

# 致理技術學院

## 資訊網路技術系 實務專題報告

### 室內／戶外影像之分類

指導老師：高楊達

學生：蘇郁傑(69334104)

陳鴻銘(69334107)

朱紘佑(69334118)

王煒棋(69334142)

中華民國 97 年 1 月

# 致理技術學院

## 資訊網路技術系 實務專題報告

### 室內／戶外影像之分類

學生：蘇郁傑(69334104)

陳鴻銘(69334107)

朱紘佑(69334118)

王煒棋(69334142)

本成果報告書經審查及口試合格特此證明。

指導教師：\_\_\_\_\_

中華民國 97 年 1 月

# 專題研究授權書

本授權書所授權之專題研究為 蘇郁傑 陳鴻銘 朱紘佑 王煒棋

共 四 人，在致理技術學院資訊網路技術系 96 學年度第 1 學期完成資  
網實務專題。

專題名稱：室內／戶外影像之分類

同意      不同意

本組同學共 四 人，皆同意著作財產權之論文全文資料，授予教育部指定送繳之圖書館及本人畢業學校圖書館，為學術研究之目的以各種方法重製，或為上述目的再授權他人以各種方法重製，

不限地域與時間，惟每人以一份為限。

上述授權內容均無須訂立讓與及授權契約書。依本授權之發行權為非專屬性發行權利。依本授權所為之收錄、重製、發行及學術研發利用均為無償。上述同意與不同意之欄位若未勾選，該組同學皆同意視同授權。

指導教師姓名：

專題學生簽名：

(親筆正楷)

學號：

(務必填寫)

中華民國 97 年 1 月 日

# 誌 謝

從三年級下學期開始專題開始分組，到現在本專題能如期完成，首先要感謝高揚達老師這兩個學期以來的熱心教導，一開始不知該如何開始進行，但老師安排了課餘時間對我們進行相關知識的課程輔導，就連寒暑假也和我們組員一起耐心的討論及指正，以致能完成此專題並能呈現給各位，謹此至上最誠摯的謝意。也同時感謝系上諸位師長與同學，熱心提供許多寶貴的意見與指教。

其次感謝郁傑和煒棋的圖片蒐集，在程式方面由鴻銘主導，文章撰寫則是紘佑與其組員一起通力合作，以致我們這組在專題論文的進度上還算順利，這也要感謝各組員的合作。

最後僅此致上最誠摯的謝意與萬分的感激，謝謝曾經幫忙協助我們的每一位，謝謝。

# 摘 要

本篇專題論文共區分為二個部份，第一部份為介紹影像分類、研究範圍與研究動機，第二部份為影像分類的理論、格式與技術研究探討，網路與資訊的蓬勃發展下，早期傳統的影像也跟隨著腳步轉為數位化，使得影像分類判別等等的影像處理技術快速的被發展出來，而這些技術在開發研究的過程中需要大量的影像圖型資料做為資料庫來測試開發使用，本專題論文準備三千餘張數位影像圖片，包括室內影像 1711 張、戶外影像 1688 張，供日後研究測試與實際評估使用。另外本專題論文探討的重點為室內與戶外的影像如何做分類，對影像的特徵做擷取，建立及訓練影像資料庫，分析其影像分類正確與失敗率，以及未來能加以改進的研究方向。

# 目 錄

授權書

誌謝

摘要

目錄

圖目錄

表目錄

第一章 緒論.....	5
第一節 室內/戶外之影像分類的重要性 .....	5
第二節 專題研究動機與目的 .....	7
第三節 研究範圍與報告書架構 .....	8
第二章 理論與技術探討 .....	14
第一節 影像格式.....	14
第二節 Color Space - RGB、YUV 及 HSI .....	18
RGB 彩色模型： .....	18
彩色空間域 YUV： .....	19
HSI 模型(Hue, Saturation and Intensity)： .....	20
RGB 與 YUV 的轉換： .....	21
第三節 影像格式轉換.....	22

第四節 向量量化原理.....	23
第三章 室內／戶外影像分類之實作 .....	26
第一節 影像資料庫架構方法 .....	26
第二節 影像訓練： .....	34
第三節 影像分類： .....	38
第四節 系統介面.....	41
第四章 實驗結果 .....	45
第五章 結論.....	51

## 圖目錄

圖 1. BMP 格式架構.....	16
圖 2. RGB 補色關係圖.....	19
圖 3. YUV 示意圖.....	20
圖 4. HIS-based Color Space.....	21
圖 5. 影像資料庫建構流程.....	26
圖 6. 影像訓練流程圖.....	29
圖 7. 影像分類流程圖.....	30
圖 8. 蒐集圖片流程圖.....	34
圖 9. 格式統一流程.....	35
圖 10. 正規化流程.....	36
圖 11 BCB6 系統介面圖.....	42
圖 12 辨別室內結果介面圖.....	43
圖 13 辨別室內結果介面圖.....	44
圖 14 測試組室內相片.....	48
圖 15 測試組戶外相片.....	49
圖 16 測試組分類錯誤相片.....	50



## 表目錄

表 1. 紅綠藍三色分配表 .....	17
表 2 測試平台設備表 .....	27
表 3 Indoor 及 Outdoor 分類狀況一覽表 .....	46
表 4 Indoor 及 Outdoor 分類狀況一覽表 .....	46
表 5 Indoor 及 Outdoor 分類狀況一覽表 .....	47
表 6 測試組室內相片實驗數據 .....	48
表 7 測試組戶外相片實驗數據 .....	49
表 8 測試組分類錯誤相片實驗數據 .....	50

# 第一章 緒論

## 第一節 室內/戶外之影像分類的重要性

在電腦科技產業蓬勃發展的時代，在知識爆炸與科技日新月異的年代，科技技術不斷的推陳出新，寬頻網路的普及，加速了資料與圖片檔案等數位資訊的交換，在資訊數位化的年代裡，帶動了所有產業對於自身產業的資訊與數位化，在網際網路（Internet）的助力下，所有產業都需要與網路有關。

而在現今一切講求效率與品質的時代，藉由網路可以提供各式各樣的服務，最基本的資料搜尋到電子商務等等，這樣的網路浪潮，的確是為公司或是一般的使用者都省去了很多查詢資料的時間，與其他可能因搜尋所需要的資訊所衍生出的成本，也因為不管是各行各業都前仆後繼的投入這充滿無限商機的網路世界裡，將其相關的商品數位化、資訊化，像是時下年輕人必備的數位相機，或是數位攝影機(DV)、儲存設備(Storage)與掃瞄器(Scanner)，也因這些可將物品數位化或是資料化的物品大量普及後，價格隨之越來越便宜，讓網路的世界從以前的單純文字模式到現在有許許多多精緻或是漂亮的圖片得以在

網路上流傳。

『網路』已成為現代人不可或缺的一部份，也因此許許多多的商家與公司將其商品數位與圖像化，放上網路提供資訊給消費者，讓消費者能在電腦前看到商品原貌與相關資訊，進而挑選自己所需要的商品，這樣的購物消費行態與模式在寬頻網路普及後，逐漸為一般大眾所接受，再者像博物館與圖書館等等擁有豐富歷史文獻與典藏文物的單位，也利用資料數位與圖像化的方式將珍藏的資料分享至網際網路上，使得網路上的資源更豐富與多元，但此時便有使用上的問題浮現出來，因網路的數位與圖像化，使得圖片的種類繁多，分類相當不易，如何將圖像的分類效率與正確性提高的技術，是急須被開發的。

影像處理是一門整合性的科技，影像處理可用於的領域包涵了醫學、數學、生物學、氣象學、工程科學等領域，其應用也廣於日常生活與各學問上，由於電腦設備技術日新月異不斷的提昇，使得影像在電腦上處理的速度與效率不再是最大的問題，也因此影像處理技術在現在是越來越重要。尤其在科技與知識爆炸的時代，WWW 與寬頻網際網路日漸發達，各式各樣的資訊與圖片檔案在網路上不停的流通，對於越來越多的影像在網路上流傳，如何處理與分類影像資料與增加影像分類的效率並正確的分類，是此專題的研究重點與核心。

此專題研究的室內/戶外影像分類，將作為人臉辨識系統的前置

作業，當前置作業處理辨識完室內/戶外的影像分類，接下來就可開始判別人臉是否於室內或是室外，增加影像的分類正確性與效率，此為本專題的研究動機，對於深入了解人臉影像辨識是相當重要的前置處理。

## 第二節 專題研究動機與目的

人類的視覺神經構造與電腦處理判別影像的處理方式不同之處，在於人類能夠對看到的影像清楚仔細的了解影像形狀、顏色及亮度，但無法將影像轉換成可參考的詳細數值，如顏色的深淺與多寡、亮度的強弱等，電腦則可替代人類處理這樣複雜的事，電腦內的影像處理程式可以將影像的各種特徵轉換成詳細的數值，且在處理數值資料的速度上也非常快，若能將人類視覺分辨影像顏色的區分方法套用至電腦，讓電腦能跟人類一樣可仔細的辨識顏色與大小等，對於往後影像辨識在電腦上的應用相信是一種非常大的突破與幫助。

在研究此專題的過程中，我們將透過 BCB6(Borland C++ builder Version 6)程式來撰寫室內/戶外影像分類程式碼，並進而與人臉辨識技術相結合，提高分類與辨識率。近年的影像辨識技術，絕大部份的研究都重視於人臉辨識生物辨識等等，對於室內與戶外之影像類別之技術尚未完全成熟，故有其開發價值，此為研究專題動機。

### 第三節 研究範圍與報告書架構

#### 研究範圍：

本專題研究範圍包括了數位影像資料格式，其中包含了空間域數位影像格式 (spatial domain) 及頻率域資料影像格式 (frequency domain)；檔案種類包含：BMP, JPEG, JPG, PNG, TIFF, ICO；色彩空間 (Color Space)：RGB, YUV, HSI；及影像資料庫的建置，前置處理 (影像格式正規化，色彩空間轉換)。

#### 特徵擷取 (Feature Extraction)：

##### 1 序列式演算法 (Sequential Algorithms) [1]

序列式演算法是一種基礎型的分類方法，其主要是希望能夠有效的彈性運用群集數 (Cluster)，希望能以較少的群集數 (Cluster) 涵蓋所有的資料，以資料本身為質心，算出資料與資料間的距離 (distance)，如果距離相近就當作同一個群集 (Cluster)，如果大於臨界值就新建立一個新的 Cluster，以不超過最大設定的 k 個 Cluster 為限。

##### 2. K-means 演算法 [2]

K-means 演算法是一種傳統的分類比對技巧，其主要是利用同一類向量與其重心彼此間的距離 (distance) 會小於同一

重心與它種類向量間的距離觀念。使用疊代的方式，將資料的向量集合加以分群（須已知族群數目），並以各群的重心作為各自族群的代表，直至分類族群重心收斂為止。

## 向量量化 (VQ)

### 1. LBG 演算法

LBG 演算法 [3] 是由 Linde、Buzo 和 Gray 三位學者所提出的方法。其主要精神是從一堆編碼字的樣本中以分群法找出代表每個群的編碼字，再由這些具代表性的編碼字形成一本編碼簿。

### 2. 標準 VQ 編碼法-Full Search[4]

標準編碼法的做法相當直覺，是將每個編碼字與欲編碼的向量一一計算它們之間的距離 (Euclidean distance)，並從中找出最接近的一組編碼字之索引值來當做該向量的編碼。

目前一些國際組織，相繼討論影像資料庫，針對影像以不同方式分類，儲存和擷取，影像中的物件，不是代表整張影像，就是屬於影像局部有意義的部份，而這些數位影像在影像系統中，依特徵以適當資料模組 (image data models) 表示。

其影像描述方式包含：

字首屬性 (meta attributes) [5] :

此屬性描述影像製作的處理方式及相關資訊，包括影像獲得的日期，影像的檔名及注解等。

語意屬性 (semantic attributes) [5] :

其屬性包含被分析影像的有關主題資訊。

色彩屬性 (color attributes) [5] :

此屬性用來代表色彩亮度的分配統計。

紋理屬性 (texture attributes) [5] :

大都用於紋理特徵集合，為 Haralick' s 灰階

co-occurrence 特徵，其他通常用來量測紋理，分述如下

1. Tamura 特徵：

包含六個特徵，粗糙 (coarseness)，對比 (contrast)，定向性 (directionality)，似直線性 (line-likeness)，規則性 (regularity)，雜亂性 (roughness)。

2. Unser 的和直方圖及差直方圖：

他所提的三十二個特徵，乃依據影像在灰階中，計算不同的和直方圖及差直方圖。

3. Galloway' s run-length 基本特徵：

計算基本的 run-length 矩陣的 20 個係數。

#### 4. Chens 的幾何特徵：

形成二進位影像順序，他所提 16 個係數，是基於影像之門檻。

#### 5. Wagner 概述 18 個不同的方法，包括 318 個的特徵及對每一個單獨紋理特徵的正確公式。

外形屬性 (shape attributes) [5]:

從技術方可區分為二類

##### 1. 量測的方法：

範圍從簡單，基本的計算如區域，由圓形至較複雜之區域，對不同的 moment invariants 計算。

##### 2. 轉換的方法：

範圍從公式轉換 (例如傅立葉描述子 Fourier descriptors)，至結構的轉換 (例如鏈碼 chain codes) 及曲度空間特徵向量 (curvature scale space feature vectors)。

其他方式也可分為下列二類：

##### 1. 全面外形特徵 (global shape features):

是依據整個影像物件的特徵。如在影像上之區域，周圍及主軸方向所符合之影像部份。

##### 2. 局部外形特徵 (local shape features)



基於影像物件的低階特徵，決定於局部特徵，通當需要更多複雜的計算。如圍繞在符合之影像區域邊緣上之曲率，邊界切割及角落點皆為此範圍。

空間屬性 (spatial attributes) [5]：

1. 拓樸集合的關係(topological set):

兩個影像與影像之間的物件或內容拓樸集合，像是兩物件間彼此的關係，例如兩物件是否有重疊，是否連接，是否相鄰，是否交錯等。

2. 向量集合的關係(vector set)：

考量影像物件的關聯位置的向量集合，如將東，南，西，北等方向，以 East，South，West，North 等 4 個方向表示。

3. 計量集合的關係(metric set of relations):

以影像物件及內容間距離關係的計量集合，如接近、遠、非常接近及非常遠。

4. 2D 字串：

每一個影像被視為符號的矩陣，而每一個符號則需符合影像物件。所符合的 2D 字串是由這些符號沿著水平及垂直軸所投射而得，並保存這些物件的相關位置。

## 報告書架構：

本文章節第一章為緒論，主要說明本專題重要性及研究動機；第二章為影像分類原理，介紹影像分類的原理與方法：如影像的正規化、Color Space(RGB、YUV、HSI)相互轉換、特徵擷取、向量量化原理等；第三章為影像資料庫之建置，說明影像資料庫建構方式及考量因素；第四章發表影像分類流程及分享實驗結果；第五章為結論，分析此研究之優缺點及未來改進方向，做為此專題的總結。

## 第二章 理論與技術探討

### 第一節 影像格式

目前影像格式有很多種，其中有一些影像格式藉由影像壓縮技術來轉換，由於影像資料在電腦上所佔的空間體積十分龐大，因此通常會使用壓縮的技術來儲存它或傳送至網路上。這樣不但節省影像資料所佔的記憶體空間，還可以加速處理影像跟傳送的時間。一般壓縮技術跟影像壓縮技術不同，一般壓縮資料的技術解壓縮後跟原來資料會完全一模一樣；然而影像資料壓縮技術則和一般資料壓縮技術不同，本章節主要是介紹影像格式，對於影像壓縮技術就不詳細說明。

影像格式種類有很多，以下我們介紹一般常用的主要幾種格式

[6]：

#### 1. JPG 格式

目前幾乎所有的數位相機都有支援 JPG 影像格式的紀錄，它精準的記錄每一個像素的亮度，並取出平均色調壓縮影像，並且用某一點顏色代表周圍附近的顏色（例如：背景是白或藍色，就可用其中一點顏色，代表附近多點的顏色），如果每一點代表的範圍小、壓縮程度

小，越接近真實影像，但儲存的記憶體則越大，比較接近原始影像；一點代表的範圍大，壓縮程度大，所需記憶體較小，影像越模糊。

## 2. GIF 格式

GIF 檔是以兩種方式壓縮影像檔案，首先，使用編碼技巧，將同一行列間顏色相近的像素當成一個單位。其次，限制檔案的索引色 (indexed color)，也就是說 GIF 檔不得超過 256 色，因為顏色組合少，GIF 格式不適合於相片或高彩度影像儲存。GIF 顏色雖少，但可進行較大的影像壓縮，讓 GIF 檔成為處理簡單線條圖形的最佳選擇。

## 3. TIF 格式

專業設計師多以 TIFF (Tagged Information File Format，簡稱 TIF) 儲存檔案，是一個不失真的 24bit 彩色影像格式，例如：麥金塔電腦處理的檔案，拿到微軟視窗系統也能相容，通常用於高階印刷，顏色比較不失真。TIF 早期的缺點是版本多，而且彼此間不一定相容。在新版 Photoshop、PhotoImpact 和 CorelDraw 都已解決這個問題；但一般人還是習慣用檔案小、低失真的 JPG 檔。

## 4. BMP 格式

任何電腦作業系統都會有主要的影像格式。微軟 Windows 是使用 BMP 格式，支援 24bit 全彩影像(1677 萬色)，也可壓縮影像至 16bit (65536 色)、8bit (256 色)、4bit (16 色) 或 1bit (單色) 索引色

(indexed color)的影像，顏色越多，影像越逼真。BMP 檔是很好的原始影像檔，要儲存圖的檔案容量很大。但 BMP 檔也不是沒有缺點，如果解析度不夠高，影像一經放大，就會出現類似馬賽克的情形。本專題由於需使用 BCB6 軟體來製作程式，所以透過 BMP 格式的圖檔無失真的特點以便撰寫程式增加辨識的效果。接下來詳細介紹 BMP 格式，Bitmap 格式可分成三大部分：檔頭(File Header)、調色盤(Palette)及影像資料(Image Date)。檔頭主要是儲存影像的特性，如影像長寬大小及使用顏色數目等。調色盤則儲存影像每一種顏色的內容，而影像資料則含有影像中每一個像素的值。

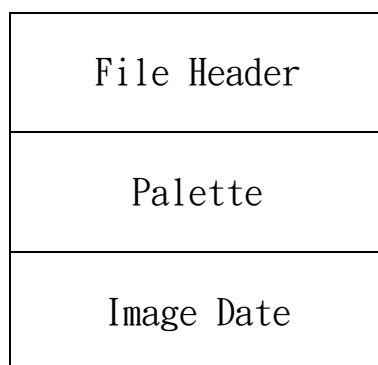


圖 1. BMP 格式架構

舉例說明，如果有一個影像圖形它的高及寬為 100 像素，每個像素可用 256 個顏色，則這些資訊就會儲存於檔頭中。而這些 256 種顏色內容則是包含於調色盤中，每種顏色都是由紅 (Red)、綠 (Green)、

藍(Blue)三原色所組成，所以我們將這些顏色依亮度由暗至亮在調色盤中由底部到頂部排列，如下表所示

表 1. 紅綠藍三色分配表

	紅色	綠色	藍色
顏色 255	顏色 255	顏色 255	顏色 255
⋮	⋮	⋮	⋮
顏色 1			
顏色 0			

在調色盤後是影像資料，其包含每一個像素的數值。在本例中每一個像素都可以使用 256 色，所以每一個像素的值範圍在 0~255 間，而每一個值代表調色盤中的一個顏色，像素 0 為黑色，255 為白色。

## 第二節 Color Space - RGB、YUV 及 HSI

### RGB 彩色模型：

人眼所見各種色彩是因為光線有不同波長所造成的，經過實驗發現[7]，人類肉眼對其中三種波長的感受特別強烈，只要適當調整這三種光線的強度，就可以讓人類感受到『幾乎』所有的顏色。

依 Young - Helmholtz 之說法，視覺感官存在三種基本色之感受視神經。經由光線傳至視神經，視神經傳達至腦部，而產生有顏色的感覺，而這三種光線要素是紅(red)、綠(green)、藍(blue)的感受作用，這三種顏色稱為光的三原色(RGB)。所有的彩色電視機、螢幕都具備產生這三種基本光線的發光裝置。因為這三種光線的混合幾乎可以表示出所有的顏色，因此電腦裡頭就用 RGB 三個數值的大小來標示顏色，每種顏色用 8 bits 來記錄，就是 2 的 8 次方，範圍是 0~255，共 256 種亮度的變化，三種乘起來就有一千六百多萬種變化，這也是我們常聽到的 24 bit 全彩。

因為光線是越加越亮，因此兩兩混合可以得到更亮的中間色：黃(yellow)、青(cyan)、洋紅(magenta)；三種等量相加可得到白色。

至於補色是指完全不含另一種顏色，例如黃色一定是由紅綠兩色合成，完全不含藍色，因此黃色稱為藍色的補色，也就是對比色，從色相圖中可以看到兩個補色隔著白色相對。將兩個補色相加會得到白色。

如圖 2 所示。

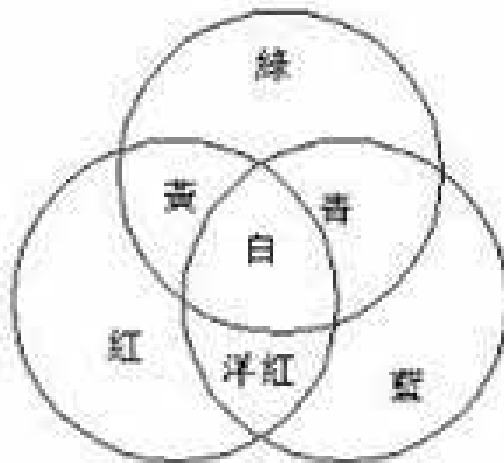


圖 2. RGB 補色關係圖

### 彩色空間域 YUV：

這是在電視工業用的顏色模型。YUV 主要用於彩色視頻信號的傳輸，亦可兼容舊式黑白電視。Y 表示顏色的明亮度 (luminance)，也就是灰階值；U、V 表示顏色的差異「色度」，作用在於描述影像色彩及飽和度。Y「亮度」是由輸入 RGB 的值所創建的，方法是將所接收



到 RGB 的信號的特定部份疊加在一起。U、V「色度」則定義了顏色兩個方面，分別用 U、V 來表示。其中，U 是反映 RGB 輸入信號紅色部分與 RGB 信號亮度值之間的差異。而 V 是反映的是 RGB 輸入信號藍色的部份與 RGB 信號亮度值之間的差異[8]。

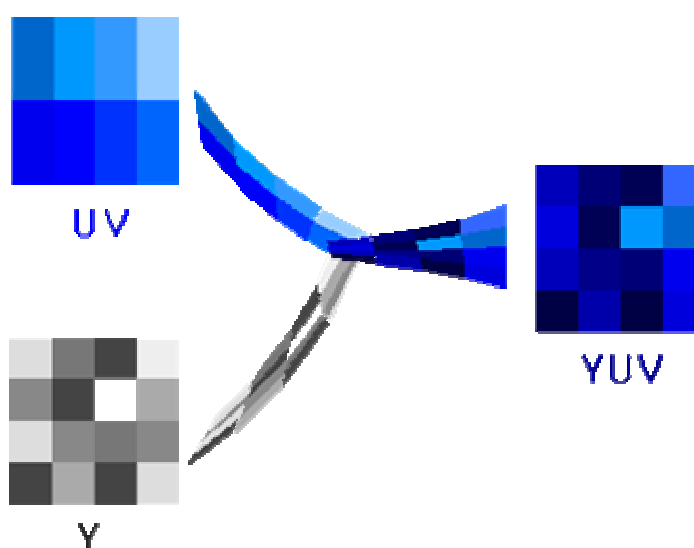


圖 3. YUV 示意圖

### HSI 模型(Hue, Saturation and Intensity)：

此模型是偏向彩色處理的應用[9]，如動畫中的彩色圖形，其中 H(Hue)表示色調，是用來表徵純彩色的一個屬性；S(Saturation)表示飽和度，是用來表示加入白光之後的純淨程度；I(Intensity)表示強度，是表現出此顏色的亮度分量。

這個模型有兩個特點：第一，亮度分量 I 與影像的彩色信息是分離的；第二，色調 H 和飽和度 S 分量與人感受顏色的方式是相連有關的。這些特點使 HSI 模型非常適合於借助人們的視覺系統來感知彩色特性的影像處理演算法的開發。

在圖 4 左邊的圖顯示顏色，右邊的圖則是顯示亮度，此模型的顏色和亮度是分開的。

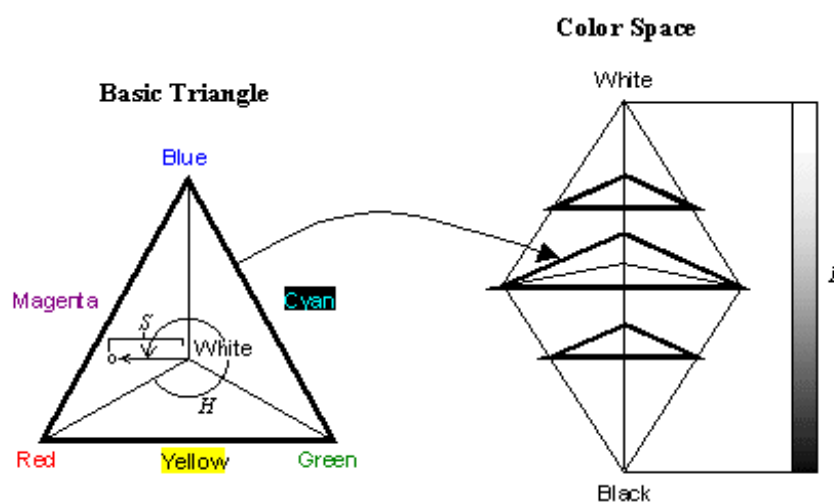


圖 4. HIS-based Color Space

RGB 與 YUV 的轉換：

RGB 轉換 YUV 的公式：

$$Y = 0.299R + 0.587G + 0.114B$$

$$U = -0.147R - 0.289G + 0.436B$$

$$V = 0.615R - 0.515G - 0.100B$$

YUV 轉換 RGB 的公式：

$$R = Y + 1.14V$$

$$G = Y - 0.39U - 0.58V$$

$$B = Y + 2.03U$$

### 第三節 影像格式轉換

在室內/戶外影像辨識中的前置工作前，必須先蒐集影像資料，資料的來源可能是網際網路或者由市集販售具版權使用之多媒體光碟的室內/戶外之圖檔，但手上所蒐集的圖檔資料格式不一，圖檔的檔案類型有 JPEG、TIFF、GIF、TGA 及 PNG 等，可利用 ACDseeV5.0 或 PhotoimpactV7.0 應用軟體將各式檔案格式以另存新檔方式轉換為 BMP 格式檔。

將手上的圖檔格式統一為 BMP 格式檔後，就要進行影像的正規化，所謂正規化是將每一個統一格式後的圖檔，都經由程式之處理過程，將其圖檔的像素大小統一為 100\*100 像素，這樣就可以讓手上的資料庫圖檔的格式與像素大小統一。

有部份圖檔經過正規化後，必須還要辨別影像物件的完整性是否明顯，例如有部份圖檔經由正規化之後，會有影像像素過小的問題 [10]，若正規化後小於 100\*100 像素之圖片就不考慮使用；或者是有邊框的影像，這時就要做去邊框處理，以增加檔案完整性，同時提昇

辨別的正確性。

#### 第四節 向量量化原理

向量量化 (Vector Quantization, VQ)，是用來減少資料庫的樣本數量，這些技術經常使用在影像壓縮及分類，此技術也可應用在本專題。本專題提出一種基於向量量化編碼法的影像檢索技術，作為辨識室內戶外的方法。以下我們先說明一些向量量化的原理。

向量量化編碼法最早由 Gary[11]提出，他先編造依本編碼 (codebook)，此編碼簿是由許多向量，稱之為編碼字 (codeword) 所組成，當針對某一向量編碼時，則直接到編碼簿中尋找與其最接近 (Euclidean distance 來決定) 的編碼字，並以該編碼字位於編碼簿中的索引值來代表原始的向量，如此即完成該項量的編碼工作。解碼方法非常簡單，只要從編碼簿中取出索引值所對應的編碼字即可完成解碼工作。

向量量化 (Vector Quantization, VQ) [12]，有下列二個優點：

- (1) 節省計算空間
- (2) 減少記憶空間

從理論觀點考量的問題：

- (1) 選擇適當的 VQ 演算法，因如果選擇不同的演算法，實際運算

後，經由收斂方式而與原先結果會產生些許誤差。

- (2) 選擇訓練初始樣本(prototypes)數量，如果樣本數量太大時，將失去快速處理的效果；如果太小時，所選擇的樣本將無法正確的與輸入的分布相吻合，容易提高錯誤率。所以初始樣本的數量必須測試評估。

在 VQ 許多演算法中我們舉其中一個演算法(LBG 演算法)為例，作為室內戶外訓練，步驟如下：

步驟 1：

將影像資料庫中所有室內戶外的影像各找出分別 100 張室內戶外影像圖片。

步驟 2：

將選取出來的圖片分別切割成  $10 \times 10$  的區塊，由於我們先前會先將圖片正規化為  $100 \times 100$ ，所以每張圖為  $100 \times 100$  的影像，一張圖則共可被切割成  $\frac{100 \times 100}{10 \times 10} = 100$  個區塊。每個區塊針對 YUV 各別取出 YUV color moment(mean and standard deviation)的特徵，將這 3 維 Y mean、U mean、V mean 視為一個向量。

步驟 3：

再將步驟 2 所取得的向量隨機抽取 500 個區塊來當作初始編碼簿，將其餘區塊以這 500 個區塊的初始向量為重心分別歸群到這 500

個重心，也就是尋找與其距離最近的向量，以形成 500 個群，歸納成一叢集。公式如下：

$$d = \sqrt{(Ym_i - Ym_k)^2 + (Um_i - Um_k)^2 + (Vm_i - Vm_k)^2}$$

$Ym_i$  : Y 向量初始 mean 值

$Ym_k$  : Y 向量第 k 個 mean 值

$Um_i$  : U 向量初始 mean 值

$Um_k$  : U 向量第 k 個 mean 值

$Vm_i$  : V 向量初始 mean 值

$Vm_k$  : V 向量第 k 個 mean 值

I 值：為隨機所取的 500 個向量，其範圍 1~500

K 值：為原有訓練 20000 個向量，其範圍 1~20000

步驟 4：

重新計算這 500 群的重心，形成一個新的編碼簿。重複步驟 3 及 4，直到每個群重心變化不再很大，即完成了編碼簿的訓練工作。

### 第三章 室內／戶外影像分類之實作

#### 第一節 影像資料庫架構方法

以下為影像資料庫建構流程：

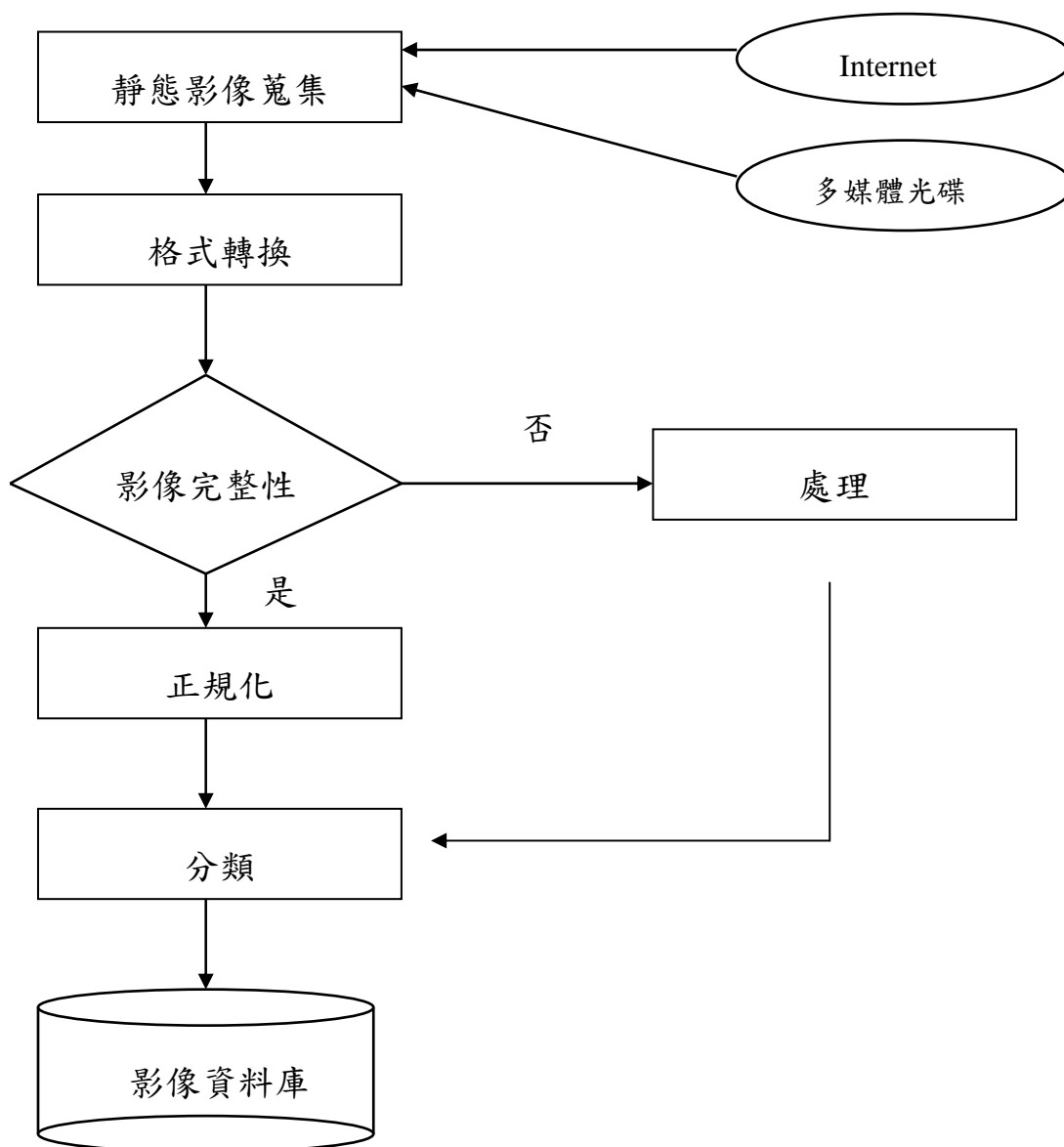


圖 5. 影像資料庫建構流程

表 2 測試平台設備表

廠牌	MACBOOK
CPU	Intel(R)1500 2.00GHZ
RAM	1GB
作業系統	WINDOWS XP

1. 影像蒐集：

影像蒐集的來源可從網際網路搜尋下載，或購買市集販售有版權的多媒體光碟圖檔，共蒐集約 2000 張影像圖檔。

2. 格式轉換：

將蒐集的各式圖檔 (TIFF, GIF, TGA, PNG, JPG, BMP)，經由 Photoimpact V7.0 軟體以另存新檔方式，轉換為 BMP 檔。

3. 判斷影像完整性：

在統一檔案的格式之後，要判斷影像的完整性，這個動作是為了考量下一步影像正規化後，其目標是否明顯，特別是動物特寫的图片，一經正規化後，目標會縮小而不易辨識，這個圖片檔案必須做裁剪影像的處理後，進行正規化才有意義。

我們的正規化動作，除了讓檔案格式統一之外，經由程式的處理，讓每一個圖檔像素的大小要統一成 100\*100 像素，但還有一些圖片正規化之後，會造成失真的情形 (像素過小)，



這種例子的圖檔就不考慮使用。

#### 4. 正規化：

正規化之影像，其固定長或寬之解析度是以相對比率長度，以保持圖像在放或縮小時不會失真。

#### 5. 分類：

因應本專題之主題，影像資料庫分類的方式採樹狀結構分類，根部將影像分為 Indoor 及 Outdoor 兩種，再依其主要特徵，細分至子目錄。

Indoor：辦公室、住家、餐廳、房間、室內運動場、會議室、劇場及醫院等。

Outdoor：動物、植物、山、水、建築物、天空、森林、海洋等。

本專題實作研究重點是參考 Vailaya 等人 [13] 所提出以 Color moment 影像特徵，向量量化(Vector Quantization :VQ)，取出訓練圖片的平均 YUV 值，再將需辨別的圖檔經由所做的程式計算，將兩數值做比較，辨別是否為室內或戶外。

影像分類的方法，我們分作兩個階段來探討，第一階段為影像的訓練，將影像正規化後（100\*100 像素），再分割成 100 個 10\*10 的

小區塊，每個小區塊對 YUV 各取 mean 因此取得 3 維的色彩矩 (Color moment) 特徵，每張影像共有 300 維度的特徵，其細節下面章節會詳細介紹；第二階段為待測影像的分類，將待測的影像經由影像訓練的前置工作，取得該張影像的 300 維度之特徵即  $Y_{mean}$ 、 $U_{mean}$ 、 $V_{mean}$  各 100 個值，分別和訓練完成的 Indoor/Outdoor YUV 向量做比對，也就是需要辨別的圖的  $Y_{mean}$ 、 $U_{mean}$ 、 $V_{mean}$ ，分別和訓練完成的室內／戶外的  $Y_{mean}$ 、 $U_{mean}$ 、 $V_{mean}$  做比對，若相似度室內較多者，則判定為室內圖片，若相似度戶外較多者，則判定為戶外圖片。

其流程圖如下所示：

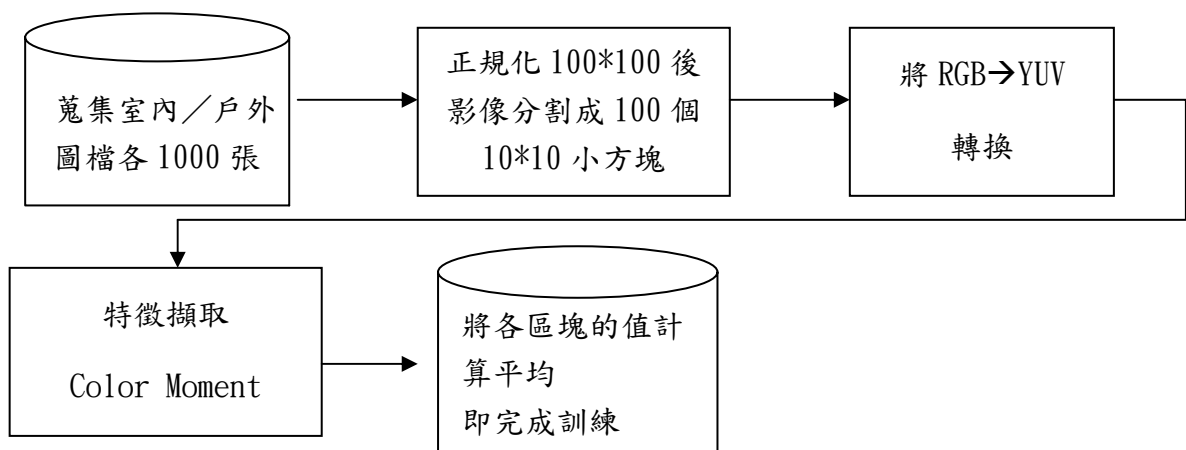


圖 6. 影像訓練流程圖

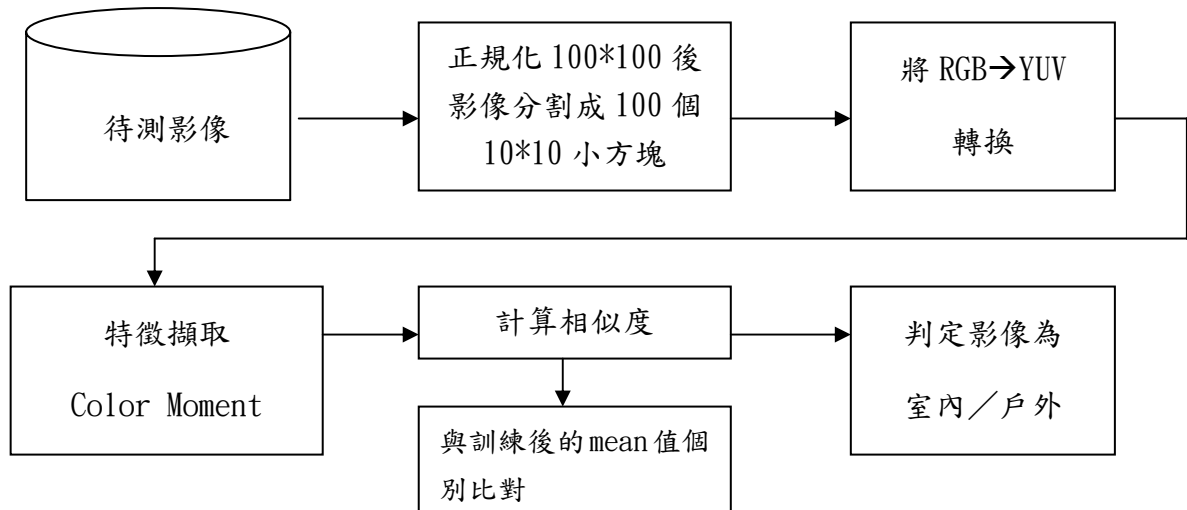


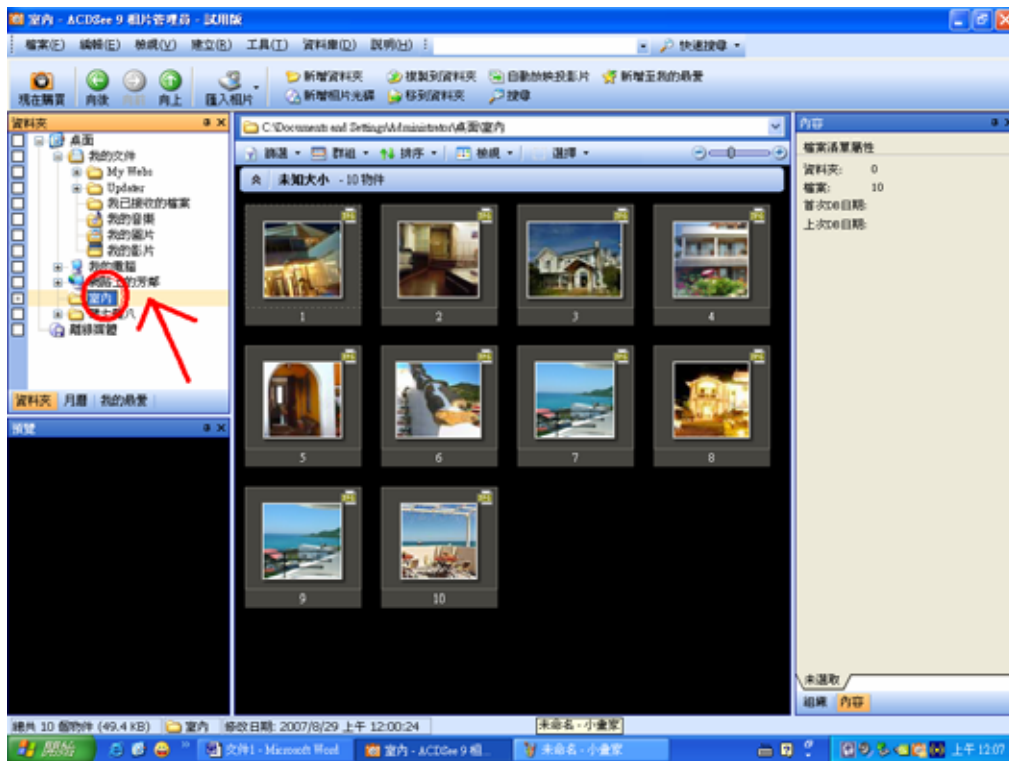
圖 7. 影像分類流程圖

現在介紹蒐集圖片的前置工作：

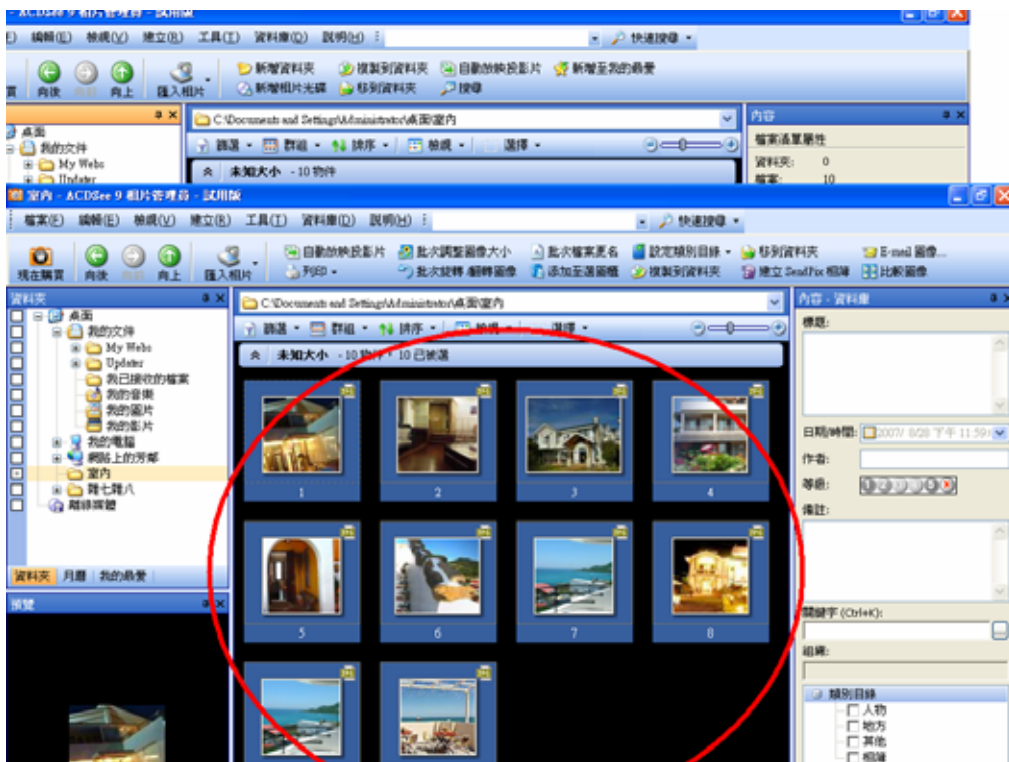
1. 開啟 ACDSSee 9 軟體：



2. 找尋檔案位置. 點選



### 3. 全選所有圖片



### 4. 更改檔案格式為 .bmp



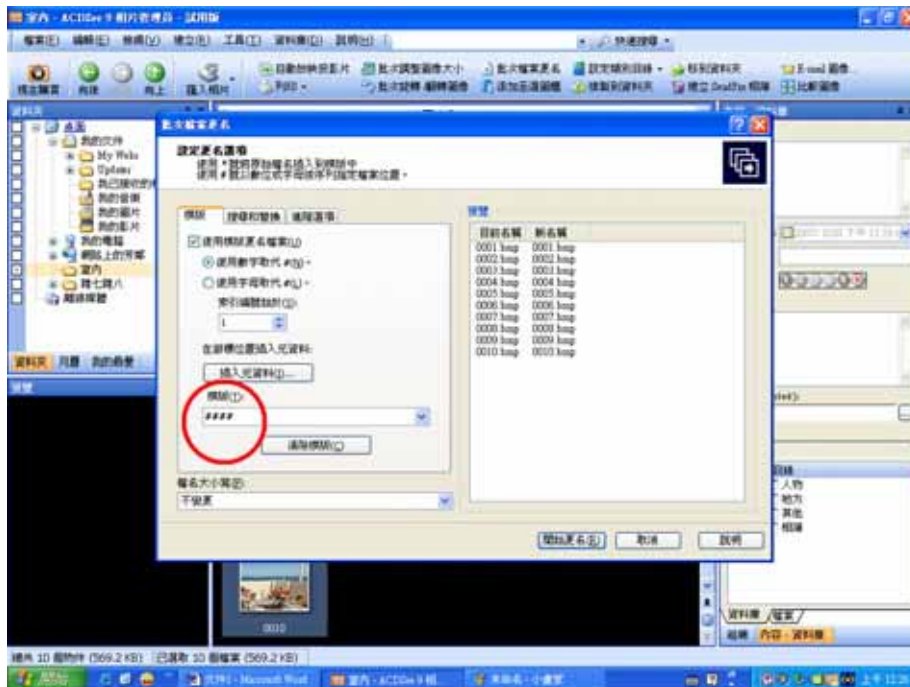
工具-轉換檔案格式-點選 bmp(開始轉換 全部皆是...)



5. 編輯 bmp 檔名為四碼



輸入 4 個 # 為 4 碼-開始更名



5. 將圖片大小設為 100\*100

點選 批次調整圖像大小-(寬設 100 高 100)



本專題實作的前置作業部份步驟如下：

## 第二節 影像訓練：

### (1) 蒐集圖片：

為了將影像訓練的參考值更為精準，必須將蒐集好的圖檔依室內／戶外以人工方式事先分類好，我們所蒐集的室內／戶外圖檔大約各一千張左右。

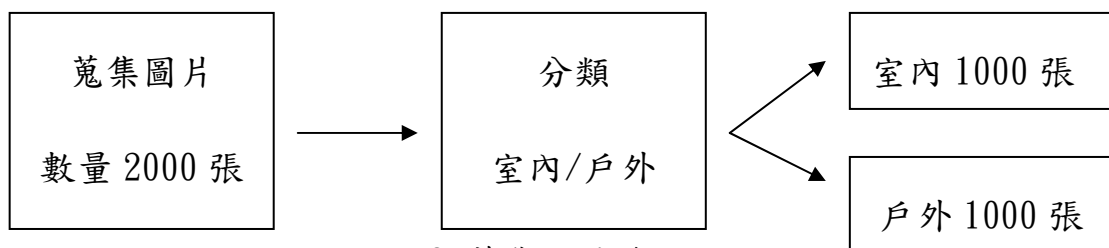


圖 8. 蒐集圖片流程圖

(2) 檔案格式統一：

這些蒐集並分類好的圖檔，檔案格式可能不一，有可能是 TIFF, GIF, TGA, PNG, JPG, BMP 等格式，由 ACDsee9.0 軟體用另存新檔的方式，將這些格式不一的檔案轉換成 BMP 檔，為什麼要轉換成 BMP 檔呢？因為使用無壓縮的 BMP 影像格式較適合做影像分析。

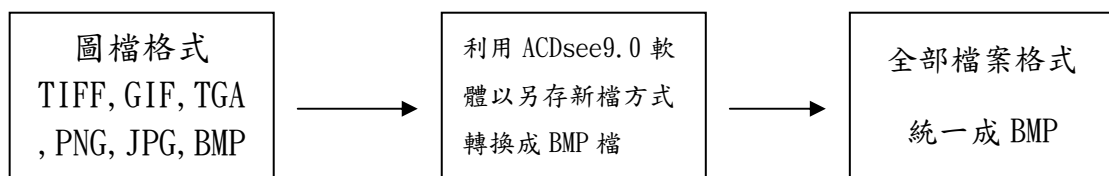


圖 9. 格式統一流程

(3) 影像的正規化：

將這些檔案轉換成 BMP 檔後，需要考慮影像像素的大小。影像的大小直接影響處理速度，較大的影像處理速度較慢，所以我們必須將影像像素大小統一，經統一格式後的檔案，要正規化成 100\*100 像素，因程式系統需要將圖片切割成等份 10\*10 的區塊；若是未經正規化的圖檔像素過小（未達 100\*100 像素），因為無法切割等份 10\*10 區塊，這些檔案就



無法做訓練，這一點是必須注意的。

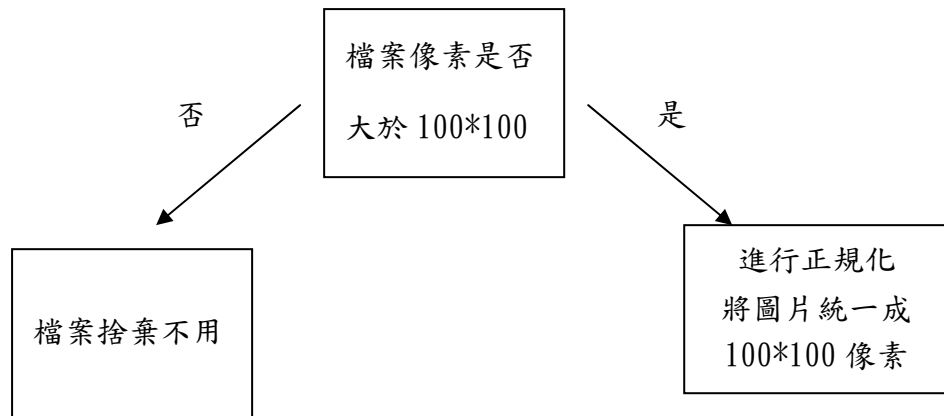
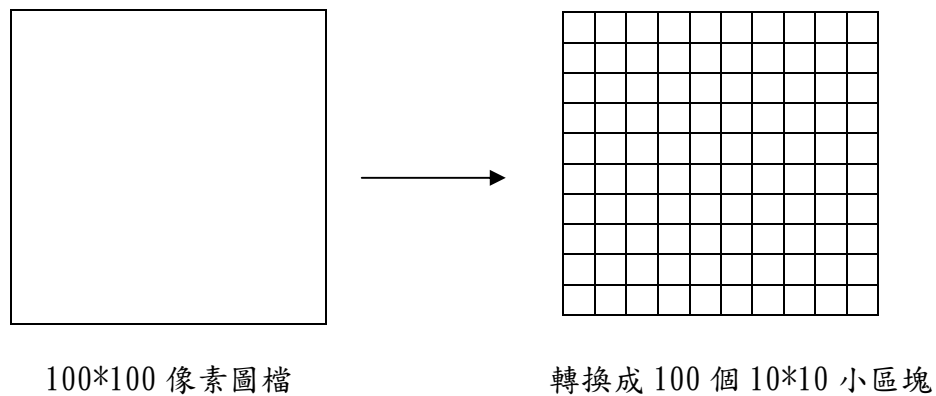


圖 10. 正規化流程

- (4) 檔案格式以及正規化的動作都完成後，就要利用 BCB6 程式來撰寫，將這些 100\*100 像素的圖檔，分割成 100 個 10\*10 像素的小區塊。



影像分割的動作完成後，接著要考慮色彩因素，一般而言，對於室內／戶外影像，在光線的因素及色彩的分佈上大為不同。所以

我們使用色彩的特徵進行比對處理，我們在此專題研究採用YUV Color Space，因為一般來說，室內的相片大多光源不充足，即使經由日光燈的照射，其光源分部也不均勻，亮度也無法達到戶外的水準。反之戶外因陽光的照射光源分部比較平均而且亮度比較亮，因此對於辨識室內戶外是一個很重要的特徵。因為Y代表光亮度元素 U、V代表兩個彩度元素，對亮度的分離性高，方便與彩度分開操作。根據研究顯示[14]，人眼察覺到的光亮度資訊(Y)的60%~70%來自於綠色光，紅色和藍色實際上只是亮度資訊的複製，因此這些重複的資訊可以完全去除，所以我們將影像經由公式及程式的處理，把RGB Color Space轉換成YUV Color Space。

RGB 轉換 YUV 的公式：

$$Y = 0.299R + 0.587G + 0.114B$$

$$U = -0.147R - 0.289G + 0.436B$$

$$V = 0.615R - 0.515G - 0.100B$$



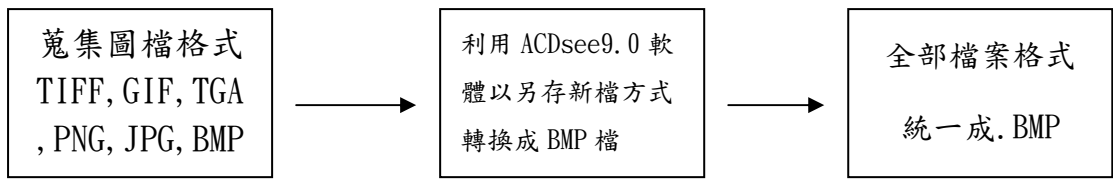
接著再把每個10\*10像素的小區塊，用程式將每個區塊中取出每個像素的Y、U、V值，全部累加；累加後的總值，在除以整個區塊的像素數目也就是10\*10(100)，計算區塊的 $Y_m$ 、 $U_m$ 、 $V_m$ 這3維的值，這樣每個圖檔有100個小區塊，把這每個小區塊依序由切割的陣列排列編號(0~99)，每個小區塊有3維的值，總共一個圖檔有300維，如此室內的1000個檔案都經由程式計算過後，依序再將0號至99號區塊的3維值做平均計算(各除1000取平均值)，這樣室內的圖檔的3維參考數據就可訓練完成；戶外亦然。訓練完成後的數據結果室內Y值、U值平均比戶外低，而V值則是平均比戶外高，由此可知訓練圖片中亮度平均是戶外較高，代表U值的彩度也是戶外較高，只有V值代表的彩度為室內較高。

### 第三節 影像分類：

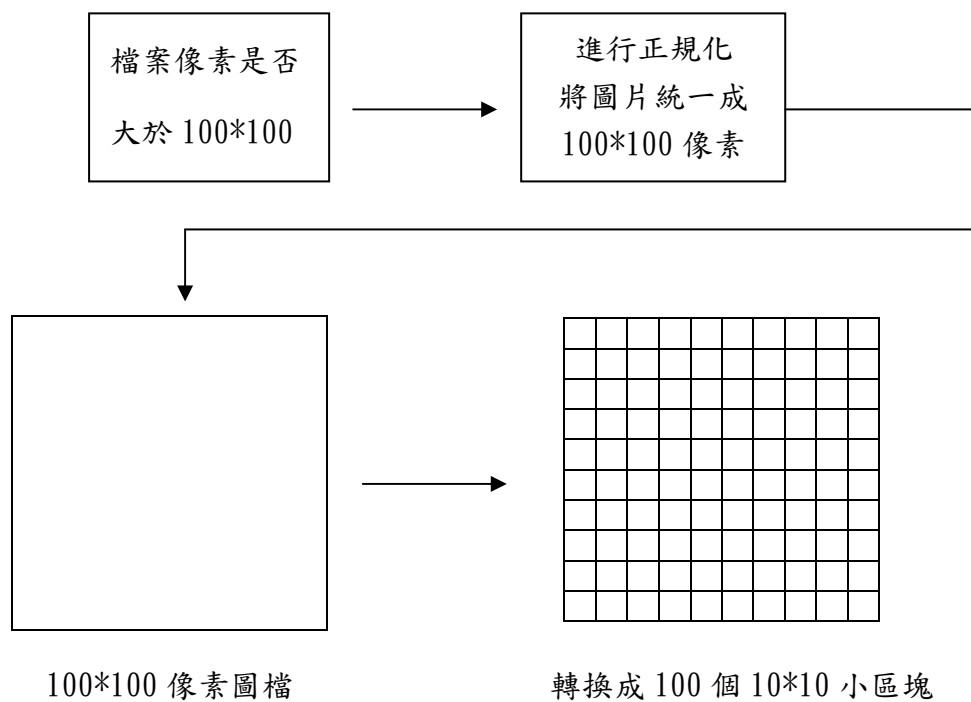
影像訓練完成後，接著可以做分類的動作，流程如下：

- (1) 將待測影像集中後，和訓練影像的前置作業相同，將各種格式之圖檔，使用 ACDsee9.0 軟體將各種檔案格式轉成 BMP

檔。



(2) 將這些檔案格式轉換完畢的 BMP 檔，正規化為 100\*100 像素，並將每個 100\*100 像素的 BMP 格式檔，分割成 100 個 10\*10 的檔案。



(3) 取出圖檔中的 RGB Color Space，將其 RGB 利用公式轉換成

YUV Color Space，公式在之前已介紹過，在此不贅述。

把 100 個 10\*10 像素的小區塊，用程式分別算出每個區塊的 Y mean、U mean、Vmean 這 3 維的值，一個完整的圖共 300 維，分別和訓練完成 Indoor/Outdoor 的 Y mean、U mean、Vmean 做比較的動作，如果辨別的圖的 Y mean、U mean、Vmean 與訓練完成 Indoor 的 Y mean、U mean、Vmean 較相近則判定為室內，反之亦然。比對公式如下：

$$d_1 = \sqrt{(Y_{m_i} - Y_{m_k})^2 + (U_{m_i} - U_{m_k})^2 + (V_{m_i} - V_{m_k})^2}$$

$$d_2 = \sqrt{(Y_{m_o} - Y_{m_k})^2 + (U_{m_o} - U_{m_k})^2 + (V_{m_o} - V_{m_k})^2}$$

$d_1$  值：需辨識圖片的 indoor 相似度

$d_2$  值：需辨識圖片的 outdoor 相似度

$Y_{m_i}$ ：訓練完成 Indoor 的 Y mean 值

$U_{m_i}$ ：訓練完成 Indoor 的 U mean 值

$V_{m_i}$ ：訓練完成 Indoor 的 V mean 值

$Y_{m_k}$ ：需辨識圖片的 Y mean 值

$U_{m_k}$ ：需辨識圖片的 U mean 值

$V_{m_k}$ ：需辨識圖片的 V mean 值

$Y_{m_o}$ ：訓練完成 outdoor 的 Y mean 值

$U_{m_o}$ ：訓練完成 outdoor 的 U mean 值

$V_{m_0}$ ：訓練完成 outdoor 的  $V_{mean}$  值

如果  $d_1$  大於  $d_2$  則辨別為室內； $d_2$  大於  $d_1$  則辨別為戶外。如此即可將影像分類完成。

#### 第四節 系統介面

軟體方面我們採用 BCB6 利用 C 語言撰寫程式，其介面如圖 6 所示，圖 7、圖 8 為辨識結果展現方式。

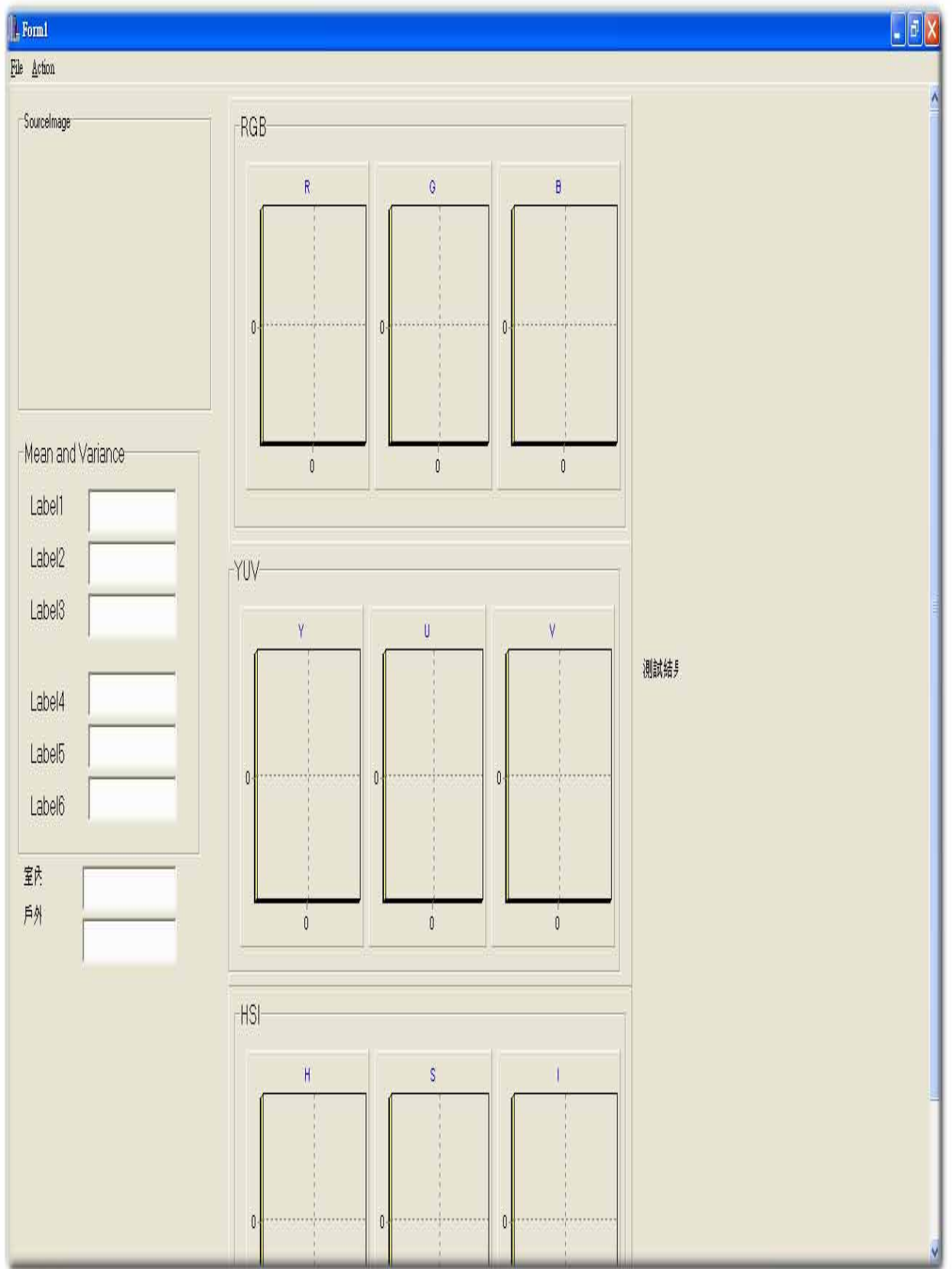


圖 11 BCB6 系統介面圖

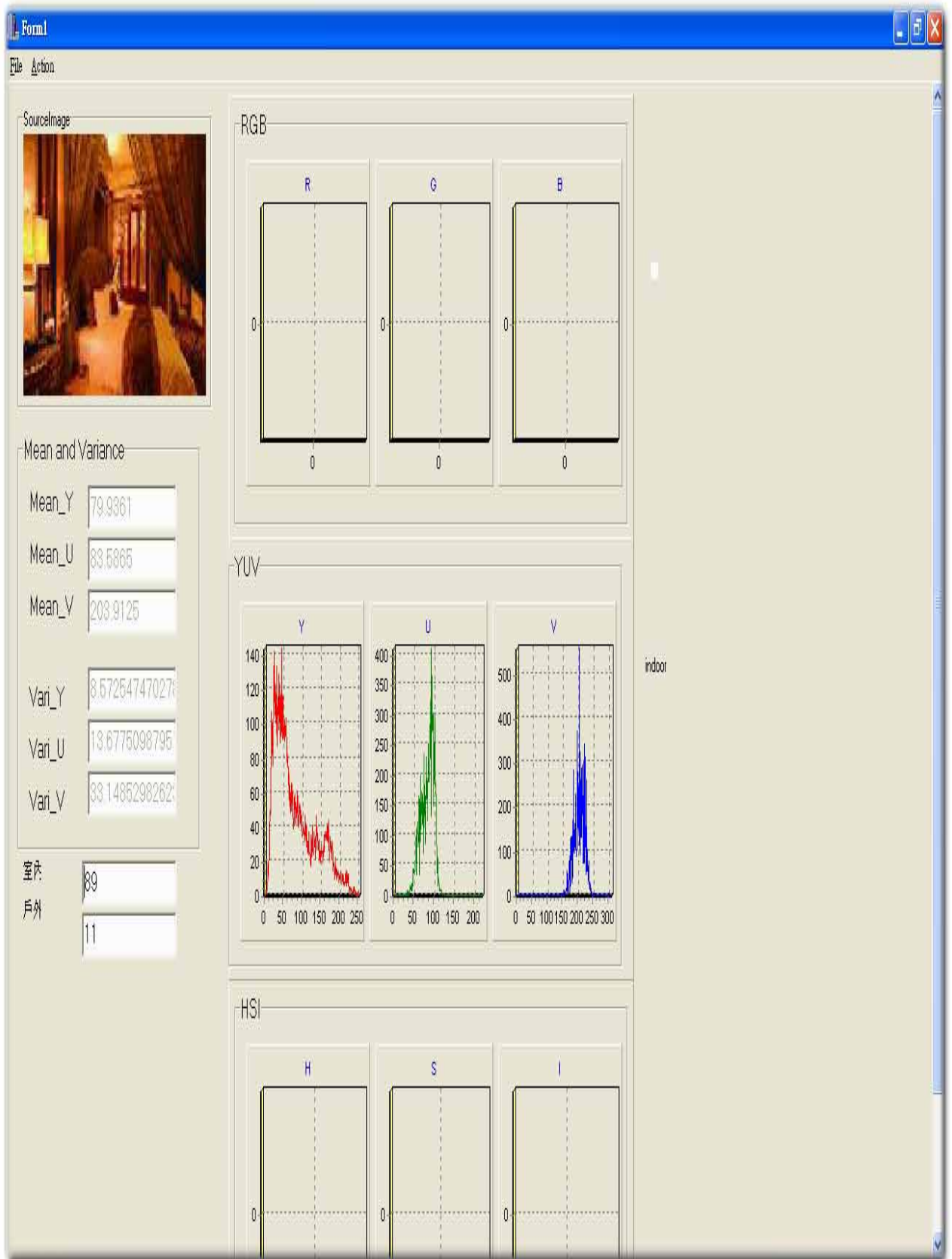


圖 12 辨別室內結果介面圖



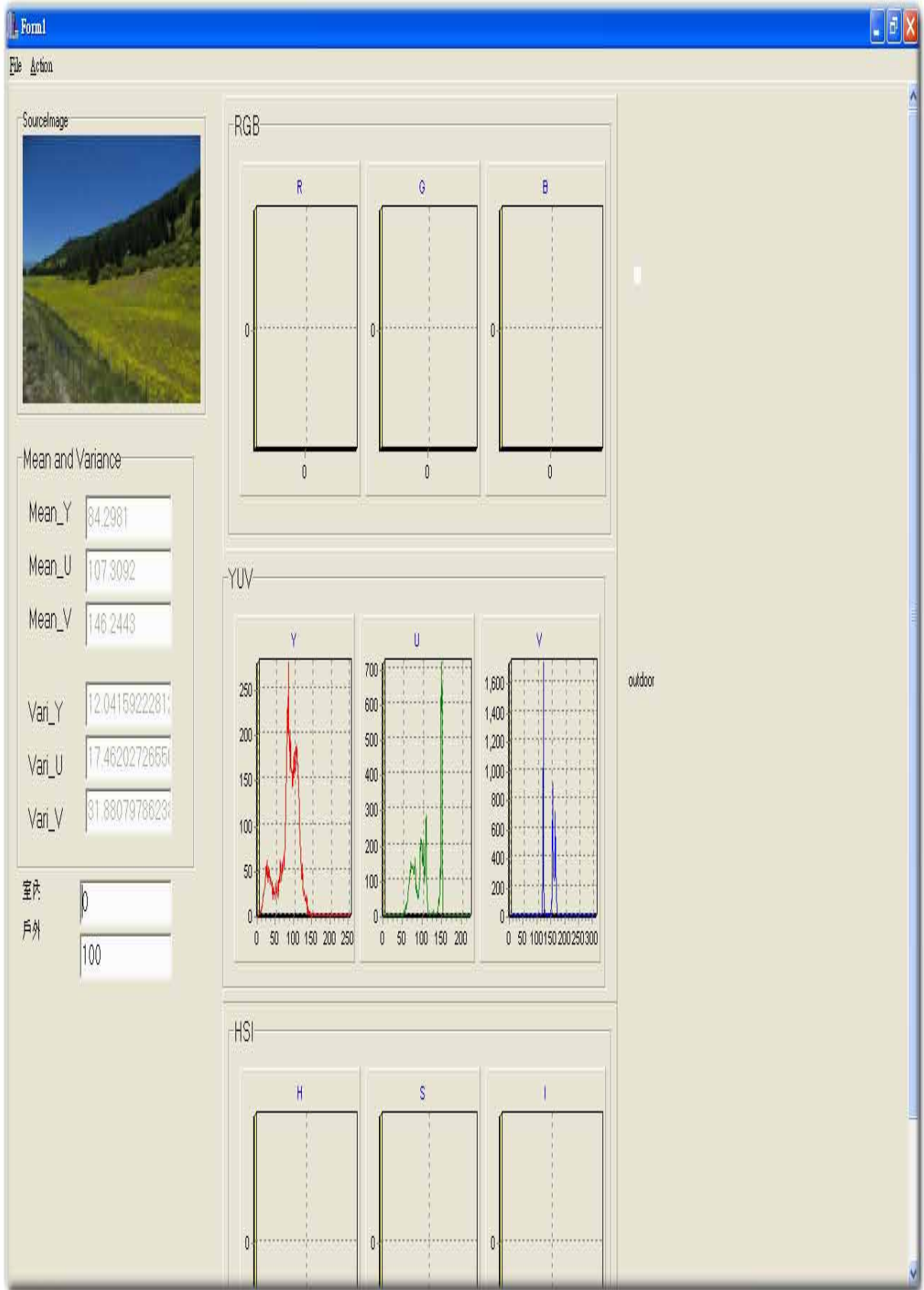


圖 13 辨別室內結果介面圖

## 第四章 實驗結果

我們所蒐集的圖片，經由人工事先分類後，將影像分別標示 indoor/outdoor，一共蒐集了室內圖片 1711 張，戶外圖片 1688 張，總共蒐集了 3399 張圖，我們使用電腦隨機抽樣，indoor/outdoor 圖片各 1000 張為訓練用樣本，其餘圖片 1399 張，包含 indoor711 張，outdoor688 張，則用來測試實驗數據。

在這上述 3399 張影像中，我們使用規格如表 2 筆記型電腦上進行分類測試，為了可以確實辨識出室內/戶外的照片。我們將我們所使用的照片分成二大類，為訓練組相片、測試組相片，其中又個別分為室內部分跟戶外部分的訓練組跟測試組。首先我們從資料庫中選取訓練組室內、戶外各 500 張照片，以及測試組室內、戶外各 200、400、600 張照片為一個階段。第二個階段我們從資料庫中選取訓練組室內、戶外各 750 張照片，以及測試組室內、戶外各 200、400、600 張照片。第三個階段我們從資料庫中選取訓練組室內、戶外各 1000 張照片，以及測試組室內、戶外各 200、400、600 張照片。為了測試我們所做的方法是否真正適用，在測試組中的照片我們甚至選取一些原先並沒有在訓練組中的相片類型。結果我們發現訓練影像資料庫，影像訓練的張數愈多，則分類成功率愈高，但測試圖片數量的多寡，也

會影響成功分類的正確性。其測試結果如表 3 所示，由實驗數據可以證明我們使用方法的效果。

表 3 Indoor 及 Outdoor 分類狀況一覽表

	Indoor image 測試張數：200		Outdoor image 測試張數：200	
	判定為 Indoor	判定為 Outdoor	判定為 Indoor	判定為 Outdoor
影像資料庫 訓練張數				
500	61%	39%	28%	72%
750	71%	29%	35%	65%
1000	76%	24%	34%	66%

表 4 Indoor 及 Outdoor 分類狀況一覽表

	Indoor image 測試張數：400		Outdoor image 測試張數：400	
	判定為 Indoor	判定為 Outdoor	判定為 Indoor	判定為 Outdoor
影像資料庫 訓練張數				
500	61%	39%	37%	63%
750	66%	34%	33%	67%
1000	63%	37%	36%	64%

表 5 Indoor 及 Outdoor 分類狀況一覽表

影像資料庫 訓練張數	Indoor image 測試張數：600		Outdoor image 測試張數：600	
	判定為 Indoor	判定為 Outdoor	判定為 Indoor	判定為 Outdoor
500	59%	41%	38%	62%
750	61%	39%	36%	64%
1000	68%	32%	35%	65%

我們將部份的實驗結果圖示如圖 9、圖 10 為測試組的室內、戶外相片，圖 11 為測試組相片判斷錯誤的。

我們嘗試解釋其分類錯誤的原因，圖 11 左上方及右上方兩張照片，因為日光燈過於明亮，且牆壁為白色，造成反光使得光源過於比一般室內光源較充足，因此被錯誤的分類成戶外的照片；而圖 11 左下方及右下方的兩張照片雖為戶外的照片，但左下方的照片為近照周圍呈現黑色的造成光源不足，且也沒明顯的戶外特徵，因此被錯誤分類成室內的相片。另外右下方的照片明顯的感覺到與一般戶外的相片光源亮度還要不足所以也被錯誤分類到室內的相片。



圖 14 測試組室內相片

表 6 測試組室內相片實驗數據

indoor 100%	indoor 93%	indoor 100%
indoor 93%	indoor 100%	indoor 100%



圖 15 測試組戶外相片

表 7 測試組戶外相片實驗數據

outdoor 96%	outdoor 94%	outdoor 98%
outdoor 93%	outdoor 65%	outdoor 76%

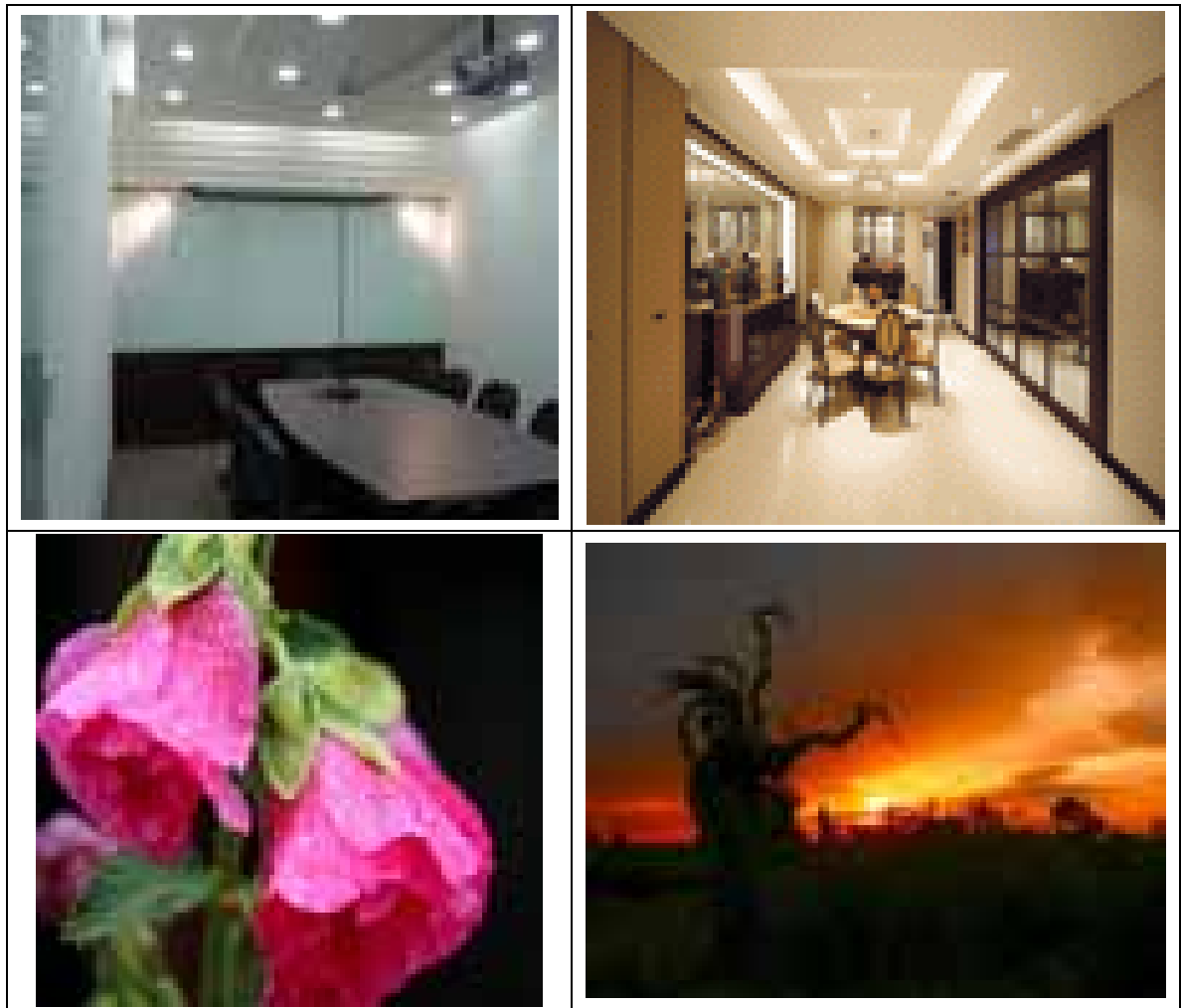


圖 16 測試組分類錯誤相片

表 8 測試組分類錯誤相片實驗數據

outdoor 77%	outdoor 63%
Indoor 60%	Indoor 54%

## 第五章 結論

從實驗結果來看，我們可以發現要提高 Indoor/Outdoor 分類的成功率，在影像訓練的張數和測試張數多寡就可以看的出來，當訓練張數及測試張數愈多，可提高分類的正確率，但訓練和測試張數數量增加愈多，分類的成功率並不會呈正比提高。

從第四章的表格資料中，能夠發現訓練張數和測試張數的比值愈高，則分類的成功率會提高，以室內的圖片為例，同樣訓練 1000 張圖片的資料庫來做測試，測試 200 張的成功率為 76%；測試 400 張的成功率為 63%；測試 600 張的成功率為 68%。

由於本專題在影像訓練所擷取的特徵，是取圖片中的 YUV 特徵，若想提高影像分類的成功率，必須擷取圖片更多的特徵，除了色彩外，還有紋理及外形等等，才能讓分類的精確度及效能提昇。

未來希望能針對窗邊的室內相片，能克服其它外在光源的問題，讓它誤認為戶外相片的機率能降到最低；在戶外相片中能克服夜晚及光線不足的問題，以達成最準確的辨識率。待未來此技術發展成熟後，我們可以將辨識出的結果儲存於資料庫中，將來以便於系統搜尋，同時也可提升搜尋的準確率，除此之外，還可再外加其他的註釋分類項目，擴充系統的便利性。由於現今網路發達，因此我們也可將



系統網路化，讓使用者可以利用我們的辨別方法，在網路上能很快的尋找到使用者所需要的圖片或相片，增加網路搜尋的便利性。

## 參考文獻

- [ 1 ] Shyi-Chyi Cheng and Ping-Huang Lee, " Content-based Access of Image and Video Libraries, 2000. Page(s):15-19
- [ 2 ] Balmelli, L. ; Mojsilovic, A. , " Wavelet domain features for texture descriptionm classification and replicability analysis" ,. ICIP99. Image Precessing, Volume:4, 1999, Page(s):440-444
- [ 3 ] Y. Linde, A. Buzo and R. M. Gray, " An Algorithm for Vector Quantizer Desing, " IEEE Transactions on Communications, Vol. 28, January 1980, pp. 84-95
- [ 4 ] Y. C. Hu, and C. C. Chang, " A Progressive Codebook Training Algorithm for Image Vector Quantization, " Proceedings of the Fifth Asia-Pacific Conference on Communications and Fourth Optoelectronics and Communications Conference(APCC/OECC 99' ), Beijing, China, 1999, Vol. 2, pp. 936-939
- [ 5 ] Grosky W. and Stanchev P., An Image Data Model, in Advances in Visual Information Systems, R. Laurini(edt.)Lecture Notes in Computer Science, Vol.

1929, (2000), pp. 14-25.

- [ 6 ] [Http://mail.mcjh.kl.edu.tw/~lin/teach/picture/picture.htm/](http://mail.mcjh.kl.edu.tw/~lin/teach/picture/picture.htm/)
- [ 7 ] 林明毅、李慶長、呂韶宜，YCbCr 色彩空間之資訊隱藏應用系統研究，臺北商業技術學院，2006。
- [ 8 ] <http://topic.csdn.net/t/20020201/09/512411.html>
- [ 9 ] 影像處理及辨識理論與實務(林金寶編著)，2003 宏有出版。
- [ 10 ] 劉大仁，影像資料庫之建立與分類，義守大學碩士論文，2004，Page. 21-30。
- [ 11 ] 電子影像技術(張真誠、黃國峰及陳同孝編著) 2003 旗標出版，Page. 35-55。
- [ 12 ] 劉大仁，影像資料庫之建立與分類，義守大學碩士論文，2004，Page. 15-18。
- [ 13 ] A.Vailaya, M. Figueiredo, A. Jain, Hong Jiang Zhang. "Ayesian Framework for Hierarchical Semantic Classification of Vacation Images." IEEE International Conference on Multimedia Computing and Systems, Page(s):518-523 vol.1999.
- [ 14 ] 蔣佳欣，室內/戶外與建築物/自然風景之影像分類研究，南

