

整合多搜尋方法之影像資料庫檢索系統

曾逸鴻* 黃吉緯

大葉大學資訊管理學系

51591 彰化縣大村鄉學府路 168 號

摘要

現存的影像檢索系統，大多是針對特定影像資料庫，抽取適當特徵值用以尋找相近的影像；然而，同樣的搜尋方法不一定適用於其他類型資料庫。以顏色特徵值為基礎的搜尋方法，分別應用在黑白、灰階及彩色影像資料庫中，檢索效果有明顯的不同；而以形狀特徵值為基礎者，則適用於前景物體存在且和背景環境差異分明的影像資料庫。有時，使用者又希望直接以影像內容的主觀描述（如影像分類、標的類型、拍攝環境等）來檢索相關影像。因此，本研究分別以影像的顏色特徵、形狀特徵與內容特性，來建構不同的影像搜尋方法。在影像資料庫建構過程，所有影像利用改良式 K-means 演算法進行分群，並於分析各群集中影像特徵之變異程度，於檢索時給予各個群集不同權重，開發出一套整合多搜尋方法並適用不同影像資料庫之檢索系統。經過完整實驗步驟之測試，可顯示整合三種搜尋方法，所得之影像檢索效果最佳。

關鍵詞：多搜尋方法，影像檢索系統，特徵分群

An Image Retrieval System Obtained by Integrating Multiple Searching Approaches

YI-HONG TSENG* and CHI-WEI HUANG

Department of Information Management, Da-Yeh University

No. 168, University Rd., Dacun, Changhua, Taiwan 51591, R.O.C.

ABSTRACT

Extant image retrieval systems are mainly aimed at specific databases which extract and match appropriate features to retrieve similar images. However, the same searching approach cannot be successfully applied in other kinds of databases. To search for images from a binary, gray or color-image database by matching color features results in a discrepant retrieval. Several shape-based searching methods are usually used to retrieve images in which foreground objects are distinct from the background environment. Sometimes users want to search for images by inputting subjective descriptions such as the type of image, object or pictorial environment. Hence, this research implements different searching modules by considering the color-, shape-, and content-based characteristics of various images. During the construction of an image database, all items are aggregated into feature-clusters by a modified K-means algorithm. Different weights are then

*通訊作者之電子郵件 aven@mail.dyu.edu.tw。

assigned to each cluster by analyzing the characteristic variations in all images in a particular group. Thus, in this study, an image retrieval system obtained by integrating multiple searching approaches is ultimately achieved. After the experimental steps were completed, the results showed that the best retrieval effect is obtained by integrating the three aforementioned searching approaches.

Key Words: multiple searching approaches, image retrieval system, feature-clustering

一、緒論

影像檢索系統 (image retrieval system) 通常以查詢影像 (query image)，或是利用輔助工具繪出圖形，來當作輸入資料，以進行資料庫中相近影像的檢索。藉著抽取輸入影像中適當的特徵值，與資料庫中每張影像的相對特徵值進行比對，將較為相似的影像及其相關資訊回傳給檢索者。一個影像檢索系統含有三個基本步驟：查詢、比對和結果顯示。其中，在「查詢」步驟可輸入文字串、查詢影像、自繪圖形、或者選擇希望檢索的影像特性（如顏色、紋理、形狀與空間關係等）。「比對」則是利用一些不同的演算法，在資料庫中找尋符合查詢條件的相近影像。對於比對出性質相似的多張影像，則在「結果顯示」步驟中，依照相關性、相似度、影像格式、建置日期等排序分頁列出，以滿足檢索者所期盼之介面呈現。

有關影像檢索的技術，早期發展重點主要以各影像的說明文字檢索為主，近期則以從影像中抽取特徵值來進行比對為主流 [1, 2]。使用關鍵字 (keywords) 為查詢字串的影像檢索方法，主要是以人類所習慣的影像描述文字做為檢索的依據；而以影像特徵值進行比對的檢索系統，則主要抽取如顏色、紋理、形狀等的影像特性，用以進行特徵向量比對。以關鍵字串來檢索影像資料庫的主要缺點，為資料庫中每張影像都需要事先以人工標注說明文字，除了人力成本的龐大花費外，也太過依賴標注者本身的主觀認知。一旦日後檢索者和當初資料庫建構者對同一張影像的認知不同，將可能使得輸入之查詢關鍵字串，並不存在於影像說明內容中，以致無法正確檢索到該影像。

以影像查詢方式的檢索系統 [4, 13]，可分為三個發展階段：第一階段為比對原始像素 (raw pixels)，當使用者輸入查詢影像後，系統先將資料庫中每張影像與此查詢影像做大小正規化 (size normalization)，再以樣版比對方式 (template matching) 進行差異累計，找出相近之影像。第二階段是語意 (semantic) 查詢 [5]，資料庫中每張影像須由人工詮釋內容並以文字註解，由系統分析並統計文字內容

描述之影像特性。在檢索階段，系統會自動分析並判定查詢影像的內容特性，並在資料庫中找尋類似特性的影像。第三階段是比對特徵值 (feature values)，從影像中擷取適當的特徵，用以與查詢影像之特徵值進行比對，找出相似的影像。其中，顏色 (color) 是最常被用來作為檢索條件的影像特徵 [10]，包括影像中的顏色比例、分佈、以及主要顏色出現位置等。紋理特性 (texture) 也是常見的影像特徵，可利用影像紋理的粗糙度、對比及方向性，作為比對的依據。而對於有明顯前景物體之影像資料庫，物體之外形也是重要的影像特徵，包括相同顏色或紋理區域的形狀，或者前景物體的外觀等 [7]。

以影像特徵值的檢索方式並非只是字串檢索的替代技術，應可著重在兩者的互補性，若能整合兩種技術應可獲得令人滿意的檢索成效。因此，本研究利用影像之顏色和形狀特徵以及內容特性，分別建構三種影像搜尋模組，並利用不同權重將之整合成為具實用性的影像資料庫檢索系統。

影像資料庫檢索系統除了搜尋核心的技術外，提供給使用者的查詢介面也是重要的課題。常用的查詢方法，主要分為下列三種：第一種為文字查詢 (query by text) [11]，使用者輸入字串來查詢影像，輸入方便且可明確表達查詢者之意圖；但是針對相同影像，不同人的描述會有所不同，系統要找出符合每人的檢索期待並不容易。第二種是選項查詢 (query by selection) [17]，使用者利用系統事先建立的影像分類，逐一利用選取項目方式，建立適當的查詢條件。由於已事先針對各項目將所有影像分類，當查詢條件確定後，系統的檢索速度甚快；然而，使用者有時無法明確理解各項目之意義，很難將檢索期待以選項方式作為描述。第三種為範例影像查詢 (query by image) [18]，使用者輸入查詢影像後，系統抽取該影像之特徵值並與資料庫中所有影像進行比對；若使用者未自備查詢影像，也可點選系統提供多種類型之樣本影像。此種方式將使用者之檢索需求，以輸入或選擇範例影像方式表達，由於操作簡單、明確，較受使用者歡迎。

本研究所採用的查詢介面為範例影像查詢方式，使用者可以選擇系統提供之樣本影像，或是自行輸入查詢影像。對

於影像內容檢索部分，則利用選項設定方式，讓使用者用以選取影像明亮度、場景、種類等相關內容特性做為查詢條件。

圖 1 為本研究建構之影像檢索系統處理流程，分別以抽取查詢影像之顏色和形狀特徵，以及輸入內容特性選項方式，來建構不同的影像的搜尋模組。同時，分析所有影像在顏色及形狀特徵分群後，各群集影像之特徵差異度，於檢索時給予不同權重。使用者可利用輸入查詢影像並選定檢索條件，系統會以影像特徵比對、並以內容特性條件過濾方式，來進行影像檢索，將檢索結果依照相似程度分頁列出。

二、整合式影像檢索系統

欲建構一套適用不同影像資料庫之檢索系統，應該同時考慮並整合不同的影像搜尋方法。本研究分別以影像之顏色特徵、形狀特徵以及內容特性，建構三種影像搜尋模組，再依照不同特徵之影像分群特性，分別給予不同權重，整合成一套完整之影像檢索系統。

(一) 顏色特徵搜尋方法

由於資料庫中影像數量龐大，為降低系統負擔、提升檢索效率，我們將每張影像由大而小，進行三層的區域分割(如圖 2)，並抽取相對的顏色特徵。首先，針對整張影像，分別算出所有像素顏色 RGB 三原色各自的平均值，當作第一層的特徵向量 $Feat_{L1}$ 。再來，將影像等分切成 4 個區塊，分別算出各區塊中所有像素顏色的平均值，當作第二層的特徵向量 $Feat_{L2}$ 。最後，將影像等分切成 16 個區塊，分別算出各區塊中所有像素顏色的平均值，當作第三層的特徵向量 $Feat_{L3}$ 。由於將像素顏色的 RGB 三原色分別計算，此三層的顏色特徵向量分別為 3 維、12 維與 48 維。影像檢索時會依序利用此三層顏色特徵向量，將資料庫中每張影像與查詢影像之相對特徵向量進行比對。

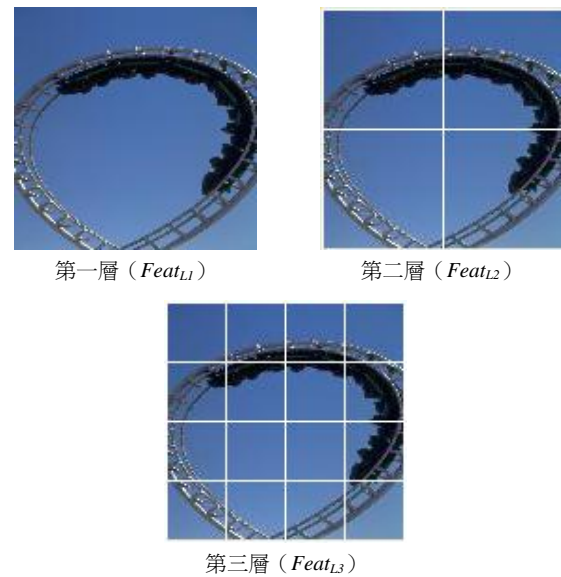


圖 2. 三層式影像分割之顏色特徵向量抽取

除了利用三層式影像分割方式，抽取各區域顏色特徵作為比對依據；為加速影像檢索，在資料庫建構階段，即利用每一層的顏色特徵，將大量的影像事先予以分群 (clustering)。實際進行影像檢索時，則先判斷查詢影像的顏色特性較接近哪個群集 (cluster)，再對該群集中的所有影像做細部比對，以檢索出資料庫中與查詢影像顏色特徵相近之影像。在判斷查詢影像落於何群集時，常因同群影像之顏色特性變化大，或者不同群集影像間有類似的顏色特性，均容易誤判查詢影像可能落居之群集。若根據不同影像資料庫的顏色特性差異度，自動決定適當的群集數目，盡可能地將性質接近的影像聚集成群，而避免特性差異明顯的影像同群，則可降低影像檢索時發生誤判群集的機率。因此，本研究利用改良式之 K-means 分群法，分別對三層之顏色特徵向量作資料庫中影像的分群，其分群結構如圖 3 所示。

1. 改良式之 K-means 分群法

「分群」的概念，是將資料集中的資料點分成幾個群集，使得同一群集中資料點間的相似程度高於相異群集間的資料點 [8, 9, 14]。最常見的分群方法為 K-means 演算法，這是一種前設式的分群方法，必須先行設定群集的數量，然後不斷以距離各群集中心的遠近，來進行所有資料點的重新分派，並持續重新調整各群集中心之位置，直至所有資料點所在群集不再變動。使用 K-means 方法來將 n 個影像特徵向量分成 k 群之演算法如下：

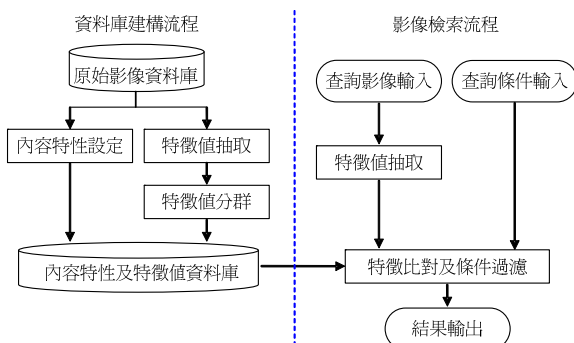


圖 1. 本研究開發之影像資料庫檢索系統流程

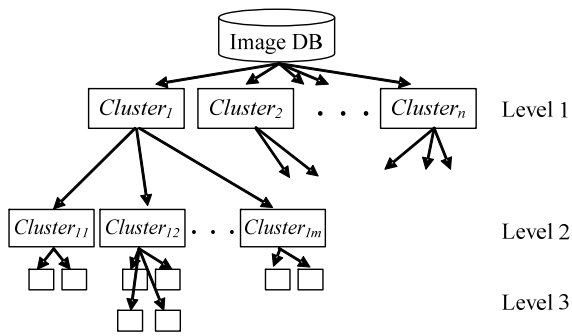


圖 3. 顏色特徵之分群結構

- Input** 影像特徵向量 F_1, F_2, \dots, F_n ，群集數 k
- Output** 分群之結果 C_1, C_2, \dots, C_k
- Step 1** 隨機找出 k 筆特徵向量值作為初始群集中心
- Step 2** 計算每一筆特徵向量 F_i 到 k 個群集中心之間的距離，並將 F_i 分派至距離最近之群集 C_j
- Step 3** 根據目前的分群結果，重新計算每一群集的中心位置
- Step 4** 不斷重複 Step 2 與 Step 3，直到每一筆特徵向量所在的群集不再變動

以圖 4 為例，假設已設定群集數量 $k=3$ ，系統先隨機找出 3 個資料點作為初始群集中心（圖 4(a)）；計算每一個資料點到 3 個初始群集中心之間的距離（圖 4(b)），並分派到距離最近的那一群。所有資料點分派後，根據目前的分群結果，重新計算每一群集的中心位置（圖 4(c)）後。以同樣的方式，持續不斷進行分群動作，直到每一個資料點所在的群集不再變動，即可得到最終的分群結果（圖 4(d)）。

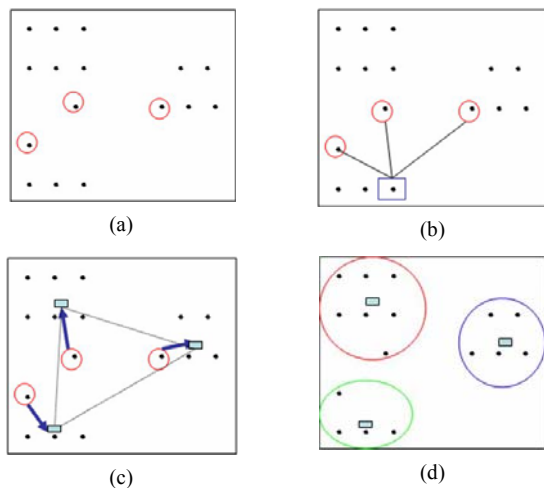


圖 4. 傳統 K-means 分群法

然而，傳統 K-means 分群法主要的缺點，在於它的分群效果會受到一開始的群集數量 k 值大小的影響。一旦資料點若變化極大，實際的分群數應該遠多於 k ，此方法還是只能分成 k 群；而資料點若相似度大，實際的分群數應該遠少於 k ，此方法則會硬性分成 k 個群集。因此，本研究採用不必事先決定群集數的改良式 K-means 方法，同樣將 n 個影像特徵向量進行分群之演算法如下：

Input 影像特徵向量 F_1, F_2, \dots, F_n

Output 分群之結果 C_1, C_2, \dots, C_m

- Step 1** 找出距離最遠的兩個特徵向量當作起初始群集中心 C_1 與 C_2
- Step 2** 計算每一筆特徵向量 F_i 到各群集中心之間的距離，若距離均大於門檻值，則將 F_i 當作新增的群集中心 C_j ；否則，將 F_i 分派至距離較近之群集
- Step 3** 根據目前的分群結果，重新計算每一群集的中心位置
- Step 4** 不斷重複 Step 2 與 Step 3，直到每一筆特徵向量所在的群集不再變動

以圖 5 為例，首先找出距離最遠的兩個資料點當作初始群集中心（圖 5(a)），計算其餘每一個資料點到兩初始群集中心的距離（圖 5(b)），並將該資料點分派至距離較近之群集。所有資料點經過初步分群之後，重新調整群集中心位置，並再度計算每個資料點至各群集中心的距離；若某資料點與所有群集中心的最近距離仍過遠（圖 5(c)），則將此資料點當作新增的群集中心（圖 5(d)）。持續不斷進行分群動作，直到每一資料點所在的群集不再變動（圖 5(e)）。

2. 應用群集分析之檢索方法

本研究開發之影像檢索系統，在資料庫建構階段先以每層抽取之顏色特徵，利用改良式 K-means 分群方法，將大量的影像各自分成數個群集。當檢索者輸入查詢影像後，系統會先抽取該張影像各層之顏色特徵，利用前兩層之特徵向量 $Feat_{L1}$ 與 $Feat_{L2}$ 來與相對層級之所有群集中心作差異度比較，來決定與該查詢影像最相近之群集。然後，再以第三層特徵向量 $Feat_{L3}$ 與此群集之每張影像進行細部比對，決定並排序出資料庫中的哪些影像為檢索結果。在差異度計算方面，假設查詢影像某層之顏色特徵向量為 (Q_R, Q_G, Q_B) ，而資料庫中某張影像之同層顏色特徵向量為 (D_R, D_G, D_B) 。先算出各顏色分項的差異值 $Dist_R$ 、 $Dist_G$ 與 $Dist_B$ 後，再利用不同的權重 w_R 、 w_G 與 w_B ，計算出兩張影像之差異度

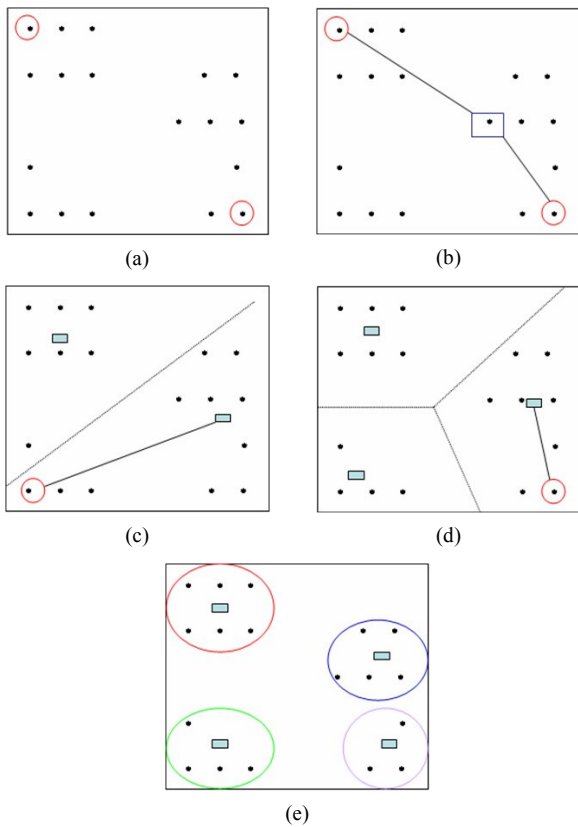


圖 5. 改良式 K-means 分群法

$Dist_{clr}$ ，如公式 (1) 所示。

$Dist_{clr} = Dist_R \times w_R + Dist_G \times w_G + Dist_B \times w_B$ ，其中

$$Dist_R = \|Q_R - D_R\|, Dist_G = \|Q_G - D_G\| \text{ and } Dist_B = \|Q_B - D_B\| \quad (1)$$

其中，RGB 三原色權重值 w_R 、 w_G 與 w_B 的決定，會在資料庫建構階段，分析各群集中心、以及每一群集內所有影像在 RGB 各原色的分佈情況，變異度越大者表示該原色對分辨不同影像有越大幫助，可給予較大之權重值。以圖 6(a) 的查詢影像為例，本系統利用顏色特徵檢索並依照差異度排序之結果，如圖 6(b) 所示。

(二) 形狀特徵搜尋方法

利用影像之顏色特徵，可以檢索出資料庫中與查詢影像顏色接近之影像，但卻可能找出拍攝標的物截然不同的影像。因此，本研究除利用顏色特徵外，同時也考慮各影像內物體的形狀特徵，來進行影像之檢索。使用者可以藉由對欲檢索影像的形狀認知，選擇或輸入相似的範例影像；系統會先抽取其形狀特徵後，再與資料庫中所有影像進行形狀特徵之比對，檢索出含有相似形狀特性的影像。欲抽取影像中物體之外形特徵，本研究先對影像進行邊緣點偵測，經過雜訊

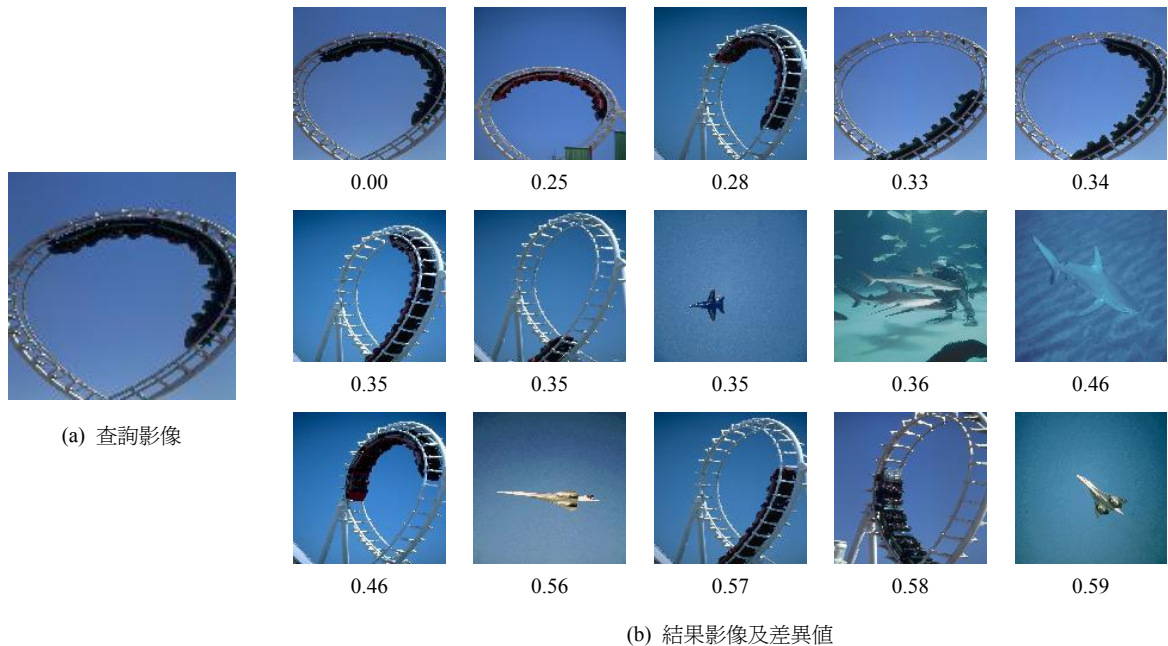


圖 6. 利用顏色特徵之檢索結果

去除後，將剩餘之邊緣點作水平及垂直方向投影，抽取兩方向之投影量特性，當作該影像之形狀特徵。

首先，我們分別利用水平以及垂直 Sobel 遮罩 [3]，算出水平梯度 (gradient) G_x 與垂直梯度 G_y ；若 G_x 或 G_y 的數值夠大，則判定該點為邊緣點。圖 7(a) 及 (b) 分別顯示一張原始影像及利用 Sobel 運算後所得之邊緣點影像。由於每張影像背景環境的複雜度不同，所產生的雜亂邊緣點可能會干擾形狀特徵的抽取；因此，本研究採取影像型態學 (morphology) 裡的斷開 (opening) 運算 [3]，以一個 3×3 的結構元素 (structuring element)，先對邊緣點影像做侵蝕 (erosion)，再做膨脹 (dilation)，即可清除部分雜訊並保留物體原貌 (如圖 7(c) 所示)。每張影像先經過邊緣點偵測及雜訊去除後，存留之邊緣點再分別往水平及垂直方向投影，並在兩方向的投影量圖 (histogram) 中，找到邊緣點的主要分佈範圍，並據以將投影量圖分成四等分 (HP_i 與 VP_i , $i=1, 2, 3, 4$)，如圖 8 所示。

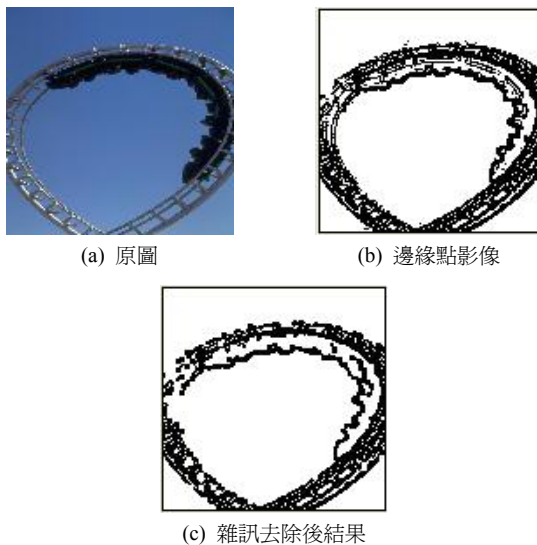


圖 7. 邊緣偵測效果

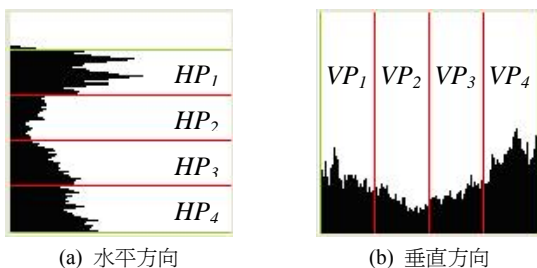


圖 8. 邊緣點投影量圖及四等分切割

由於相似形狀的物體，在不同影像中其大小會有所不同，所產生的投影量分佈也會不同；因此，本研究將兩方向每等分的投影量做正規化 (normalization)。利用公式 (2) 的計算，將原先投影量由 HP_i 轉為 HP'_i 、 VP_i 轉為 VP'_i 後，抽取出該影像之形狀特徵向量 $Feat_{shape} = (HP'_1, HP'_2, HP'_3, HP'_4, VP'_1, VP'_2, VP'_3, VP'_4)^T$ 。同樣地，結合改良式 K-means 分群法的應用，可將資料庫中每張影像的形狀特徵值，事先分群後存入特徵值資料庫。

$$HP'_i = \frac{HP_i}{\sum_{j=1}^4 HP_j} \quad \text{and} \quad VP'_i = \frac{VP_i}{\sum_{j=1}^4 VP_j}, \quad i \in \{1, 2, 3, 4\} \quad (2)$$

假設查詢影像之形狀特徵向量為 Q_{shape} ，資料庫中某張影像形狀特徵向量為 D_{shape} ，公式 (3) 為利用形狀特徵來計算兩影像差異程度 $Dist_{shape}$ 的計算方式。當 $Dist_{shape}$ 小於門檻值時，該張影像可判斷為與查詢影像相近，系統再將這些影像依差異程度高低排列，成為最終檢索結果。圖 9(b) 為與查詢影像 (圖 9(a)) 相似之數張依差異程度高低排列的檢索結果影像。

$$Dist_{shape} = \|Q_{shape} - D_{shape}\| \quad (3)$$

由圖 9 的檢索結果中，可以看到一些結果影像與查詢影像似乎差異甚大。為了過濾這些影像，我們針對形狀特徵之抽取與比對方式進行調整；在兩方向投影量分佈圖中，分別計算每等分中投影量分佈的變異數 (如圖 10 所示)，產生另一個外形變異度特徵 $Feat_{shapeVar} = (\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2, \sigma_5^2, \sigma_6^2, \sigma_7^2, \sigma_8^2)^T$ ，用以過濾原先利用形狀特徵 $Feat_{shape}$ 檢索出的結果影像。假設查詢影像之外形變異度特徵為 Q_{var} ，某檢索結果中某影像的特徵為 D_{var} ，一旦其差異值 $\|Q_{var} - D_{var}\|$ 過大，則可將該影像從檢索結果中移除 (如圖 11 所示)。

(三) 內容特性搜尋方法

利用影像之顏色或形狀特徵做檢索的依據，往往是搜尋技術上的考量；然而，並非所有使用者都可從技術角度，明確選出適當的範例影像，導致檢索結果不如預期。對於一般使用者來說，若能以選項方式主觀地設定檢索條件，應該更能夠明確表達心中期盼所得的影像特性。影像檢索系統配合



圖 9. 利用形狀特徵之檢索結果

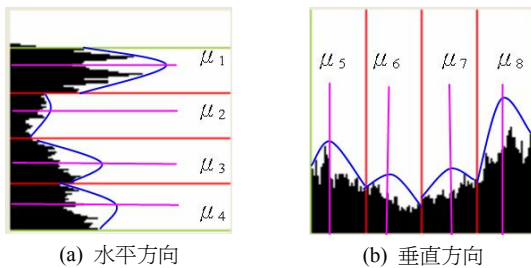


圖 10. 計算投影量中各分格之邊緣點分佈特性

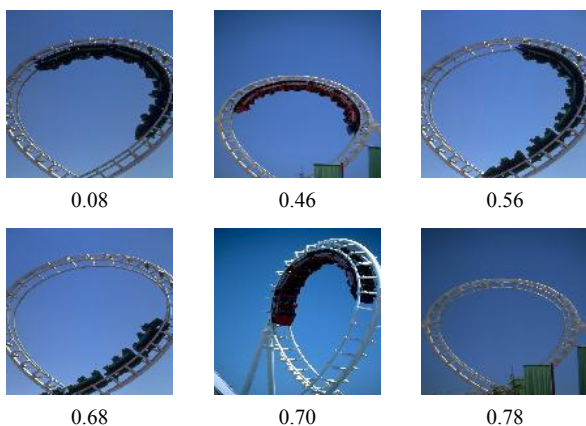


圖 11. 加入變異數過濾後，圖 9(a) 影像之最終檢索結果

使用影像內容特性的分類查詢方法，將使系統更加完整成熟 [6, 12, 15, 16]。本研究根據所收集的大量影像資料庫，採用

明顯的影像內容特性，提供相關選項供使用者選擇。第一種內容特性是「場景」，提供的選項有 [室內] 與 [室外] 兩個類別；第二種內容特性是「明暗度」，提供的選項有 [昏暗]、[適中] 與 [明亮] 三個類別；第三種內容特性是「種類」，提供的選項有 [人類]、[動物]、[植物]、[交通工具]、[建築物]、[鳥類]、[海洋生物]、[風景] 與 [其他] 等九個類別。

1. 內容特性建構方式

針對影像內容提供適當的特性選項讓使用者選擇，對影像檢索者而言相當方便；但是，針對資料庫中的所有影像，必須逐一以人工方式設定其相關特性。為避免不同人員對相同影像的主觀認知有所差異，本系統對內容特性的設定並非二分法，而是提供一個簡單介面（如圖 12 所示），讓資料庫建構者可針對每張影像三種內容特性（「場景」、「明暗度」與「種類」）的所有類別，分別設定其程度等級（有 [極相似]、[相似]、[普通]、[相異] 與 [極相異] 五種）。如此一來，對於一些特性模糊的影像，可以設定某特性的不同類別有相同的等級。例如，某影像是主人與寵物合照，則該影像之「種類」特性的 [人類] 與 [動物] 類別，其等級均可設為 [極相似]。

2. 內容特性比對方式

內容特性的比對，其用途主要是從資料庫中過濾影像，只擷取出符合使用者設定之內容特性選項者。然而，系統建



圖 12. 影像內容特性建構介面

置者與使用者對相同影像之認知可能有所差異，當影像檢索者選定相關內容特性之程度等級後，本系統會自動從資料庫中挑選特性等級相近之影像，而非等級須完全相同者。例如，當使用者設定欲檢索影像之某特性等級為 [普通]，則本系統會自動挑選該特性之等級為 [普通]、[相似] 或 [相異] 之所有影像；而使用者若設定 [極相似] 等級，系統則會挑選 [極相似] 與 [相似] 兩個等級者。因此，利用本研究提出之影像內容特性檢索機制，可容忍個人認知上的差異，以增加檢索的效果。

(四) 權重設定

不同影像資料庫其各種特徵值的顯著性並不相同，對於色彩繽紛且對比明顯者，適合使用顏色特徵；而對於物體外形單純且明顯者，則適合使用形狀特徵。因此，如何彈性地決定權重，將利用不同影像特徵的搜尋方法做結合，對於建立一個適用於各種影像資料庫的檢索系統，是個很重要的課題。因此，本研究利用不同特徵進行資料庫中的影像分群

後，分析各群集內特徵值的差異程度，來決定各特徵的重要程度，而給予每一群集不同的權重。

假設利用顏色特徵做影像分群的結果，在第 k 群中共有 n 張影像 $Cluster_k = \{img_1, img_2, \dots, img_n\}$ ，每張影像 img_i 有各自的顏色特徵向量 (R_i, G_i, B_i) ，計算出 n 張影像顏色特徵向量的變異數 $(\sigma_R^2, \sigma_G^2, \sigma_B^2)$ 後，以公式 (4) 的計算方式，設定此群集的權重 $wt_{clr}(k)$ 。

$$wt_{clr}(k) = \begin{cases} 0.8, & \text{if } \bar{\sigma}_k^2 < MinThr_{clr_}\sigma \\ 1, & \text{if } \bar{\sigma}_k^2 \in [MinThr_{clr_}\sigma, MaxThr_{clr_}\sigma] \\ 1.2, & \text{otherwise} \end{cases}, \text{ 其中}$$

$$\bar{\sigma}_k^2 = \frac{\sigma_R^2 + \sigma_G^2 + \sigma_B^2}{3} \quad (4)$$

其中，兩門檻值 $MinThr_{clr_}\sigma$ 與 $MaxThr_{clr_}\sigma$ 的決定方式，先將各自影像群集內所有影像之顏色特徵變異數 $\bar{\sigma}_1^2, \bar{\sigma}_2^2, \dots, \bar{\sigma}_N^2$ 算出，並根據大小排序後，以中位數 (median) Val_{med} 當作基準，將 $MinThr_{clr_}\sigma$ 設為 $0.5 * Val_{med}$ ；而 $MaxThr_{clr_}\sigma$ 則為 $1.5 * Val_{med}$ 。當查詢影像輸入後，分析該張影像之顏色特徵值落於哪一群集後，在計算與該群集內各影像之顏色差異程度時，將差異值分別乘上該群集的權重。如此一來，顏色特徵差異較明顯的影像群集，其顏色差異值就會佔較高權重。圖 13(b) 為尚未對各群集設定權重時，針對查詢影像 (圖 13(a)) 檢索後依差異程度高低排列之結果；加入各群集之權重考量後，輸入與圖 13(a) 相同的查詢影

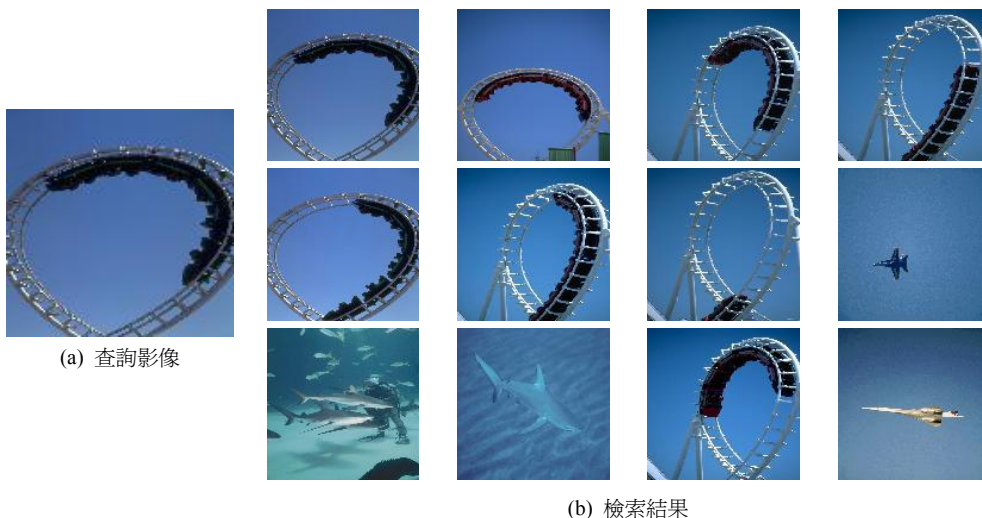


圖 13. 加入群集權重前之影像檢索結果 (利用顏色特徵)

像，其檢索結果如圖 14 所示。

除了利用群集內影像顏色特徵的差異程度決定權重外，對於利用形狀特徵分群的各群集，同樣可利用群集內所含影像形狀特徵的變異程度，來決定權重（如公式（5）所示），亦可提升利用形狀特徵進行的影像檢索效果。

$$wt_{shape}(k) = \begin{cases} 0.8, & \text{if } \bar{\sigma}_k^2 < MinThr_{shape_σ} \\ 1, & \text{if } \bar{\sigma}_k^2 \in [MinThr_{shape_σ}, MaxThr_{shape_σ}] \\ 1.2, & \text{otherwise} \end{cases} \quad (5)$$

其中，兩門檻值 $MinThr_{shape_σ}$ 與 $MaxThr_{shape_σ}$ 的決定方式，類似計算顏色特徵群集權重時的 $MinThr_{clr_σ}$ 與 $MaxThr_{clr_σ}$ 。

三、檢索介面與操作方法

本研究擷取影像之顏色與形狀特徵，並分析資料庫之影像分群效果，訂定每群集之權重，建構影像檢索系統；使用者也可設定欲檢索影像之內容特性，系統可據以過濾影像檢索結果。除了整合多搜尋方法來建構影像檢索系統外，本研究同時也考量使用之方便性，來設計檢索介面以及檢索步驟，使得系統可根據使用者輸入的查詢影像及相關設定，來做更符合期待的影像檢索及結果呈現。

（一）檢索介面

本研究開發之影像檢索系統（介面如圖 15 所示），使用者可利用按選『瀏覽範例影像』，輸入或點選查詢影像，並以勾選『顏色』、『外形』或『內容特性』選項，決定是

否採用影像顏色特徵、形狀特徵或者內容特性，來進行影像之檢索。針對「內容特性」部分，使用者可進一步勾選要考量的特性有哪些，不管是 [場景]、[明暗度] 或 [種類]，均可選取適當之類別，並選擇相似程度。而「顏色」與「外形」部分，則可設定重要等級（高、中、低），以利後續將兩搜尋方法所得的不同檢索結果，加上權重做交叉排序。查詢設定完成後，按選『檢索』鈕，即可開始進行影像檢索，並將檢索結果影像按照與查詢影像之差異程度，以分頁方式顯示。

（二）差異度合併計算

在檢索介面的設定中，使用者可能同時選取利用查詢影像的顏色特徵與形狀特徵，來進行資料庫中相關影像的檢索。由於利用兩特徵檢索出的結果並不完全相同，根據使用者針對兩特徵所設定的重要等級，本系統分別賦予不同的主觀權重（低、中、高分別設定為 0.2、0.5 與 0.8）。例如，使用者輸入一張查詢影像後，勾選「顏色」及「外形」特徵並分別設定重要性為 [高] 與 [低]。首先，系統將分別利用影像的顏色特徵與形狀特徵進行影像檢索，並各自產生檢索結果。若某影像 img_k 同時出現在兩檢索結果中，其與查詢影像之差異程度分別為 $D_{clr}(k)$ 與 $D_{shape}(k)$ ；再來，系統對兩差異值 $D_{clr}(k)$ 與 $D_{shape}(k)$ 分別做正規化（normalization），使其值均介於 0~1 之間（以 $D'_{clr}(k)$ 與 $D'_{shape}(k)$ 表示）。最後，再利用兩特徵不同的重要性設定，分別對 $D'_{clr}(k)$ 與 $D'_{shape}(k)$ 進行加權，得到該影像 img_k 最後之檢索差異值 $Dist(k)$ ，其計算如公式（6）所示。

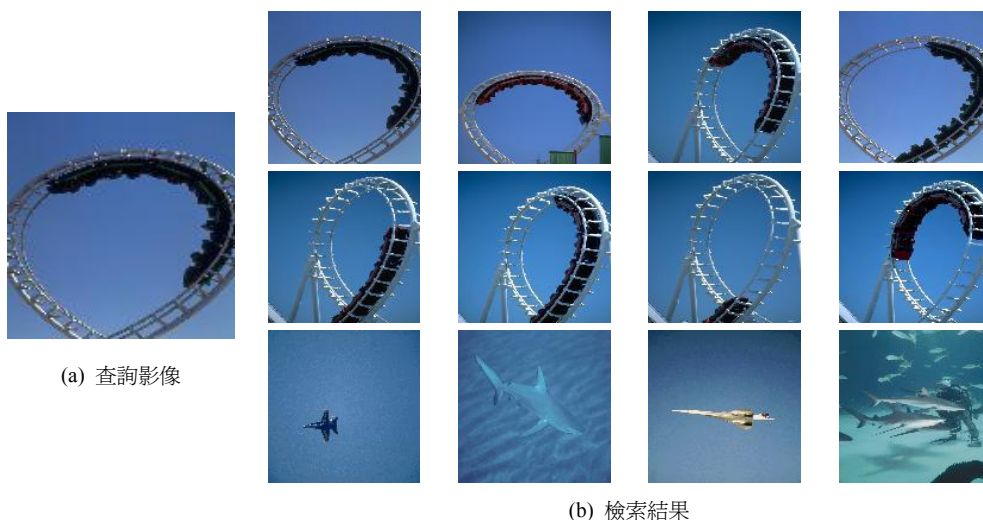


圖 14. 加入群集權重後之影像檢索結果（利用顏色特徵）



圖 15. 查詢介面設計圖

$$Dist(k) = D'_{ctr}(k) \times 0.8 + D'_{shape}(k) \times 0.2 \quad (6)$$

以此方式分別將利用顏色特徵與形狀特徵所得之檢索結果作交叉排名，即可得到最終之影像檢索結果。

四、實驗結果與分析

本研究分別利用影像之顏色特徵與形狀特徵開發並整合成資料庫檢索系統，也輔以使用者對影像內容特性的檢索條件設定，提高檢索效果與使用方便性。此影像檢索系統目前為 Windows 應用程式，使用 Microsoft Visual Studio 2005 開發環境，以 C# 程式語言撰寫。欲測試所提方法之影像檢索效果，本研究將此系統安裝在含有 Pentium4-2.6G CPU 與 512MB RAM 配備的桌上型個人電腦上，並以來自美國加州大學河濱分校 VISLAB 實驗室 (<http://www.vislab.ucr.edu>) 的多張影像為測試資料庫。此影像資料庫中包含人類、動物、植物、建築物、交通工具、風景等不同主題之影像約 1,500 張，均為 128x128 pixels 大小的 JPEG 格式檔案。本研究之實驗方式為依照各主題不同之影像數量，按比例抽取共 100 張影像作為測試用查詢影像；每張查詢影像均利用不同的設定做檢索，檢索結果再由 5 位成人測試者，檢驗其檢索效果後予以平均。本研究評估影像檢索效果的量測方式，採用最為普遍的回收率 (recall rate) 與精確率 (precision rate)，

其計算方法分別如公式 (7) 與公式 (8) 所示。

$$\text{回收率} = \frac{\text{檢索到相關影像數量}}{\text{資料庫中所有相關影像數量}} \quad (7)$$

$$\text{精確率} = \frac{\text{檢索到相關影像數量}}{\text{檢索所得影像數量}} \quad (8)$$

回收率愈高，代表資料庫中被檢索到相關的影像數量愈多；而精確率愈高，代表檢索所得之所有影像符合使用者期待者愈多。

為了驗證本系統單獨應用顏色特徵或形狀特徵的影像檢索效果，我們在檢索介面中僅勾選「顏色」或「外形」項目其一後，針對測試用 100 張查詢影像進行檢索，並由 5 位測試者主觀地統計各查詢影像之 [資料庫相關影像]、[檢索出相關影像] 及 [檢索所得影像] 等數量後，分別計算出兩種特徵各自的平均檢索效果 (回收率與精確率)。另外，由於本研究也提供使用者可設定欲檢索之影像內容特性，並與顏色與形狀特徵之檢索方式結合應用；因此，各種不同檢索方式組合經過實驗測試，所量測之檢索效果如表 1 所示。

使用者對影像資料庫檢索結果之接受度，往往顏色相近者重於形狀接近者。雖然檢索出與查詢影像不甚相符者，若色彩分佈極為接近，使用者往往能接受；若形狀接近但顏色差異較大時，使用者會覺得較無法接受。由表 1 所列之影像

表 1. 不同檢索方式組合之檢索效果

檢索方式	檢索效果	回收率 Recall rate	精確率 Precision rate
	顏色特徵	83.0%	67.6%
	形狀特徵	75.5%	61.6%
	顏色特徵+形狀特徵	80.8%	64.4%
	顏色特徵+內容特性	87.1%	71.0%
	形狀特徵+內容特性	85.7%	68.6%
	顏色特徵+形狀特徵+內容特性	92.0%	74.1%

檢索效果，可發現單獨使用顏色特徵搜尋方法較優於僅用形狀特徵。然而，由於單用形狀特徵的效果並不算理想，所以導致「顏色特徵+形狀特徵」的檢索效果反而低於單用顏色特徵。其實，形狀特徵並非完全無用，只是有時不同類型的影像恰好有類似的形狀特性（邊緣像素分佈），而容易發生檢索錯誤。根據本研究的觀察，若能事先知道欲檢索的影像分類，形狀特徵的檢索效果其實還不錯。因此，從表 1 的檢索效果列表中，可看出使用者若勾選了「內容特性」選項再結合形狀特徵，可得到較單獨使用顏色特徵更佳的影像檢索效果。

因為影像的內容特性描述為主觀認知，只要建置者與檢索者的認知差異不至過大，本系統會將不符檢索特性的影像先過濾掉，使得檢索出的影像均符合設定之內容特性。而當三種搜尋方法整合在一起時，可將系統檢索效果發揮到最高，回收率可達 92%；雖然準確率僅有 74.1%，但對使用者而言，檢索效果的認知以回收率重於準確率。本系統利用顏色或形狀特徵檢索出之相關影像，會利用差異程度做排序，即使會檢索到一些與查詢影像相似度不夠高的影像，但因相似度較低而被排在較後序位，使用者感受到之檢索準確率將高於系統測試所得之數值。

除了以量化方式測試所提方法之影像檢索效果，根據本研究開發之檢索系統所提供的功能，我們分別與開放使用之幾個國內外影像檢索網站（包括台灣蝴蝶外觀檢索系統 http://dlm.ncnu.edu.tw/butterfly/query_html/query_index.htm、美國哥倫比亞大學的 WebSEEK 網站 <http://persia.ce.columbia.edu:8008> 以及美國 IBM 公司的 QBIC 網站 <http://www.qbic.almaden.ibm.com>）進行功能上的比較，歸納分析結果如表 2 所示。

「台灣蝴蝶外觀檢索系統」的用途僅限於蝴蝶影像資料庫的檢索，提供了以關鍵字來查詢影像的文字檢索介面，以及讓使用者點選外型特徵的介面，來進行蝴蝶影像的檢索。

表 2. 檢索功能比較

檢索系統 查詢方式	蝴蝶外觀檢索	WebSEEK	QBIC	本系統
關鍵字	√	√	√	
顏色查詢	√	√	√	√
形狀查詢	√		√	√
紋理查詢	√		√	
範例查詢	√	√		√
內容特性				√
權重設定		√		√

而美國哥倫比亞大學網站的「WebSeek」功能，則可依目錄和關鍵字來搜尋；該影像資料庫收藏了近七十萬張影像，找到一張影像後，可再依內容進一步搜尋。雖可達到階層式之檢索效果，但往往必須對影像資料庫的存放目錄與相關領域關鍵字有所瞭解，方能得到滿意之檢索效果。

IBM 公司的「QBIC」網站，則是第一個以內容為主（content-based）的影像資料庫，主要是以色彩長條圖（color histogram）、形狀與紋理（texture）特性為檢索的依據。使用者並不容易根據自己的主觀認知，設定符合個人需要的檢索方式。而本研究開發之影像檢索系統，則讓使用者可根據需要，自行決定單獨或合併使用「顏色特徵」、「形狀特徵」以及指定「內容特性」，並設定相關的權重或等級，來進行影像檢索。

五、結論

本研究提出以顏色特徵與形狀特徵來開發不同的影像搜尋方法，並輔以使用者輸入影像內容特性，建構完整之影像檢索系統。首先，針對影像資料庫的每張影像，分別抽取顏色與形狀特徵值後，利用 K-means 分群法將所有影像分成數個群集；然後，分析各群集影像之特徵值變異度，於後續檢索時給予不同權重。在影像檢索過程中，使用者先輸入查詢影像，並設定欲檢索的影像內容特性。系統會抽取該查詢影像的相關特徵值，並與資料庫之每個群集中心比對；決定落居之群集後，再對該群集的每張影像進行細部比對，最後，依差異程度高低排序輸出檢索結果。

利用完整之實驗步驟來測試本研究開發之影像檢索系統，可發現單獨使用顏色特徵之檢索效果較優於僅用形狀特徵；而利用使用者設定之內容特性條件，系統會先過濾掉不相關的影像，可提昇檢索效果。考慮依差異程度排序後之檢

索結果，不管是回收率或準確率數據，均顯示當三種搜尋方法整合在一起時，可將系統檢索效果發揮到最高。

誌謝

本文承蒙大葉大學校內服務型專題計畫補助（編號：ORD-9405），特此致謝。

參考文獻

1. Chen, Y. and Z. Wang (2002) A region-based fuzzy feature matching approach to content-based image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(9), 1252-1267.
2. Fuh, C. S., S. W. Cho and K. Essig (2000) Hierarchical color image region segmentation for content-based image retrieval system. *IEEE Transactions on Image Processing*, 9(1), 156-162.
3. Gonzalez, R. C. and R. E. Woods (2002) *Digital Image Processing*, 2nd Ed., Prentice-Hall, New Jersey, NJ.
4. Goodrum, A. (2000) Image information retrieval: An overview of current research. *Informing Science*, 3(2), 63-67.
5. Kuroda, K. and M. Hagiwara (2002) An image retrieval system by impression words and specific object names-IRIS. *Neurocomputing*, 43(1-4), 259-276.
6. Luo, J. and A. Savakis (2001) Indoor vs. outdoor classification of consumer photographs using low-level and semantic features. International Conference on Image Processing, Thessaloniki, Greece.
7. Mahmoudi, F., J. Shanbehzadeh, A. M. E. Moghadam and H. S. Zadeh (2003) Image retrieval based on shape similarity by edge orientation autocorrelogram. *Pattern Recognition*, 36(8), 1725-1736.
8. Ng, K. (2000) K-means-type algorithms on distributed memory computer. *International Journal of High Speed Computing*, 11(2), 75-91.
9. Pelleg, D. and A. Moore (1999) Accelerating exact k-means algorithms with geometric reasoning. International Conference on Knowledge Discovery and Data Mining, San Diego, CA.
10. Qiu, G. and K. M. Lam (2003) Frequency layered color indexing for content-based image retrieval. *IEEE Transactions on Image Processing*, 12(1), 102-113.
11. Rui, Y., S. Huang and S. F. Chang (1999) Image retrieval: Current techniques, promising directions, and open issues. *Journal of Visual Communication and Image Representation*, 10(1), 39-62.
12. Serrano, N., A. Savakis and J. Luo (2002) A computationally efficient approach to indoor/outdoor scene classification. International Conference on Pattern Recognition, Quebec City, Canada.
13. Smeulders, W. M., M. Worring, S. Santini, A. Gupta and R. Jain (2000) Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12), 1349-1380.
14. Su, M. C. and C. H. Chou (2001) A modified version of the k-means algorithm with a distance based on cluster symmetry. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6), 674-680.
15. Vailaya, A., A. Jain and H. J. Zhang (1998) On image classification: City images vs. landscapes. *Pattern Recognition*, 31(12), 1921-1935.
16. Vogel, J. and B. Schiele (2006) Performance evaluation and optimization for content-based image retrieval. *Pattern Recognition*, 39(5), 897-909.
17. Wu, J. K., A. D. Narasimhalu, B. M. Mehtre, C. P. Lam and Y. J. Gao (1995) CORE: A content-based retrieval engine for multimedia information systems. *Multimedia Systems*, 3(1), 25-41.
18. Zhong, Y. and K. Jain (2000) Object localization using color, texture and shape. *Pattern Recognition*, 33(4), 671-684.

收件：97.07.25 修正：97.10.22 接受：98.02.19