

中華民國 第 49 屆中小學科學展覽會

作品說明書

高中組 生活與應用科學科

佳作

040809

符合風景照片觀看感覺之音樂產生器

學校名稱：國立臺灣師範大學附屬高級中學

作者： 高二 吳季儒	指導老師： 林麗玉
---------------	--------------

關鍵詞：電腦作曲、Canny 演算法、MIDI 音樂

摘要

本研究結合文獻資料及個人創意，探討圖片轉換音樂的可能性。由本研究撰寫而成的軟體，可藉由使用者的問答，及圖片特徵的擷取，搭配建構好的樂句資料庫，成功產生具有悅耳性、圖片貼切性的創新音樂。

壹、研究動機

看到五線譜音高的排列，忍不住幻想是山巒的起伏。看到繪圖作品，耳中不禁會響起莫名的旋律。我開始為他們之間的關係有了聯想。繪圖是種靜態的美，音樂則是動態的美。乍看之下是種分隔、不同的兩件事物，但感覺卻可以將他們有所牽連。而接連兩者的關鍵，理所當然的是觀看者的情緒。因為藝術作品應該都是以「人」作為對象而產生，所以用電腦程式作媒介，架起它們之間的「橋」。

利用圖片產生音樂的過程，主要牽涉到程式設計、影像處理、認知科學和樂理等部分，可視為美術上較為獨特的一個旁支，和音樂上電腦自動作曲的實踐。

貳、研究目的

- 一、經由文獻理論，確認影像對於電腦音樂創作的依循理論
- 二、抓取影像特徵技術的實踐
- 三、經由圖片特徵及與使用者對於圖片觀看情緒，產生其符合圖片觀看感覺之音樂

參、研究設備及器材

- 一、硬體設施：個人電腦、筆記型電腦
- 二、程式撰寫軟體：Dev-C++、Borland C++Builder6
- 三、音樂製作軟體：SYTMP

肆、研究過程或方法

一、文獻探討

以下先針對現有關於音樂和影像之關係的理論加以討論，以佐證影像對於電腦音樂創作的理論依據。

(一)

相關研究記載最早可追溯到西元八世紀的春秋時代政治家管仲，所著之書《管子·五行篇》中說：「昔黃帝以其緩急，作五聲，以政五鐘。令其五鐘，一曰青鐘，大音，二曰赤鐘，重心，三曰黃鐘，麗光，四曰景鐘，昧其明，五曰黑鐘，隱其常。五聲既調，然後作立五行，以正天時。五官以正人位，人與天調，然後天地之美生。」其中五音（五聲）為角、徵、宮、商、羽，五鐘的音高為角、徵、宮、商、羽五音，五鐘的名稱為青、赤、黃、景（白）、黑五色。

(二)

西元前六世紀畢達哥拉斯的「顏色與音樂調性之間頻率共振」的理論，音樂就像七彩的彩虹一樣，是有色彩的。在大調音階中有 Do、Re、Mi、Fa、Sol、La、Si 七種調，這七種調的運用除了取決於樂器本身的音域條件及音色特性外，作曲家在選擇來作一首樂曲時，也考慮到它的光度。像 Do 調代表紅色，Re 調代表橙色，Mi 調代表黃色，Fa 調代表綠色，Sol 調代表藍色，La 調代表青色，Si 調代表紫色。每一個調性的明暗會影響欣賞的感覺，如果再加上小調的應用，這種感覺會更明顯，比如覺得聽起來開朗活潑，或是黯淡沉滯等等。因為色彩性也往往代表某種特徵，如紅色代表體力充沛、熱情、獨立，橙色代表勇敢、外向，黃色代表思考、多情、內向，綠色代表平衡、寧靜、健康，藍色代表鎮定、沉著，青色代表乾淨，紫色代表服務…等

(三)

西元前四世紀亞里斯多德提出的理論。他認為各種色彩就如音樂之和聲，可按比例分配，做最愉悅的安排。

(四)

十七世紀，知名物理學家牛頓也曾認為紅、橙、黃、綠、藍、靛、紫七色，相當於 C、D、Eb、F、G、A、Bb 七個音。十九世紀物理學家更進一步發展出色彩的屬性對應於音樂屬性，並依聲音震動次數比較波長，以此對應出色彩與音樂的共感覺，表 1 為音波與光波震動次數對照表。

音高	音頻(每秒音波震動次數)	顏色	光波(A.U.)
C	256(國際高度)258	暗紅	7360(Å)
C#	273(國際高度)274	紅	6900(Å)
D	288(國際高度)290.3	橘紅	6428(Å)
D#	307(國際高度)307.5	橘	6130(Å)
E	320(國際高度)325.8	黃	5890(Å)
F	341(國際高度)345.2	黃綠	5520(Å)
F#	352(國際高度)365.8	綠	5352(Å)
G	384(國際高度)387.5	藍綠	4906(Å)
G#	410(國際高度)410.6	藍	4600(Å)
A	427(國際高度)435	藍紫	4416(Å)
A#	448(國際高度)460	紫	4204(Å)
B	480(國際高度)488	暗紫	3920(Å)
C	510(國際高度)517	不可見	3680(Å)

表 1 音波與光波震動次數對照表

(五)

總和以上幾點前人的理論，可發現其中偏重於色彩對音樂的連結較為普遍。如果利用圖片色彩整體出現頻率或是分段出現的位置來對應表一「音波與光波震動次數對照表」，此方法產生的音樂為第四點的實踐方法之一。但誠如第二點後半所說，普遍會認為紅色代表熱情，藍色代表鎮定、沉著。如果比照表則會發現藍色對應出的音頻比紅色來的高。在物理上，頻率較高通常顯示是動能較大，有較活潑的本性。所以但純就這兩點來進行圖片與音樂的轉換，似乎有其衝突點存在。

二、研究背景

由上述文獻探討結果可以證明，圖片轉出音樂是有其可能性及理論。此研究最主要目的是要讓電腦「自動」產生出符合觀看情緒的音樂。但受限於現有人工智慧的技術發展，尚無法讓電腦具備人性化的創作能力。故此研究著重於參數對應產生出音樂所具備的元素，而這些參數大多是從圖片中特徵萃取而來。情緒是種很活很主觀的人類感情，因此本研究加入一部分由使用者自行選擇的參數，產生更精確的樂曲。

對人類來說，判斷一段音樂屬於樂音或是噪音是簡單而直覺的能力。但對電腦來說相當的困難，對此篩選產生出有悅耳性的音樂是很重要的問題。本研究選擇以演唱會中「組曲」的概念作解決。組曲是爲了在有限時間內唱出許多歌曲，故只選擇最爲膾炙人口的部份(一般來說是副歌)接連演唱。演唱會式的組曲單位是「一首歌的某部份」，而本研究是以「樂句」爲單位。樂句是旋律的組成部分，由幾個小節組成。它可表現一個樂曲的特定氣氛，有自己的特性，但構思並不完整，所以單獨的樂句並不能構成獨立的樂曲。將經過程式挑選出的幾段樂句連接起來，成爲一首完整的歌，此爲本研究的組曲產生音樂概念。而樂句連接方法，則是以「模仿」現有完整音樂，視完整的樂章裡樂句連接爲有人爲手法的連接。此連接方法就會有悅耳性的連接，不至於出現過於誇張的不和諧。

本研究限定圖片主題為風景照，因為風景照的主要線條較具有代表性，且音樂風格相近，故選之為此研究限定的主題。而研究範圍是以圖片的線條、亮度、及內容為基礎，示意圖如圖 1。



圖 1 研究範圍示意圖

(一)以圖片中的線條形狀及使用者情感來挑選樂句

當看到海洋的圖片時，我們會希望能聽到的是較為平緩的浪潮聲，而遇到山巒的圖片時，會與音高起伏較大的音樂產生情感上的連結。本研究即利用上述想法來選擇主線條為產生音樂的依據。

分析文獻裡電腦音樂創作的方式，大多是在某些基礎上，演算出下一個音符。然而，此法較無法顧及創作樂理及樂曲的行進走向，原因為作曲家作曲的想法並不能完全建構模型去詮釋。故本研究以「樂句」為單位，讓經作曲家寫出的不完整音樂片段，能在此依照圖片主線條的起伏，產生出相對應跟隨其軌跡的「創新」音樂。

(二)以圖片中的亮度來決定演奏速度

本研究參考多數人及文獻中提及對圖片亮度的觀看感情，進而選擇以圖片中的亮度來決定演奏速度。

圖片當中的亮度，對於觀賞者情緒上直接的影響雖然會有許多不同，但是大致上亮度較低的圖片容易讓人有陰沉鬱悶的感覺，而亮度高的圖片，比較容易使人產生愉悅、開朗的心情。

(三)以圖片內容決定樂器

本研究限定圖片主題為風景照。在風景照中，有海浪、下雨等較為實際的具體物像，以及季節景色等抽象主題。在此根據音樂家對圖片主題和樂器的連結，決定最後演奏的樂器。

三、研究方法

以下針對針對本研究使用之邊緣偵測及 MIDI 相關理論加以說明。

(一)邊緣偵測

1.Sobel 演算法簡介

Sobel 演算法是圖像處理中的技術，主要用作邊緣偵測。在圖像的任何一點使用此演算法，將會產生對應的梯度向量。本研究以 Sobel 演算法做特徵萃取的線條部份處理方法。

此演算法包含兩組 3x3 的矩陣，分別為橫向及縱向，將之與圖像作平面卷積，即可分別得出橫向及縱向的亮度差分近似值。如果以代表原始圖像，及分別代表經橫向及縱向邊緣檢測的圖像，其公式如下：

$$\mathbf{G}_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * \mathbf{A} \quad \text{and} \quad \mathbf{G}_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * \mathbf{A}$$

G_x 及 G_y 分別負責檢知 X 與 Y 方向的邊緣變化，因為 ΣG_x 與 ΣG_y 皆為零所以如果一個點其相鄰的八個點都是相同值，表示沒有邊緣變化，其值也會是零。如果左右的值不同， G_x 輸出結果就不會是零，但 G_y 仍會是零直到上下的值不同為止。

圖像的每一個像素的橫向及縱向梯度近似值可用以下的公式結合，來計算梯度的大小。

$$G = \sqrt{G_x^2 + G_y^2}$$

然後可用以下公式計算梯度方向。

$$\Theta = \arctan\left(\frac{G_y}{G_x}\right)$$

此演算法可適用於灰階及全彩影像，惟獨全彩影像牽涉到多維梯度的計算，且跟全彩轉灰階後的邊界偏失不大，故此研究選擇將全彩轉灰階再進行 Sobel 演算法，較為簡單也可使程式運行更精簡。

在本研究實踐 Sobel 演算法後，發現其邊界線段較為破碎，所以須再經由連結邊界線段的步驟，來抓出主線條。此步驟將會花費許多的時間，故評估效能後，改以抓出邊界線較不易破碎的 Canny 演算法來抓出邊界線。

2.Canny 演算法簡介

Canny 的邊緣偵測器原理：

- (1) 應用高斯濾波器在灰階影像 $f(x,y)$ ，得到平滑影像 $g(x,y) = f(x,y) * w_G(x,y; \sigma)$
- (2) 應用微分濾波器 $\nabla g(x,y)$ 計算邊緣強度和方向

Canny 提出兩個方法來滿足邊緣偵測器單一響應和定位正確性的要求：

(1) Non-maxima suppression

- ① 對每一點 $C(x,y)$ ，選定垂直於 orientation 方向兩個側邊的鄰近點，記作 A 和 B；
- ② 如果 $M(A) > M(C)$ or $M(B) > M(C)$ ，則 C 不為邊界(設定 $M(C(x,y))=0$)；
- ③ 輸出(edge)強度影像 $M_{NMS}(x,y)$

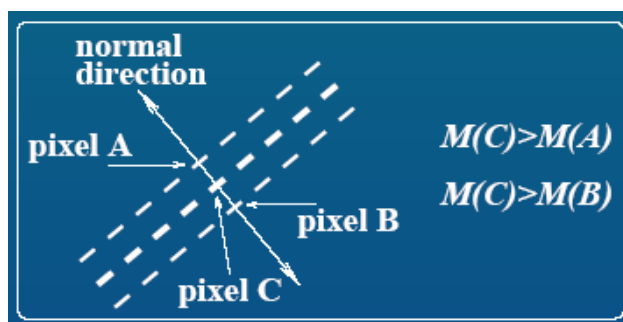


圖 2 Non-maxima suppression

non-maxima suppression 輸出仍會帶有一些非邊界的 local maxima，同時連接性質不明顯的邊界區域。以下的 Hysteresis thresholding 法提供了解決方案。

(2) Hysteresis thresholding

- ① 定義兩個 thresholds, T_{high} and T_{low} ，
像素 (x,y) 如果 $M_{NMS}(x,y) > T_{high}$ ，該像素就稱為 strong，
像素 (x,y) 如果 $M_{NMS}(x,y) \leq T_{low}$ ，該像素就稱為 weak，
所有其他的像素稱為 candidate；
- ② 如果像素 (x,y) 是 weak，則略去；如果是 strong，則輸出為邊界像素；
- ③ 如果像素 (x,y) 是 candidate，而且 $M_{NMS} > T_{low}$ ，則判斷是否沿著 local maxima 相連的邊界方向有穿過 (x,y) ，若是，則輸出為邊界；
- ④ 如果 candidate 像素 (x,y) 與一 strong 像素，則輸出該 candidate 為邊界。

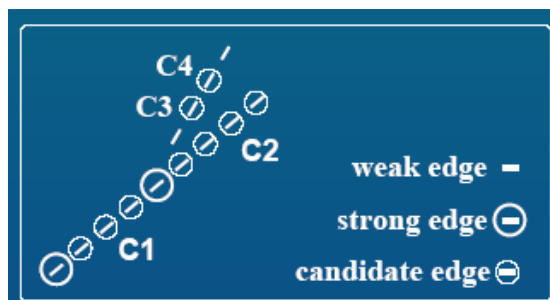


圖 3 Hysteresis thresholding

(二)MIDI 與 SYTMP 簡介

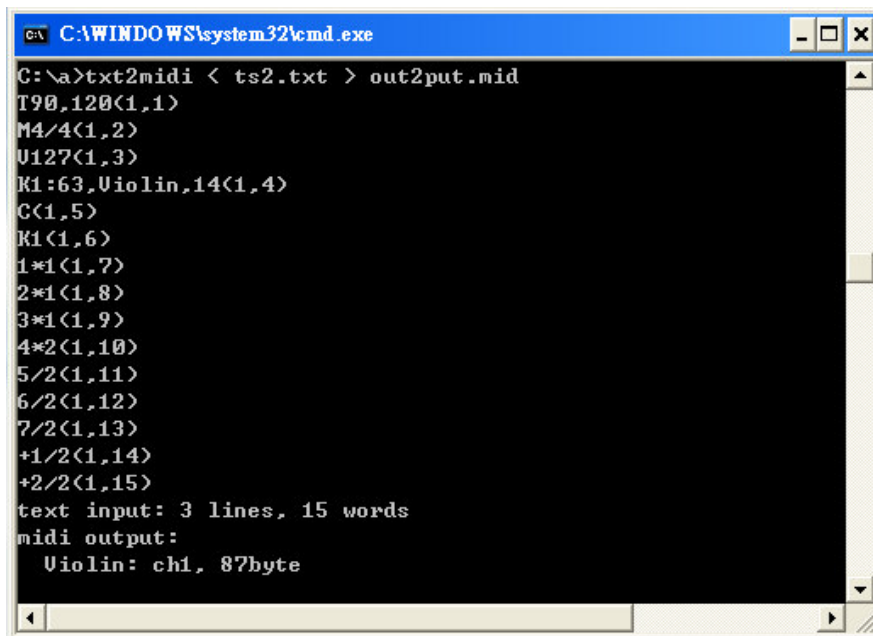
本程式最後產生的音樂會以 MIDI 格式呈現。以下簡介 MIDI 及 MIDI 音樂製作軟體 SYTMP。

1.MIDI 簡介

MIDI(Musical Instrument Digital Interface)是「電子樂器數位介面」的意思。為一種數位電腦式語言，主要用途是統一的音樂資料，使之可以互通消息。其可演奏的音色(樂器)多達 128 種，詳見附錄一。

2. SYTMP 軟體簡介

SYTMP (Text Music Project from Samuel Yang) 是設計用來『以文字模式編輯樂譜』的軟體，讓使用者可以在文字模式下輸入樂譜，轉換成可以播放的 MIDI 檔。此軟體為開放軟體，以 c 語言寫成，操作畫面如圖 4。因與本研究主程式部份撰寫程式語言相同，故選之做為最後音樂呈現部份的輔助軟體。



```
C:\WINDOWS\system32\cmd.exe
C:\>txt2midi < ts2.txt > out2put.mid
T90,120<1,1>
M4/4<1,2>
U127<1,3>
K1:63,Uiolin,14<1,4>
C<1,5>
R1<1,6>
1*1<1,7>
2*1<1,8>
3*1<1,9>
4*2<1,10>
5/2<1,11>
6/2<1,12>
7/2<1,13>
+1/2<1,14>
+2/2<1,15>
text input: 3 lines, 15 words
midi output:
  Violin: ch1, 87byte
```

圖 4 SYTMP 操作畫面

四、程式架構與研究過程

圖 5 為本實驗程式整體架構圖，接下來將針對架構圖的各部份加以說明。

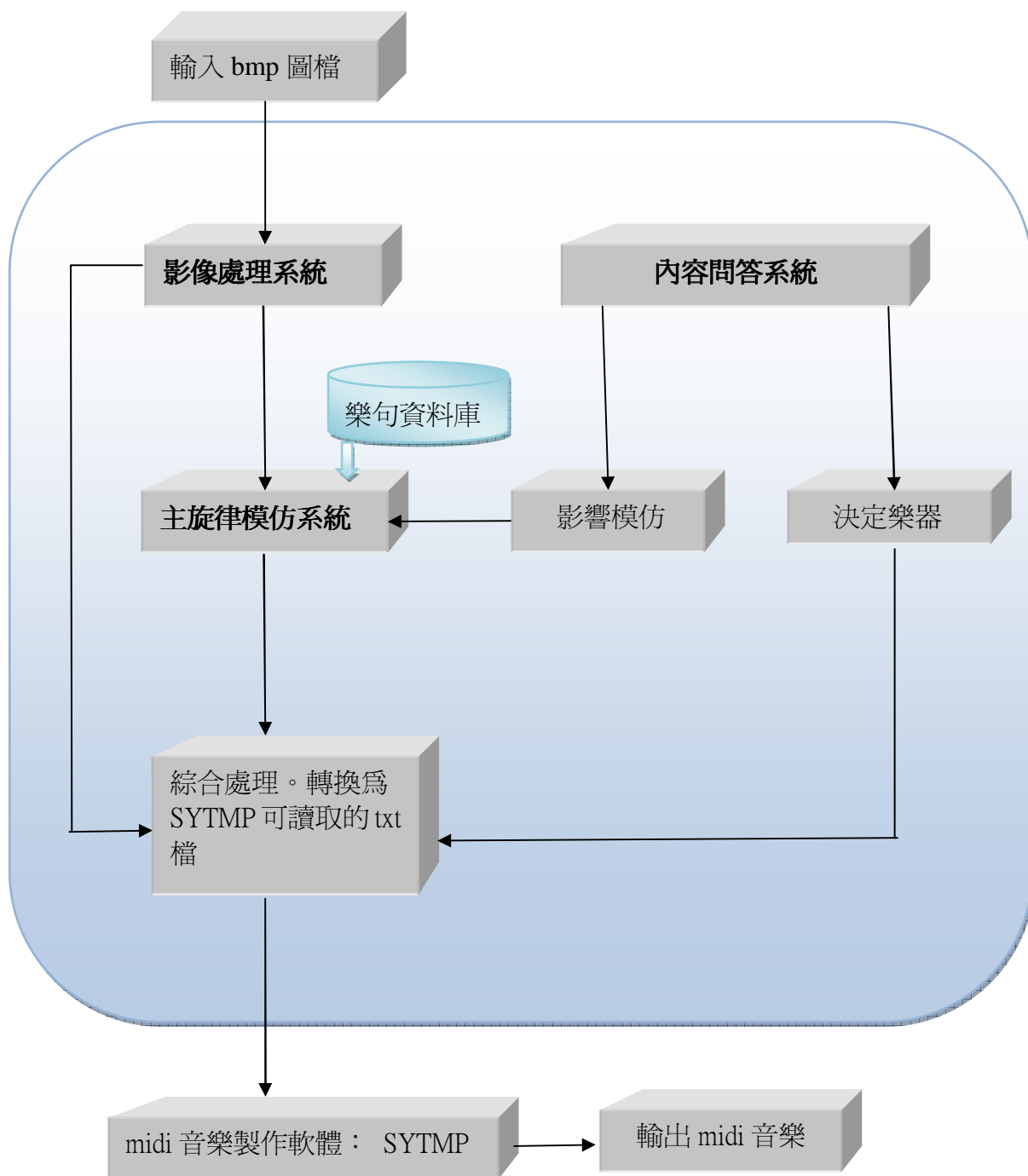


圖 5 程式架構圖

(一)內容問答系統

內容問答系統為程式的與使用者做接觸的步驟，讓使用者的想法有機會能參與音樂產生。這是最直接的讓音樂產生程式捕捉使用者情緒的步驟。

本系統會詢問使用者下列兩個問題：

Q1.請問您觀看此圖片是屬於何種感情：

A0 無感情	A1 愉悅	A2 生氣	A3 沉鬱
--------	-------	-------	-------

※此答案提供主旋律模仿系統的情緒加權的參數

Q2.請問此風景照是屬於下列何種主題：

A0	無法分類	A1	春天	A2	夏天	A3	秋天	A4	冬天
A5	雪景	A6	雲	A7	霧	A8	岩石	A9	山壁
A10	山巒	A11	海洋	A12	草地	A13	樹林	A14	建築(東方)
A15	建築(西方)	A16	花	A17	海岸	A18	農田	A19	河口
A20	河川	A21	水庫	A22	湖泊	A23	河床	A24	沙漠

※此答案提供主旋律模仿系統的特定樂句加權的參數，以及決定演奏的樂器

由於限定於風景照的類別，所以是問使用者風景照是何種類型。此問題決定最後音樂要以何種樂器呈現，並也會影響主旋律模仿系統的參數。雖說影像辨識已有一定程度的技術水平，但很有可能判別錯誤圖片主題。以攝影角度來說，拍攝最大面積的不一定是要呈現的主題，且其中可能有隱諱的故事性，故直接詢問使用者圖片內容的主題是較好的方式。

Q2.請問此風景照是屬於下列何種主題：

※此答案提供主旋律模仿系統的特定樂句加權的參數，以及決定演奏的樂器

2.風景照的主題配對樂器之決定

請教樂齡超過七年之三人，共同討論出各種風景照主題該用何種樂器演奏較為貼切，以作為預設的風景照主題配對樂器，表 2 為圖片主題與配對樂器的對應表。

答案	配對之預設樂器	答案	配對之預設樂器
A0 無法分類	平台鋼琴	A13 樹林	高音薩克管
A1 春天	鐵琴	A14 建築(東方)	古箏
A2 夏天	小號	A15 建築(西方)	豎琴
A3 秋天	音樂盒	A16 花	排笛
A4 冬天	手風琴	A17 海岸	大提琴
A5 雪景	嘹唳鋼琴	A18 農田	中音薩克管
A6 雲	豎琴	A19 河口	小提琴
A7 霧	音樂盒	A20 河川	小提琴
A8 岩石	法國號	A21 水庫	法國號
A9 山壁	雙簧管	A22 湖泊	大鍵琴
A10 山巒	大鍵琴	A23 河床	尼龍吉它
A11 海洋	大提琴	A24 沙漠	鋼線吉它
A12 草地	小提琴		

表 2 圖片主題對應之配對樂器

(二)影像處理系統

此部分包含基本的影像開檔、讀檔的程式碼。進行圖檔的開檔，讀取其 RGB 值。其中，亮度計算的部份，是利用亮度公式： $Y=0.299 \times R+0.587 \times G+0.114 \times B$ ，計算出整體圖片之亮度平均。以其範圍劃定演奏時的速度。通常面對明度比較低的圖片時，會比較希望聽到的是較慢的音樂，所以此處以亮度決定演奏速度。表 3 為亮度與演奏速度的對應表。

亮度		速度(拍/分鐘)
0~51	慢板	56
52~102	行板	60
103~153	小行板	66
154~204	中板	96
205~255	稍快板	108

表 3 亮度對應之演奏速度

(三)主旋律模仿系統

圖 6 為主旋律模仿系統架構圖，以下將針對系統內各步驟加以說明。

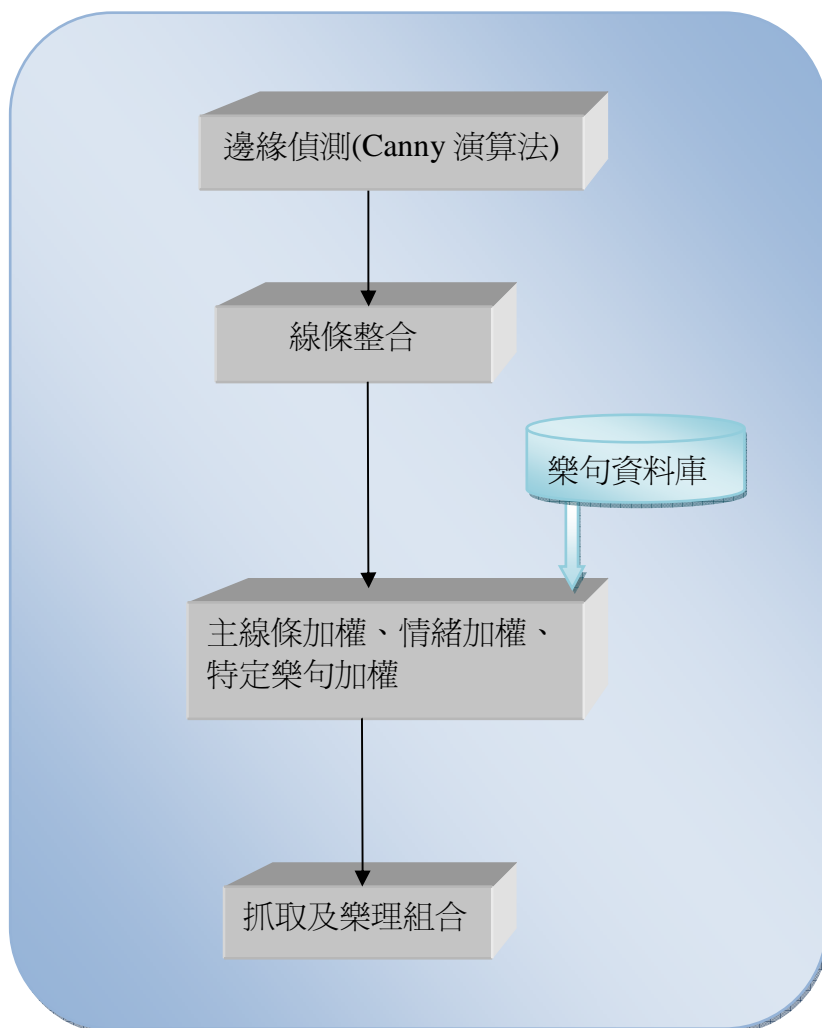


圖 6 主旋律模仿系統架構圖

1.邊緣偵測

以圖 7 的風景照為例，利用 Canny 演算法進行邊緣偵測，可得到圖 8 的結果。

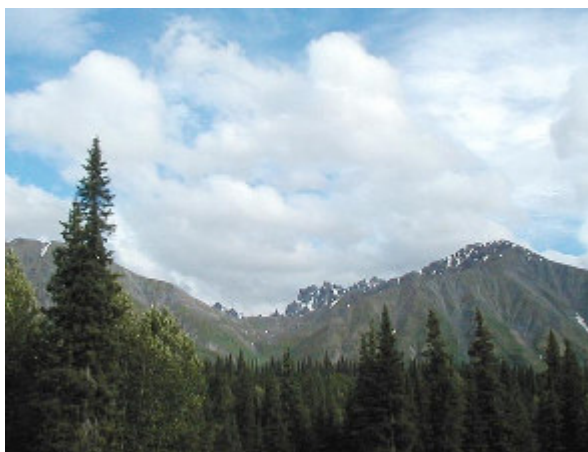


圖 7 原圖



圖 8 在 Canny 參數值為(150,100)的輸出圖

經 Canny 演算法處理後，邊界線較不破碎。故此演算法可省去連結破碎線段的步驟，有助提升程式效能。

爲了抓出一條圖片的主線條，故選擇長度最長，且起伏須超過圖片高度六分之一的線條，做爲具有代表性的主線條，結果如圖 9

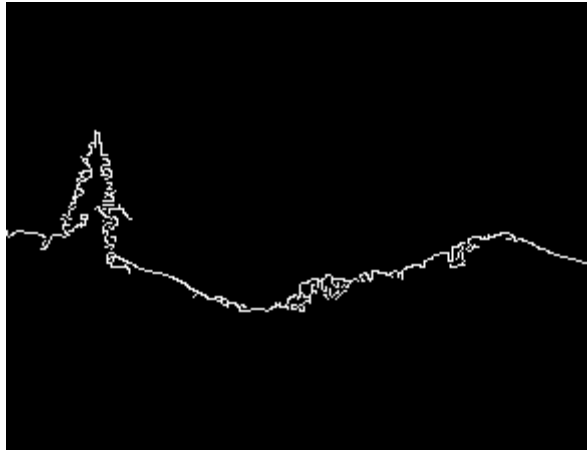


圖 9 抓出最長且起伏超圖片高度六分之一的主線條

2. 線條整合

將找出之主線條依照橫坐標分成 8 等分，計算每等分內所有點的縱座標平均。將線分段是爲了要看出線條的起伏。如果分段太少會看不太出起伏且頂點位置可能不確實。分太多段的線條起伏精準性雖已達許可水準，但會造成系統運算上有較大的負擔。經測試比較，拆成八份較爲合適。以圖 9 結果爲例，示意圖如圖 10。

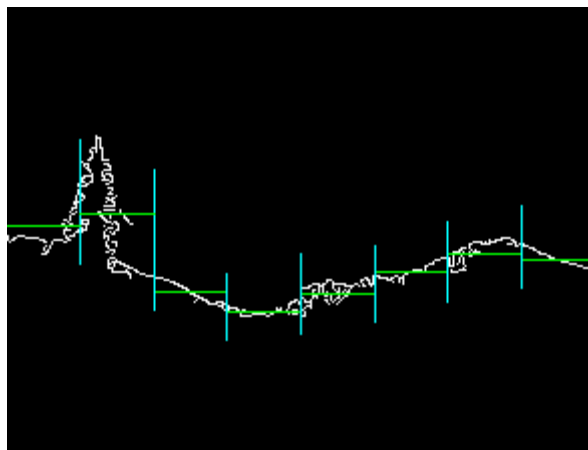


圖 10 各等份內平均高度以綠色線表示

3.樂句資料庫

由於之後需針對樂句資料庫裡各樂句進行加權處理，在此先介紹樂句資料庫的建構方法。

將市售各種樂譜，依人工判斷將之分成樂句，並紀錄其：(1)音高、(2)音長、(3)屬於哪個樂章、(4)樂章的第幾個樂句，以及輸入進資料庫時就決定的(5)此音章聆聽情緒(愉悅、生氣、沉鬱、無感情)、(6)特定主題。

其中，聆聽情緒為輸入資料庫時，感受此樂句所屬的情緒而決定。而特定主題好比從韋瓦第(1678—1741)所寫的「四季」春天樂章裡輸入進去的樂句。此特定主題就為春天。

5.加權

針對資料庫裡各樂句的抓取優先數，進行數據加權。其目的為利用主線條及內容問答系統的答案所輸進來的情緒和特定主題參數，搭配樂句資料庫的音高、屬於哪個樂章及此樂章聆聽情緒，進行更進一步的樂句挑選。開始時將賦予每個樂句一個「抓取優先數」，初始值為 50。由數值運算的結果對樂句抓取優先數進行加權。

(1)主線條加權

以主線條的八個平均大小關係的比對資料庫樂句的音高大小關係，其相似度越高，則該樂句越可能被挑選。

加權規則如下：

將樂句的音分成八等份，計算每份音高的平均。



圖 11 上面為原始樂句，下面為各等份的平均音高

主線條的前後兩個平均之間大小關係與樂句資料庫裡的樂句音高平均，有一處相同則該樂句抓取優先數加權 5%(乘以 1.05)。

主線條數字最大者出現在第幾個位子(1~8)，與樂句資料庫裡的樂句最高音出現位置，有一處相同則該樂句抓取優先數加權 5%(乘以 1.05)。接著是次大者，依此類推。

(2)情緒加權：

接受內容問答系統答案，將資料庫所有樂句的聆聽情緒跟內容問答系統第一個問題答案相同者，該樂句抓取優先數加權 10%(乘以 1.1)。但假如內容問答系統的答案為無情緒，則不加權。

舉例來說，假如使用者對此圖片所感受到的情緒為愉悅，在內容問答系統應會選擇 A1.愉悅。若在資料庫又有一個樂句，其聆聽情緒亦為：愉悅。則該樂句抓取優先數加權便乘以 1.1。

(3)特定主題加權：

接受內容問答系統答案，將資料庫所有樂句的特定主題跟內容問答系統第二個問題答案相同者，該樂句抓取優先數加權 15%(乘以 1.15)。

舉例來說，假如此圖片主題為 A1 春天。樂句資料庫裡從韋瓦第 (1678—1741) 所寫的「四季」春天樂章裡輸入進去的樂句，因其特定主題一樣為:春天。則該樂句抓取優先數加權便乘以 1.1。

5.抓取和樂理組合

此部份分為抓取及樂句加權、樂理組合，分述如下：

(1)抓取及後句加權

首先選取抓取優先數經加權後最大者，為第一個樂句。抓出第一個後，對資料庫裡的被抓出來的那句，如果其在同一樂章有下一句，對下一句的抓取優先數加權 8%(乘以 1.08)。此步驟稱為後句加權。

舉例來說，假設第一次抓出為標號第 350 的樂句，假設它同一樂章有後面一句，則後面那個樂句(標號 306)的抓取優先數 $\times 1.08$ 。

加權完後，再選取抓取優先數第二大者為第二句，再一次進行後句加權。直到抓出五句為止。

(2)樂理組合

現在抓取出的五個樂句，彼此之間沒有樂理的連接，兩樂句任意連接的話可能會出現不協調感。對此問題本研究選擇以模仿的方式解決，模仿對象就是現有資料庫裡的樂曲。因為資料庫裡的樂句，如果在同一樂章下，前後兩樂句是經過音樂家的樂理排列而成。所以與樂句資料庫度相似越大，則越可能是有樂理的音樂。

在此定義銜接時只看前面樂句最後 2 個音及後面樂句最前面 2 個音的音高及音長。比較其兩個樂句的四個音與樂句資料庫裡同一樂章前後兩樂句銜接，相似度越大則其組合則可能較具有樂理性。將五個樂句排列出最具有和資料庫相似的旋律。此旋律為主旋律模仿系統輸出的成品。

伍、研究結果

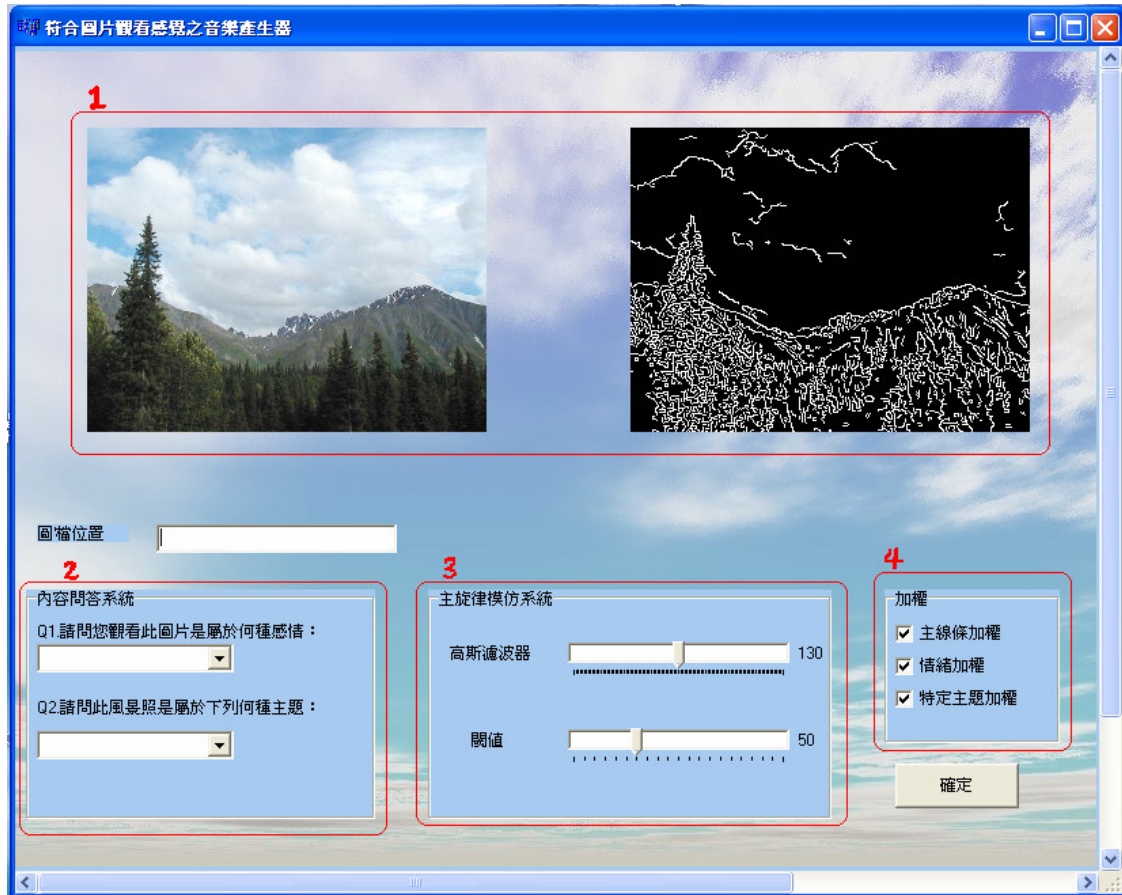


圖 12 程式畫面

本研究使用 Borland C++Builder6 來做為本程式介面撰寫軟體，程式介面如圖 12。在程式的第 1 部分分別以兩個圖形方塊呈現原圖及經演算法處理過後的影像。第 2 部分為內容問答系統，以下拉選單的方式來選擇答案。第 3 部分的主旋律模仿系統，可經由使用者調整參數，改變抓出的邊界點數及相連程度。第 4 部分的加權，則可以選擇使用哪些加權方式。

將程式輸入 30 張準備好的風景照(見附錄一)，由測試者依個人觀感回答並產生音樂後，進行以下三種測試，總測試人數為 25 位。

一、測試產生出的音樂是否好聽，以人為測試方式，驗證此研究方法能否實際產生出有悅耳性的音樂成品。

從結果很顯然的看出，以本研究的實驗方法所發展出的音樂產生程式，能有中上程度的水準產生有悅耳性的音樂作品。此方法可加以延伸為新的電腦音樂創作手法。

	悅耳度評分(1~5 分)
本研究之成品	3.47 分
隨機亂數音樂	1.25 分

二、測試程式所費時間與使用者可能會失去耐性之實驗。由於程式所花費時間跟圖片內容有很大關係，包括圖片大小與邊界線段的多寡皆會影響程式運行時間。每張圖片產生出音樂的時間相差有些大。故用受試者個人的耐性程度來評比。

經測試後，16%的受試者認為過半的測試資料所費時間過久，在程式使用上很容易失去耐性。

三、為了檢驗此程式是否能產生出符合圖片觀看情緒的音樂，以同樣的測試資料對測試者進行測試。對照組為程式內所有參數皆由亂數決定的作品。

	貼切度評分(1~5 分)
測試組(依圖片產生音樂參數)	2.88 分
對照組(亂數參數)	1.08 分

四、表 4 為的圖片以及其產生之音樂片段表，為本研究其貼切度排名前五名的成品。








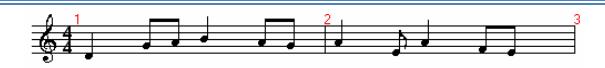


排名	圖片	悅耳度	貼切度	音樂片段
1		4.36	4.4	
2		4.08	4.28	
3		4.72	3.88	
4		4.6	3.72	
5		3.96	3.4	

表 4 本研究產生貼切度排名前五名的成品

陸、討論

以下針對本研究進行幾點討論：

一、此音樂產生的概念除了以圖片的特徵萃取作為參數使用外，還引入了「組曲」的概念完成樂曲。有別於大部分電腦音樂程式採用的演算式作曲法，故此研究方法有一定程度的創新性。

二、由測試一中可看出，程式的悅耳性有中上程度的水準。不過還有許多提升悅耳機率的空間，應能由更進一步的篩選機制提升品質。

三、比較 Canny 演算法及 Sobel 演算法的效能，Canny 演算法抓出主線條比起 Sobel 演算法快了約五倍。原因是 Sobel 演算法須經過連結破碎線段的步驟，此步驟將會花費許多時間。而 Canny 演算法則因為抓出之邊界線不易破碎，故不須此步驟。而兩演算法抓出主線條正確率差不多，故 Canny 演算法對於本研究較佳。

Sobel 演算法時間過久的問題，可能在於連結破碎線段的步驟用了許多遞迴的寫法，加快運算效率於此有兩種方法：(一)將遞迴寫法改成迴圈寫法(二)改以組合語言撰寫

四、由測試三的結果，證明以圖片的特徵萃取提供參數的音樂產生，能有過半的水準讓測試者認為產生的音樂貼切圖片主題。此結果可說明圖片的特徵及使用者的觀感輸入，有助於圖片轉出音樂的貼切性。若將此方法配上不同的電腦音樂創作手法，相信能有更多元的結果

五、樂句資料庫的資料量在程式中負責相當大的音樂產生貼切度及悅耳性。其中，有需要在輸入時就紀錄的「音樂聆聽情緒」。此部分的好壞也會影響最後成品的貼切性。

柒、結論

一、結論與應用

此研究試圖在圖像與音樂之間開闢一條新的道路，使圖像和音樂有了聯結。然而藝術本身是要具有人類情感的作品，只能大致上以人為制定的規則讓電腦實行，音樂的接受度及與圖片的貼近性，則端看使用者看法。本研究所開發的程式具備以下特色：

- (一)使用圖片的特徵
- (二)開發由圖片產生音樂的技術
- (三)產生出的音樂符合部分音樂常規，且具有一定比例悅耳性
- (四)將使用者觀感融入於音樂創作中
- (五)產生出讓使用者認為符合圖片觀看情緒的音樂

在研究時試圖從參考文獻包括：電子音樂、演算式作曲、繪圖美學等，在之中找出更為精確的圖片音樂轉換。但此部份文獻十分缺乏，僅能略為參考，加上請教專家意見，配上創意實踐而成。

本研究由圖片產生出之音樂，其技術可應用於音樂治療方面。音樂可以紓解身心的緊張，使人放鬆，而風景照片也可以使人心曠神怡。因此，本研究能在這方面，結合風景照與其相對應的音樂，雙管齊下的進行治療。

二、未來展望：

- (一)結合類神經網路演算法，進行音樂產生
- (二)藉後設作曲使片段音樂發展成完整樂章
- (三)結合自動配和弦系統，以求更有變化的音樂

捌、參考資料及其他

- 一、陳建文，2002，「彩色影像邊緣萃取之研究」，國立成功大學測量工程學系碩士論文。
- 二、陳若涵，2006，「以音樂內容為基礎的情緒分析與辨識」，國立清華大學資訊系統與應用學系碩士論文。
- 三、王威欽，2005，「運用影像分析實現電腦音樂之研究」，國立台藝大科技藝術研究所碩士論文。
- 四、傅思維，2004，「基於中止式理論的自動譜和弦系統」國立台藝大科技藝術研究所碩士論文。
- 五、色彩在聆聽音樂的歷程中，
http://ed.arted.gov.tw/uploadfile/periodical/1675_pr0001_0160_000640067.pdf
- 六、維基百科，sobel 運算子
<http://zh.wikipedia.org/w/index.php?title=%E7%B4%A2%E8%B2%9D%E7%88%BE%E7%AE%97%E5%AD%90&variant=zh-hant>
- 七、黃嫻嬪，閒談音樂治療，<http://www.geocities.com/bihlian/musicure.html>
- 八、轉貼_數位藝術，<http://zrs001001.spaces.live.com/blog/cns!9B6E75B106971463!249.entry>
- 九、平田 豐，C 程式設計-500 個應用範例技巧大全集
- 十、清華大學 資訊工程學系 張智星教授網站 <http://neural.cs.nthu.edu.tw/jang/>
- 十一、SYTMP，<http://www.geocities.com/labourvanity/sytmp/index.html#example>
- 十二、陳慶瀚，邊緣偵測，<http://140.115.11.235/~chen/course/vision/ch6/ch6.htm>

標準 MIDI 樂器排序表 (GENERAL MIDI INSTRUMENT MAP) - GM

No.	Instrument	No.	Instrument	No.	Instrument
-	Piano 鋼琴	-	Ensemble 合奏	-	Synth Effects 合成聲效
001	Acounstic Grand Piano 平台鋼琴	049	String Ensemble 1 弦樂合奏 1	097	FX 1 (Rain) 雨
002	Bright Acoustic Piano 嘹亮鋼琴	050	String Ensemble 2 弦樂合奏 2	098	FX 2 (Soundtrack) 五度重疊音
003	Electric Grand Piano 電平台鋼琴	051	Synth Strings 1 合成弦樂合奏 1	099	FX 3 (Crystal) 水晶
004	Honkey-Tonk Piano 美國西鋼琴	052	Synth Strings 2 合成弦樂合奏 2	100	FX 4 (Atmosphere) 氣氛
005	Electric Piano 1 電鋼琴 1	053	Choir Aahs 詩詞合唱	101	FX 5 (Brightness) 氣氛
006	Electric Piano 2 電鋼琴 2	054	Voice Oohs 人聲合唱	102	FX 6 (Goblins) 小妖精
007	Harpichord 大鍵琴	055	Synth Voice 合成聲音	103	FX 7 (Echoes) 水滴擴音
008	Clav 古鋼琴	056	Orchestra Hit 管弦合奏	104	FX 8 (Sci-Fi) 科幻
-	Chrom Percussion 音階敲擊	-	Brass 銅管	-	Ethnic 民族樂器
009	Celesta 銅片琴	057	Trumpet 小號	105	Sitar 西塔琴
010	Glockenspiel 鐵琴	058	Trombone 長號	106	Banjo 班鳩琴
011	Music Box 音樂盒	059	Tuba 低音號	107	Shamisen 三味線
012	Vibraphone 電鐵琴	060	Muted Trumpet 悶音小號	108	Koto 古箏
013	Marimba 馬林巴琴(立奏木琴)	061	French Horn 法國號	109	Kalimba
014	Xylophone 木琴	062	Brass Section 銅管合奏	110	Bagpipe 風笛
015	Tubular Bells 管鐘琴	063	Synth Brass 1 合成銅管 1	111	Fiddle 古提琴
016	Dulcimer 德西馬琴	064	Synth Brass 2 合成銅管 2	112	Shanai
-	Organ 風琴	-	Reed 簧管	-	Percussive 敲擊樂
017	Drawbar Organ 電風琴	065	Soprano Sax 高音薩克管	113	Tinkle Bell 清脆鈴聲
018	Percussive Organ 敲擊風琴	066	Alto Sax 中音薩克管	114	Agogo 阿哥哥
019	Rock Organ 搖滾風琴	067	Tenor Sax 次中音薩克管	115	Steel Drums 鋼鼓
020	Church Organ 教堂風琴	068	Baritone Sax 上低音薩克管	116	Woodblock 木魚
021	Reed Organ 簧片風琴	069	Oboe 雙簧管	117	Taiko Drum 日本太鼓
022	Accordion 手風琴	070	English Horn 英國號	118	Melodic Drum 旋律中鼓
023	Harmonica 口琴	071	Bassoon 巴松管	119	Synth Drum 合成鼓
024	Tango Accordion 探戈手風琴	072	Clarinet 單簧管	120	Reverse Cymbal 擴音鈸

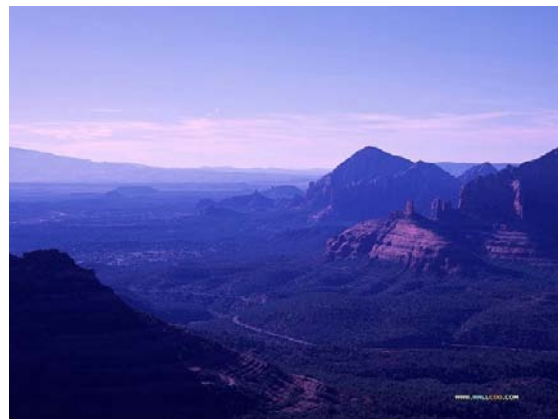
-	Guitar 結他	-	Pipe 管	-	Sound Effects 聲效
025	Acoustic Guitar(Nylon)尼龍結他	073	Piccolo 短笛	121	Guitar Fret Noise 結他擦弦雜音
026	Acoustic Guitar(Steel) 鋼線結他	074	Flute 長笛	122	Breath Noise 吹氣聲
027	Electric Guitar (Jazz) 爵士結他	075	Recorder 牧童笛	123	Seashore 海浪聲
028	Electric Guitar (Clean)清音結他	076	Pan Flute 排笛	124	Bird Tweet 小鳥聲
029	Electric Guitar(Muted)悶音結他	077	Blown Bottle 吹瓶子聲	125	Telephone Ring 電話鈴聲
030	Overdriven Guitar 疊音結他	078	Shakuhachi 尺八	126	Helicopter 直升機聲
031	Distortion Guitar 破音結他	079	Whistle 口哨	127	Applause 掌聲
032	Guitar Harmonics 結他泛音	080	Ocarina 洋埙	128	Gunshot 槍聲
-	Bass 低音	-	Synth Lead 合成主音	-	
033	Acoustic Bass 低音提琴	081	Lead 1 (Square) 正弦波	-	
034	ElectricBass(finger)低音結他-指撥	082	Lead 2 (Sawtooth) 鋸齒波	-	
035	ElectricBass (Pick)低音結他-片撥	083	Lead 3 (Calliope)	-	
036	Fretless Bass 無格低音結他	084	Lead 4 (Chiff)	-	
037	Slap Bass 1 低音結他(拇指敲擊) 1	085	Lead 5 (Charang)	-	
038	Slap Bass 2 低音結他(拇指敲擊) 2	086	Lead 6 (Voice) 人聲	-	
039	Synth Bass 1 合成低音結他 1	087	Lead 7 (Fifths) 五度鋸齒波	-	
040	Synth Bass 2 合成低音結他 2	088	Lead 8 (Bass+Lead)低音+主音	-	
-	Strings 弦樂	-	Synth Pad 合成背景氣氛	-	
041	Violin 小提琴	089	Pad 1 (New Age) 新世紀	-	
042	Viola 中提琴	090	Pad 1 (Warm) 溫暖	-	
043	Cello 提琴	091	Pad 1 (Polysynth) 多音合成	-	
044	Contrabass 低音提琴	092	Pad 1 (Choir) 合唱	-	
045	Tremolo Strings 弦樂顫音	093	Pad 1 (Bowed) 玻璃	-	
046	Pizzicato Strings 弦樂撥奏	094	Pad 1 (Metallic) 金屬	-	
047	Orchestral Harp 豎琴	095	Pad 1 (Halo) 光冕	-	
048	Timpani 定音鼓	096	Pad 1 (Sweep) 刮風	-	

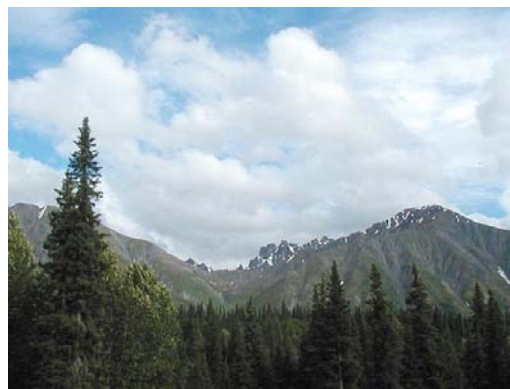
附錄二

此部分為本研究測試所用之圖片。









【評語】 040809

本作品對風景照片作分類，加上兩個主觀判斷的問題，然後依其不同情境產生音樂，是一個很有趣的研究。照片分類只用了風景照片中呈現的天際線作為參數，是太簡單。分析得到的情境與樂曲的對應，目前是依據文獻建立對照表，若能經由夠多被測人實驗得出對照表，會更符合科學研究的精神。