful ^a	0	0	0	0	5 (5.1)	0	0	0	0	0	0	3 (3.1)	17 (17.3)
Happy ^b	21 (21.4)	74 (75.5)	28 (28.6)	10 (10.2)	21 (21.4)	31 (31.6)	11 (11.2)	36 (36.7)	13 (13.3)	26 (26.5)	0	0	0
Hopeful ^b	0	0	8 (8.2)	0	0	0	0	0	5 (5.1)	0	6 (6.1)	0	0
Innocent ^b	0	0	0	0	0	0	0	0	0	0	33 (33.7)	0	0
Lonely ^a	0	0	0	3 (3.1)	0	0	0	0	3 (3.1)	0	6 (6.1)	4 (4.1)	0
Loved ^b	15 (15.3)	0	0	0	0	0	0	0	0	17 (17.3)	0	0	0
Peaceful ^b	0	0	12 (12.2)	4 (4.1)	0	0	0	0	8 (8.2)	0	13 (13.3)	0	0
Powerful ^b	0	0	0	0	7 (7.1)	0	0	0	0	7 (7.1)	0	2 (2.0)	14 (14.3)
Sad ^a	4 (4.1)	0	0	8 (8.2)	13 (13.3)	0	0	0	10 (10.2)	0	0	30 (30.6)	24 (24.5)
Sick ^a	0	0	0	0	0	0	32 (32.7)	0	0	0	0	0	0
Tired ^a	0	6 (6.1)	0	0	9 (9.2)	0	0	0	5 (5.1)	0	0	5 (5.1)	7 (7.1)
No emotion	3 (3.1)	0	4 (4.1)	3 (3.1)	3 (3.1)	7 (7.1)	4 (4.1)	3 (3.1)	4 (4.1)	8 (8.2)	2 (2.0)	3 (3.1)	2 (2.0)