

結合顏色形狀與空間關係之影像擷取方法 Content-based Image Retrieval Using Color, Shape and Spatial Relationships

李進德
Jin-Der Lee,
義守大學資訊工程研究所

錢炳全
Bee -Chian Chien,
義守大學資訊工程系

章定遠
Din-Yuen Chan
義守大學資訊工程系

{m873002m, cbc, dychan}@csa500.isu.edu.tw

摘要

本篇論文提出一結合顏色、物件型態和空間關係等特性的影像內容查詢與相似影像擷取方法。以圖像內容作為查詢條件的影像擷取方法是近年來影像資料庫的主要研究方向，該方式提供了最方便、豐富的影像查詢能力。然而圖像本身的特性相當多，而且不易掌握，如何結合不同的影像特性，以期能滿足使用者的各種不同需求是一相當重要的研究主題。在本文中我們結合多種影像特性，包括顏色、形狀與圖像物件之空間關係等三種影像內容做為影像資料庫的索引與查詢之依據。我們藉由影像處理的萃取程序擷取不同的影像特性，並以適當的表示法有效率地加以記錄於資料庫中，然後設計一相似影像擷取演算法，有效的由影像資料庫中擷取出符合使用者所期望的圖像。我們的特性萃取程序不但改進了萃取影像特性的正確性，所使用的表示法也解決了一致性的問題，使特性資料可以輕易的利用多維資料結構來加以記錄並製作索引，經由實際的實驗結果，我們的方法可以更精確的擷取到相關的圖像。

關鍵詞：影像擷取、空間關係、相似評估、影像分割、資訊擷取

Abstract

Content-based retrieval is a common and convenient way for querying multimedia data recently. It has become one of the most important research directions on multimedia information retrieval. Many related results have been proposed for different features of images. In this paper, we propose a new method for content-based image retrieval and design a prototype system to evaluate the efficiency and effectiveness of our method. Three types of features in images including color, shape and spatial relationship are used to measure the average of similarity between two images. We design a new method to extract different features of images automatically and store these features in a unified representation. An object-pair matching algorithm and a similarity matching algorithm are also presented. By experiments, we find that the spatial relationship is an important for image retrieval from human's perception. The experimental results show that our new retrieval method is more efficient and effective than the previous methods.

Keywords: content-based, image database, spatial relationship, image segmentation, similarity measure, information retrieval

1 簡介

隨著製作多媒體資料工具與設備的演進，影像資料的數量日趨龐大，種類也越來越複雜，影像資料庫的需求也日趨殷切。影像資料庫所具備查詢之功能中，如何能夠讓使用者簡單而快速的擷取到所期望之圖像是相當重要的課題。在目前的研究中，以影像內容為主(Content-based)的影像擷取方法是目前的主要研究方向。目前有許多不同研究領域的專家學者，均對此一議題從不同的角度，有許多深入的研究與探討，並且得到不少成果。而關於影像內容的研究，可以將其分為兩類：一是以影像本身像素(pixel)分佈的客觀特性，如：顏色(color)的分佈、影像所表現出來的紋理(texture)、影像中形成單純物件的形狀(shape)等來作為內容查詢的依據；另一類則以影像中有意義的物件之分類(classification)、結構(structure)或空間關係(spatial relationship)等屬性來記錄其相關的影像內容。而目前的擷取技術偏向結合不同影像特性做為處理與比對的重點，例如，結合顏色與形狀[9][10]、結合顏色與空間關係[2][11]、結合顏色、形狀與紋理[8]等方法，這些方法嘗試結合各種不同的影像特性，希望由不同特性的組合能夠提供更符合人類視覺與適合使用者的影像擷取系統。通常這些系統均包含了幾個不同的影像處理程序，包括影像分割(image segmentation)、顏色分類、物件萃取(object extraction)、形狀判別等方法，這些不同的方法各有其優缺點以及其適用的時機與環境。本文提出一新的方法，結合影像顏色、物件形狀與圖像中物件之空間關係等不同種類的影像內容做為影像資料庫的索引與查詢之依據。我們藉由以前的相關研究，分析其優缺點，探討其適用時機，根據不同方法所表現出來不同的影像特性處理程序設法加以結合，使能萃取出不同影像物件及個別的形狀特性，最後再配合影像物件配對演算法及相對的空間關係來判斷圖像所代表意義內容的相似性。新的方法中，除了改進了原來各種方法的缺點與問題外，我們以更精確及更有效率的空間關係比對方法，設計一相似影像擷取演算法，使圖像中相關物件的排列更適合人類視覺的感覺，以提高擷取圖像的正確率。我們的影像特性萃取程序不但改進了萃取影像特性的正確性，這些所萃取出來的影像特性資料並可用一致性的表示法(unified representation)有效率地加以記錄於資料庫中，使所有的影像特性資料可以輕易的利用多維資料結構(multi-dimensional data structure)如：R-tree等，來加以記錄並製作索引(index)，增進蒐尋的效率，因此可以應用於大量影像資料庫的搜尋。所製作的離型系統經由實際的測試實驗結果，證實我們的方法可以更精確的擷取擷取出符合使用者所期望的圖像。本文結構組織如下：第2節中將對過去的研究做一個簡

單的回顧與整理。第3節中將描述我們所採用的影像特性、萃取方法與程序。第4節則介紹如何自動建立影像資料庫以及我們所提出新的相似影像比對及擷取的演算法。第5節為實驗的相關結果與討論，最後一節中則做一結論並說明未來的研究方向。

2 相關研究

QBIC(Query By Image Content) 是先期影像資料擷取系統中較具代表性的系統，由 IBM Almaden Research Center[8]所發展，採用影像之顏色、紋理、形狀以及概圖(sketch)等內容來代表一張圖像。使用者可以設定期望圖像之影像像素(image pixel) 特性來查詢圖像，並透過系統擷取出與所描述的像素特性相近的圖像。在QBIC系統中，提供三種查詢方式：第一，使用者可以指定查詢圖像中物件之主要顏色，然後系統以每個物件與資料庫圖像物件的顏色空 (color space) 差距，(如 RGB, YIQ, LAB 及 MTM)來計算兩圖之相似度；第二個方法採用查詢影像中像素顏色出現之機率來表示影像之特性並藉以做影像相似度的查詢；第三種查詢方法則是由使用者畫出期望圖像中所需顏色物件之方塊，並設定其代表之形狀，由所獲得查詢圖像物件之顏色、形狀、位置、大小等特性，進而計算兩張圖像之差距值。QBIC系統所用的方法主要均著重於影像中低階像素層次的特性，至於較高層次的特性，如：影像中之物件或甚至於所代表的意義(semantic)內容等，則無法以其判別。

關於如何結合低階影像像素特性與高階影像物件間的特性一直是相當困難的課題，在此一問題上傳統的圖形識別(pattern recognition)領域中有相當多的探討，但要找到一具普遍性及快速的物件辨識方法實非易事。Chang[2]提出以影像之顏色做為區分物件的方法，並結合顏色區塊物件之空間關係做為查詢相似圖像的依據。Chang所提出之方法則只採用顏色表示圖像的特性資料，作法為利用 Lin[7]提出之影像顏色物件叢集法(clustering)，將影像分割成為數個最佳顏色叢集區塊成為顏色物件，並以最小包圍區塊(minimum boundary rectangle; MBR)的資料與顏色建立資料庫。在查詢方面，先利用顏色區塊之位置導出查詢圖像與資料庫圖像之 2D String[1]表示式，再利用兩圖像的 2D String 計算出兩圖之空間關係差異度，進而擷取出相似之圖像。Chang的方法優點為所使用的顏色叢集法相當容易，且所判別出來的顏色物件與人類的視覺感知相當一致，然而其主要缺點為無法考量顏色物件的形狀，造成顏色優勢的現象，無法完全反應人類對物件特性的認知，並且所使用的 2D String 空間關係比對需將萃取之特性資料作一定程序的轉換，無法直接提供做為索引之用，而且不夠精確，對需處理大量的圖像資料庫而言，並非十分合適。

Mehre[10]則為一以顏色與形狀的組合來表示與擷取相似圖像的影像資料擷取系統。在 Mehre 的方法中，運用顏色與形狀為查詢的條件。其中顏色如同 QBIC 的第二個查詢方式，以顏色值與其色階表示，而形狀則根據每個顏色在圖片上的分佈，分割出每個顏色分佈所代表的不同區塊後，再採用 Hu[6]提出之不變動量(Invariant Moment)，計算出每個區塊之七個動量不變值來表示其形狀。比對的方式為以影像中每個物件的顏色與形狀差

值和，做為查詢圖像與資料庫圖像之差距值。此系統的優點為以簡易的動量不變值來代表相關圖形形狀，並且可以不受物件本身的旋轉、放大與縮小等效應影響；不過，此方法的顏色分割並無法確實分割出符合使用者需求的顏色區塊，造成許多比對上的困擾，而且每塊物件的比對均以同顏色的物件為主，並未考慮物件之空間關係，所以查詢結果常是局部的顏色物件相似，與整體影像的期望擷取圖像有所誤差，此為其主要的缺點。為了改進先前幾種影像擷取方法的缺失，我們提出結合顏色、形狀與空間關係的影像擷取方法，藉著整合不同方法的優點來克服其相對的缺點，並改進空間關係的比對，將可使系統能擷取到更符合使用者要求的圖片。下節將詳述我們所採用的各種影像特性萃取方法，第4節則探討如何結合這些方法。

3 影像特性之萃取方法

為了改進前人方法的缺點，並能夠將不同影像特性加以整合，以達到我們擷取符合使用者需要的圖像的目的，對於三種不同的影像特性，我們所使用的顏色空間採適合人類視覺的 CIE LU 的顏色表示法[5]，而形狀則是延用由 Hu 所提出之不變動量(Invariant Moment)表示法[6]，另外為達到人類對物件間空間關係的敏感度，我們則以 Chien[3][4]所提出之 2D Vector 來做空間關係的相似處理。以下詳述三種影像特性與評估差異的方法。

3.1 CIE LU 顏色空間

RGB 表示法為電腦系統中最常用來表示像素顏色結構的顏色空間，但 RGB 顏色空間並不符合人類眼睛對顏色差異的敏感度，因此不適合用來當做人類視覺查詢的顏色空間。在我們的影像資料擷取的方法中，根據許多相關研究與實驗結果[5]顯示，利用 LUV 顏色空間所評估出來的兩顏色差異，與人類視覺感應到的差異較為接近，因此決定採用 CIE LUV 顏色空間來作為評估物件間顏色的差異值，以期可以得到更符合人類視覺感受的影像顏色特性。RGB 與 LUV 顏色空間的轉換公式如下[5]：

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.490 & 0.310 & 0.200 \\ 0.177 & 0.813 & 0.011 \\ 0.000 & 0.100 & 0.990 \end{bmatrix} \times \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

$$u' = \frac{4X}{X + 15Y + 3Z}, \quad v' = \frac{9Y}{X + 15Y + 3Z},$$

$$u_0' = \frac{4X_0}{X_0 + 15Y_0 + 3Z_0}, \quad v_0' = \frac{9Y_0}{X_0 + 15Y_0 + 3Z_0}$$

$$L = 116 \times (Y / Y_0)^{1/3}, \quad Y / Y_0 > 0.01$$

$$u = 13 \times L (u' - u_0'),$$

$$v = 13 \times L (v' - v_0').$$

其中 R, G, B, 代表 RGB 空間的顏色值，而表示式中的 (X₀, Y₀, Z₀)則代表做為參考色的顏色，在此我們以白色為參考顏色，也就是 (R, G, B) = (1.0, 1.0, 1.0) 時所計算出來的 (X, Y, Z) 值。

顏色差異的評估方面，我們則利用兩顏色所計算出來的 L, u, v 值的方均差來估計。例如：假設有二不同顏色物

件 O_1 與 O_2 ，兩物件的 CIE LUV 顏色空間表示法為 (L_1, u_1, v_1) 與 (L_2, u_2, v_2) ，則兩物件的顏色差異值定義如下：

$$\Delta_c(O_1, O_2) = \sqrt{(L_1 - L_2)^2 + (u_1 - u_2)^2 + (v_1 - v_2)^2} \quad (1)$$

3.2 動量不變值(Moment Invariants)

Hu[6]所提出之動量不變值，是影像的形狀的重要特性之一。圖像經過旋轉縮放或平移之後，其新圖與原圖的動量不變值將不受影響，若將像素點誤差考慮進去的話，所表現出來的差距非常微小，所以可用來做為圖像辨識的重要特徵。假設圖像大小為 $M \times N$ ，像素位置 (x, y) ，其顏色值為 $f(x, y)$ ，則在 x 軸方向第 p 階與 y 軸方向第 q 階的動量(moment) m_{pq} 為：

$$m_{pq} = \sum_{x=0}^M \sum_{y=0}^N x^p y^q f(x, y)$$

令中心點座標為 (\hat{x}, \hat{y}) ，則正規化中心動量(normalizing central moment) η_{pq} 為：

$$\eta_{pq} = \frac{\sum_{x=0}^M \sum_{y=0}^N (x - \hat{x})^p (y - \hat{y})^q f(x, y)}{\left[\sum_{x=0}^M \sum_{y=0}^N f(x, y) \right]^{[(p+q)/2+1]}}$$

我們所使用來表示物件形狀的七個動量不變值(moment invariants) 分別代表不同的意義，以 $\phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6, \phi_7$ 表示，定義如下式：

$$\begin{aligned} \phi_1 &= \eta_{20} + \eta_{02}, \\ \phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2, \\ \phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{32})^2, \\ \phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{32})^2, \\ \phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\ &\quad (3\eta_{21} - \eta_{03})(\eta_{21} - \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2], \\ \phi_6 &= (\eta_{20} + \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + \\ &\quad 4\eta_{11}^2(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{32})^2, \\ \phi_7 &= (\eta_{21} - 3\eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\ &\quad (\eta_{30} - 3\eta_{21})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]. \end{aligned}$$

我們可以利用 $\phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6, \phi_7$ 這七個動量不變值來評估兩物件的形狀差異，作法如下：假設有二不同形狀的物件 O_1 與 O_2 ，代表兩物件形狀的動量不變值分別為 $\phi_1^{O_1}, \dots, \phi_7^{O_1}$ 與 $\phi_1^{O_2}, \dots, \phi_7^{O_2}$ ，則 O_1 與 O_2 形狀差異值定義為：

$$\Delta_M(O_1, O_2) = \sqrt{\sum_{i=1}^7 (\phi_i^{O_1} - \phi_i^{O_2})^2} \quad (2)$$

3.3 2D Vector

2D String[1]是用來表示空間關係索引的方法，有許多空

間關係的相關研究均採取 2D String 的方法，不過由於其比對方式是採符號的(symbolic 處理方式，若要與其他數值型的影像特性相結合將有表示法與索引上的困難，並且相關的比對處理相當費時。如 Chang[2]的方法就必需將所記錄的 MBR 坐標轉換成 2D String 再進行空間比對，而比對結果也無法再與原來的評估方法結合做整體排名(ranking)的考量，因此這個問題必需設法另行解決。

2D Vector 由 Chien[3]所提出，其主要作法為以圖像中物件間相對的比例值來記錄其空間關係，並提供具彈性的相似圖像擷取演算法。2D Vector 可以完全提供與 2D String 相同的空間判斷能力，而且更具效率[4]。其中物件先由先前的影像處理技術萃取出，再利用物件的位置與大小，運用 2D Vector 計算物件間的空間差異值，其空間差異評估即可作為查詢的依據。

以圖 1 為例，圖像中有兩個物件 O_1 及 O_2 ，首先對圖像

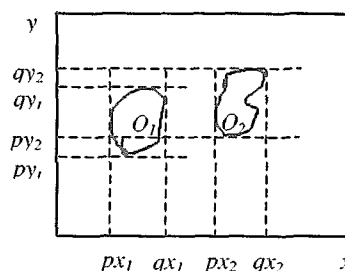


圖 1：2D Vector 圖像物件投影圖

中的物件分別取 x 軸及 y 軸方向的垂直投影， px_i 和 qx_i 為物件 O_i 於 x 軸上的投影邊界坐標， py_i 和 qy_i 為物件 O_i 於 y 軸上的投影邊界坐標。將物件 O_1 和物件 O_2 於圖像中在 x 軸和 y 軸方向的絕對距離定義為 Dx_{12} 和 Dy_{12} ，其值如下：

$$Dx_{12} = px_2 - px_1, \quad Dy_{12} = py_2 - py_1.$$

另外，令 Ix_i 和 Iy_i 為物件 O_i 於 x 軸方向和 y 軸方向的投影大小 (i 為物件之編號)，可以得到

$$Ix_i = qx_i - px_i, \quad Iy_i = qy_i - py_i.$$

假設以 O_1 為參考物件，定義以下的相關關係：

$$\begin{aligned} \alpha_x &= \frac{Dx_{12}}{Ix_1}, \beta_x = \frac{Ix_2}{Ix_1}; & \alpha_y &= \frac{Dy_{12}}{Iy_1}, \beta_y = \frac{Iy_2}{Iy_1}; \\ \gamma_x &= \frac{\alpha_x + \beta_x}{\beta_x + 1}; & \gamma_y &= \frac{\alpha_y + \beta_y}{\beta_y + 1}; \end{aligned}$$

則兩物件 O_1 及 O_2 對於 x 軸與 y 軸的相對距離關係差異值 $f(X_{12})$ 與 $f(Y_{12})$ 定義為：

$$f(X_{12}) = \begin{cases} f_1 = \frac{1}{(1 - \gamma_x)} & \text{if } \alpha_x + \beta_x < 0; \\ f_2 = (1 - \frac{1}{\gamma_x}) + 3 & \text{if } \alpha_x > 1; \\ f_3 = 2\gamma_x + 1 & \text{otherwise}; \end{cases}$$

$$f(Y_{12}) = \begin{cases} f_1 = \frac{1}{(1-\gamma_y)} & \text{if } \alpha_y + \beta_y < 0; \\ f_2 = (1 - \frac{1}{\gamma_y}) + 3 & \text{if } \alpha_y > 1; \\ f_3 = 2\gamma_y + 1 & \text{otherwise}; \end{cases}$$

因此對於兩張不同的圖像 P_1 與 P_2 ，假設分別包含兩個不同的物件 O_1 及 O_2 ，而兩圖中的物件對於 x 軸與 y 軸的相對距離關係評估值分別為 $f(X_{12}^A)$ 、 $f(Y_{12}^A)$ 與 $f(X_{12}^B)$ 、 $f(Y_{12}^B)$ ，則 P_1 與 P_2 中物件 O_1 及物件 O_2 的空間關係差異值定義如下：

$$\Delta_s(O_{12}^A, O_{12}^B) = \sqrt{[f(X_{12}^A) - f(X_{12}^B)]^2 + [f(Y_{12}^A) - f(Y_{12}^B)]^2} \quad (3)$$

4 影像資料建立與相似影像擷取演算法

相似影像擷取(similarity image retrieval)的主要目的為由影像資料庫中擷取出與使用者所查詢或要求的相似圖像來。在以內容為主的影像擷取技術中，資料庫建立時，除了儲存影像本身外，需要儲存代表影像內容的相關特性資訊，並加以製作索引，利用這些資訊及相關索引，在做相似比對(similarity measure)時才能避免耗時重複的影像處理與全域搜尋，進而在龐大的影像資料庫中有效率的評估可能與查詢圖像相似的圖像以及判斷其相似程度，以擷取出相似程度高的結果。以下詳述自動影像資料庫建立方法與相似影像擷取演算法。

4.1 影像資料的建立

在我們的方法中，將每一影像視為由數個物件所組成，每個物件以顏色與型狀來代表其相關特性。如上節所述，顏色特性所使用的資訊為每個物件的 LUV 顏色空間值，形狀則由動量不變值所描述，至於 2D Vector 所描述物件的空間關係則可以由物件在影像中的位置與大小等直接經過簡單的推算得到，所以只需記錄物件之 MBR 坐標即可。以下詳述如何利用影像處理技術萃取這些影像特性，並自動建立影像資訊的方法，步驟如下：

(1) 採用 Lin 所提之顏色區塊叢集(clustering)法[7]，將影像做適當的分割(segment)。首先，將圖像分成大小同為 $X_b \times Y_b$ 的小區塊(block)，並設定其顏色為區塊範圍內所有像素的平均顏色。然後，計算每個區塊與鄰近區塊(neighboring blocks)的 LUV 顏色差距，再合併差距最小的兩區塊，而合併後的顏色為兩區塊的平均顏色。重複合併的動作，直到只剩一個全圖的平均顏色區塊。Lin 的方法以合併後的影像與原圖的均方差(mean square error)的變化率來決定最適合的區塊數目。令 Q_k 為合併成 k 個區塊後的圖像與原圖的 MSE，圖像區塊變化率的評估定義如下：

$$\delta(k) = \frac{Q_{k-1} - Q_k}{Q_k - Q_{k+1}} \quad (4)$$

決定最適合的切割區塊標準為取式(4)為最大值 $\max(\delta(k))$ 時的 k 值，該值即為最適合的區塊數目 $knee$ 。例如圖 2 中，(a)為原圖，式(4)計算後的結

果如圖 2(b)所示，其最大值 $\max(\delta(k))$ 落在 $k=7$ ，根據觀察圖 2(b)中 Q_k 與 k 的曲線，可以發現約是在曲線的最大斜率處，叢集之後的圖像如圖 2(c)，分割成為 7 個不同的顏色區塊，每個區塊即可代表圖像中的一個物件，從此案例中可以發現經切割後的物件區塊相當接近人類的視覺結果。

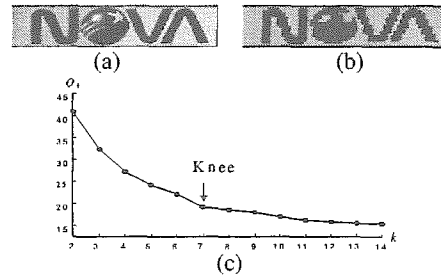


圖 2：(a)原圖；(b)叢集後的圖像 (c) Q_k 曲線圖；

- (2) 計算出代表物件形狀的 7 個動量不變值，以決定每個區塊物件的形狀特性。利用區塊所包含的顏色像素位置，代入不變動量(invariant moment)運算式計算出代表物件形狀的 7 個動量不變值。
- (3) 以多維資料結構型態記錄圖像中物件的 LUV 空間顏色值、MBR 位置與動量不變值於資料庫中，其資料結構表示如下：

$[Object_i, L, u, v, \phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6, \phi_7, LD_x, LD_y, RU_x, RU_y]$

其中 LD_x, LD_y, RU_x, RU_y 代表物件 i 之 MBR 的左下角與右上角坐標。

根據以上的三個步驟，可以自動萃取出代表影像的各個物件之特性，並進而表示一張影像的內容。由於以上所萃取出來的特性其差異值均以 Euclidean distance 來作為評估的依據，此一致性對於建立資料索引(index)將相對的容易，如何建立多維空間索引(multi-dimensional index)的方法亦是一重要的研究主題，由於並非本文的重點，我們在此並不加以深入探討。

4.2 相似圖像擷取演算法

在建立完圖像資料之後，由影像資料庫中擷取出符合使用者需求的圖像需要一相當有效率的相似圖像比對方法。通常使用者須先設定所欲查詢的圖像資訊(query image)，然後根據所設定的要求透過相似比對演算法(similarity matching algorithm)找出相似圖像後，再依相似程度加以排列(ranking)。我們先討論查詢圖像之設定方法，而後說明相似比對演算法。

4.2.1 查詢圖像之設定

在我們的影像擷取方法中，查詢圖像需要設定的物件資訊與資料庫圖像的萃取資訊一樣，必需包含影像物件的顏色、動量不變值、位置與大小。為了方便使用者設定查詢圖像，我們的系統提供一使用者查詢介面，如圖 3 所示，而系統所提供的設定方法主要有兩種：

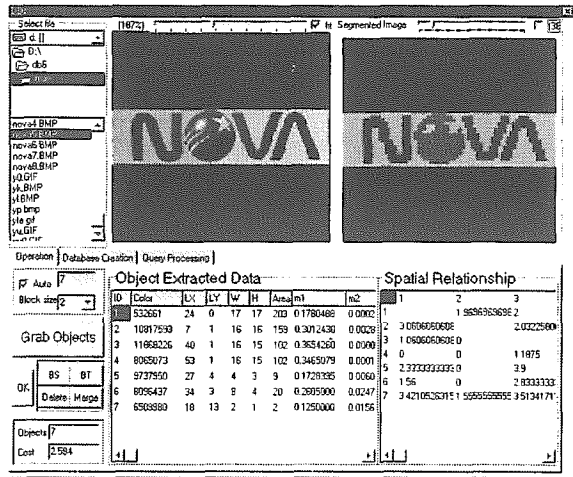


圖 3：使用者查詢介面

1. 範例影像查詢法(query by example; QBE)：給定一欲找的類似圖片樣本，利用前述建立資料庫圖像中萃取出物件資訊的方法，萃取出查詢圖像中物件的顏色值、位置、大小以及動量不變值，然後進行比較的動作。
2. 概圖繪畫查詢法(query by sketch)：若使用者目前並無相關的影像，系統亦允許使用者自行繪出所欲查詢圖像的概圖(sketch)，由概圖獲得相關的查詢資訊然後進行比較。當然，由此法找到的圖像將可以當做 QBE 的範例影像，做再更詳細的查詢。

4.2.2 相似比對演算法

在我們的方法中，為了達到能夠部分查詢(partial query)的目的，查詢圖像與每張資料庫圖像比對之前須找出相對之物件配對(object-pairs)，也就是兩圖中一組至數組最相似的物件。利用這些物件配對，才可以進一步評估查詢圖像與每張資料庫圖像中部分圖像或整張圖像的相似度，以下分別說明物件配對演算法與相似比對演算法。

物件配對演算法(object-pairs matching algorithm)：

我們以顏色與形狀的差異度來評估物件的相似度，進而決定兩物件是否可為配對。假設查詢圖像 P_1 有 m 個物件，分別為 O_1, \dots, O_m ，資料庫圖像 P_2 有 n 個物件，分別為 O_1, \dots, O_n ，則找尋物件配對的步驟如下：

步驟一：建立一 $m \times n$ 矩陣 D_{mn} ，令 $D_{mn} = [d_{ij}]_{m \times n}$ ， $1 \leq i \leq m$ ， $1 \leq j \leq n$ 矩陣中的元素 d_{ij} 為 P_1 中第 i 個物件 O_i 與 P_2 中第 j 個物件 O_j 的物件相對差距值，定義為：

$$d_{ij} = \omega_c \Delta_C(O_i, O_j) + \omega_s \Delta_M(O_i, O_j) \quad (5)$$

其中 $\Delta_C(O_i, O_j)$ 為式(1)計算物件 O_i 與物件 O_j 的顏色差距值， $\Delta_M(O_i, O_j)$ 為式(2)計算兩物件的形狀差距值，而(5)式中的 ω_c 、 ω_s 分別為顏色與形狀差距值的權重(weight)。

步驟二：搜尋 D_{mn} 中差距值最小的元素，假設位於 (a, b) ，則可設定 P_1 中第 a 個物件與 P_2 中第 b 個物件為物件配對。

步驟三：除去 D_{mn} 中的第 a 列與第 b 行，重複步驟(2)，

直到 P_1 或 P_2 沒有未配對的物件，或相似程度太小為止。之後即可得到數組相似物件配對。

在顏色與形狀差距值的權重方面，可依資訊圖像的應用領域而調整。例如，當資訊圖像為自然圖片時，形狀不夠確定，所以可增加顏色的比重，找出顏色分布較接近的圖像。而在商標等邊緣強烈的圖像方面，視覺較受形狀的影響，形狀的權重可以增加。

相似比對演算法(similarity matching algorithm)：

步驟一：利用物件配對演算法將所要查詢的圖像 P_1 與資料庫圖像 P_2 相互配對找出 k 組相似物件配對，假設分別為

$$(O_1^1, O_1^2), (O_2^1, O_2^2), \dots, (O_k^1, O_k^2)$$

步驟二：計算 P_1 與 P_2 的差異度 $Dis(P_1, P_2)$ ，差異度 Dis 之定義如下：

$$Dis(P_1, P_2) = \omega_{sr} \psi_{spatial}(P_1, P_2) + \omega_c \psi_{color}(P_1, P_2) + \omega_s \psi_{shape}(P_1, P_2)$$

其中

$$\psi_{color}(P_1, P_2) = \sum_{i=1}^k \Delta_C(O_i^1, O_i^2);$$

$$\psi_{shape}(P_1, P_2) = \sum_{i=1}^k \Delta_M(O_i^1, O_i^2);$$

$$\psi_{spatial}(P_1, P_2) = \sum_{i=1}^{k-1} \sum_{j=i+1}^k \Delta_S(O_i^1, O_j^2).$$

步驟三：搜尋可能的圖像特性資料，重複步驟二的計算，將之由小而大排序輸出。

以上的相似度評估方法(similarity measure)考慮查詢圖像與資料庫圖像的顏色、形狀與空間關係的差異評估值。其中顏色評估值 ψ_{color} 為配對物件間的顏色差異值的和，形狀評估值 ψ_{shape} 為配對物件間的形狀差異值的和， $\psi_{spatial}$ 則是利用配對物件的位置與大小，代入 2D Vector 所計算出的空間關係評估值，而查詢圖片與資料庫圖像的差異度為三項評估值的加權和。差異度 $Dis(P_1, P_2)$ 定義中的 ω_c 和 ω_s 與尋找物件配對中所設定的權重相同，都代表對形狀或顏色的重視程度，而 ω_{sr} 代表對於物件間的方位關係的重視程度。兩圖的 $Dis(P_1, P_2)$ 值越小者，相似度越高，使用者可以設定一臨界值(threshold)，當 $Dis(P_1, P_2)$ 小於使用者設定之臨界值時，該圖像即是可擷取之相似圖像之一，或是設定欲找尋的相似圖像個數的上限，取其排名最相似者均可。

5 實驗與比較

我們在 PC Windows 環境下製作雛型系統，並對我們所提出的方法進行評估。初期的影像資料庫中，目前儲存 100 張測試圖像，以及經由自動建立影像資訊資料庫後，所萃取出來的圖像特性資訊。

為了評估與比較結果，我們選擇 Mehre[9]的方法建立另一個測試圖像的萃取資訊，並且利用他們的影像擷取

方法與我們的方法做擷取效果的比較。其中為了配合 Mehre 的系統作公平的評判, 測試影像資料庫中的圖像以商標類為主, ω_c 、 ω_s 與 ω_{sp} 等比重的設定值為 0.2、0.4、0.4, 而 Mehre 的系統所有設定值採用其論文之建議值。以下以兩個查詢範例來說明我們所測試的結果。該二範例為方便說明起見, 只列出查詢後最相似的前 7 張圖像的結果。

查詢一為以概圖繪畫查詢法來給定查詢圖像, 如圖 4 所示。對於圖 4 的查詢圖像, 兩個系統的擷取結果及排名如圖 5, 真正相關的圖像為圖 5 中我們方法所擷取出來的前五張。我們發現在這個例子當中, 雖然兩種方法皆可擷取出相似的圖像, 但是以正確性而言, 我們的方法能夠很精確的將結果擷取到, 而且以符合人類的視覺方式排列。而由於 Mehre 的方法只考慮顏色、形狀與面積比例, 導致有些不相關但具有相同影像特性的圖片被選取並認為較相似。我們的方法在顏色、形狀等特性外多考慮物件間的空間關係, 所以在相似度的排列方面, 我們的方法能夠較符合人類的視覺。由此可見物件空間關係對人類視覺意義上的影響相當大, 所以空間方位的特性是有結合的必要。

查詢二則以範例影像查詢法來給定查詢圖像, 圖 6 為第二個查詢圖像, 而影像資料庫中我們發現會影響判斷的兩張重要圖像如圖 7 所示, 其中顏色與形狀與查詢圖像差不多, 而方位作大福的改變。查詢的結果在圖 8, 擷取結果也顯示因為我們的方法考慮空間關係, 使受到的影響較小。另外若沒有考慮方位關係的話, 我們方法所擷取出來的第 4 張影像將不在 Mehre 方法的擷取結果中, 取而代之的是另一張意義上較不相干的圖樣。

6 結論

本文中提出了結合顏色、形狀與圖像物件之空間關係等三種影像內容做為影像資料庫的索引與查詢之依據。我們的方法有以下優點:

1. 能自動萃取與建立較符合人類視覺的影像分割結果與影像特性。
2. 提供具一致性且易於製作索引的影像資訊特性表示法。
3. 具相當有效率的結合顏色、形狀與圖像物件之空間關係特性的相似比對演算法。

由實驗結果與比較得知, 圖像中物件間的空間關係也和其他影像特性一樣左右圖像擷取結果。未來我們將依此一雛型建立更精確的影像測試平台(testbed)與模型(model), 藉以做更深入的效益探討與精確度分析。另外為了再增加影像擷取的效率與精確度及擴大應用範圍, 我們將陸續加入不同的影像特性, 如紋理(texture)等重要特徵的萃取, 以及利用模糊理論(Fuzzy theory)建立之多重解析架構(multiresolution structure)儲存影像物件, 以增加圖像部分查詢(partial query)的準確度。

7 參考文獻

[1] S. K. Chang, Q. Y. Shi and C. W. Yan, "Iconic indexing by 2D -String", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol 9, NO 3, pp413-428, 1987

[2] C. C. Chang, "Color Image Retrieval Based on 2D String", *Master thesis, Department of Computer Science*, National Tsing Hua University, Taiwan, 1998.

[3] B. C. Chien and P. C. Chieng, "A New Approach of Similarity Measurement for Iconic Image Retrieval", in *Proceeding of International Conference on Computational Intelligence and Multimedia Applications*, 9-11, Feb. 1998, Churchill, Australia.

[4] P. C. Chieng, B. C. Chien, "A New Similarity Metric Based on 2D Vector for Iconic Image Retrieval", in *Proceedings of 1998 International Computer Symposium, Workshop on Software Engineering and Database Systems*, 16-19, Dec., 1998, Tainan, Taiwan, pp.85-92.

[5] J. D. Foley, A. V. Dam, S. K. Feiner, and J. F. Hughes, "Computer Graphics: principles and practice," 2nd edition in C, Addison-Wesley, 1996.

[6] R. C. Gonzalez and R. E. Woods, "Digital Image Processing", third edition, Addison-Wesley, Reading, MA, 1992.

[7] H. C. Lin, L. L. Wang, and S. N. Yang, "Color image retrieval based on hidden Markov models," *IEEE Transactions on Image Processing*, Vol. 6, No. 2, 1997.

[8] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, P. Yanker, and C. Faloutsos, "The QBIC project: querying images by content using color, texture, and shape," *Research Report*, No. 9203, IBM Almaden Research Center, 1993.

[9] A. K. Jain and A. Vailaya, "Image Retrieval Using Color and Shape", *Pattern Recognition*, Vol. 29, No. 8, pp. 1233-1244, 1996.

[10] B. M. Mehre, M. S. Kankanhalli and W. F. Lee, "Content-Base Image Retrieval Using A Compost Color-Shape Approach", *Information Processing And Management*, Vol. 34, Issue. 1, 1998, pp. 109-120

[11] J. R. Smith and S. F. Chang, "Vis ually Searching the Web For Content", *IEEE Multimedia*, Vol 3, 12-20, 1997.

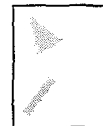


圖 4: 查詢圖像(一) 包含藍色三角形、黃色圓形與粉紅長條形

Rank	我們的方法擷取結果	Mehre 的方法擷取結果
------	-----------	---------------

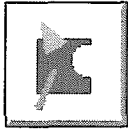
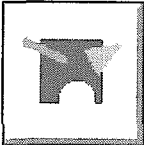
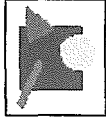


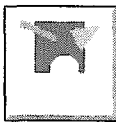
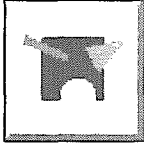
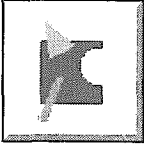




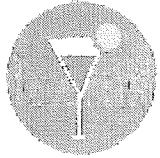
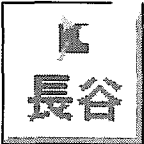
1		
2		
3		
4		
5		
6		
7		

圖 5: 圖 4 之查詢結果




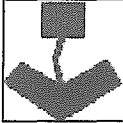






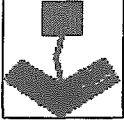



1		
2		
3		
4		
5		
6		
7		

圖 8: 圖 6 之查詢結果

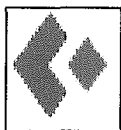


圖 6: 查詢圖像(二)

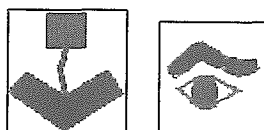


圖 7: 影響判斷之圖像

Rank	我們的方法擷取結果	Mehre 的方法擷取結果
------	-----------	---------------