

行政院國家科學委員會專題研究計畫 成果報告

植基於 VQ 編碼指標與色彩質心空間特徵之影像檢索系統

計畫類別：個別型計畫

計畫編號：NSC93-2213-E-034-002-

執行期間：93 年 08 月 01 日至 94 年 07 月 31 日

執行單位：中國文化大學資訊科學系

計畫主持人：孫振東

報告類型：精簡報告

處理方式：本計畫可公開查詢

中 華 民 國 94 年 10 月 30 日

摘要

影像資訊檢索在許多應用上扮演著重要角色，諸如衛星影像資料庫、地理資訊系統、診斷醫學影像資料庫、及商業商標資料庫，因此，發展一可靠而有效率的影像檢索系統是非常重要的。本計畫提出一個基於影像向量量化(VQ)指標與色彩質心空間位置的影像檢索系統，VQ 是將原影像色彩長方圖再量化成低解析度，資料庫內影像根據 VQ 編碼指標分成多個群組，本系統利用四個色彩質心的相對位置來描述一個影像的色彩空間分布特徵，一個像素的質量即是它的色彩灰階，我們將一個影像分成四個象限，每個象限有一個色彩質心，利用四個色彩質心彼此之間的歐基里得距離作為搜尋影像的索引，實驗結果顯示此方法的影像辨識能力受影像的平移、翻轉、旋轉、及縮放等變化的影響不明顯。

關鍵詞：影像檢索、向量量化、植基於色彩的影像檢索、色彩長方圖、色彩質心

Abstract

Image information retrieval plays a significant role in many application areas, such as satellite image databases, geographic information systems, diagnostic medical image databases, and commercial logo databases. Hence, developing an effective and efficient image retrieval system is very important. This project proposed an image retrieval system based on the VQ (vector quantization) index and spatial locations of color-mass-center in images. VQ is used to construct a lower-resolution version of the original color histograms of images. The images in the database are divided into clusters based on VQ code index. This system describes the color-spatial feature of an image using the relative location of four color-mass-centers. The color-mass of a pixel is defined by the gray level of the pixel. We divide an image into four quadrants. Each quadrant has a color-mass-center. The Euclidean distances between each pair of four color-mass-centers are the indexing key to search the image. Experimental results show that the recognition ability is insensitive to the shift, rotation, and scale variants of images.

Key words: image retrieval, VQ, color-based image retrieval, color histogram, color mass center

一、前言

隨著數位影像技術的進步，數位相機、掃描器等影像捕捉系統已是價廉物美而唾手可得，再加上網際網路的普及，對於資訊的多媒體展現已是基本需求，畢竟，視覺化的表現方式對人類來說遠比文字描述要來的易於接受，而數位影像在多媒體資料裡數量龐大佔了舉足輕重的地位，如何將這些數位影像有效率的管理與快速的存取成為很重要的課題，目前數位影像檢索系統是重要研究項目之一。影像檢索系統可應用於各種不同的領域，例如用於醫療方面：可將病人之X光或超聲波相片以數位影像存檔，除可減少儲存空間，遇到診斷或教學需要時可透過網路快速調出；用於商標圖像方面：可在短時間內由商標資料庫查詢新式商標是否與已註冊商標雷同；其他諸如地球科學衛星照片、指紋圖庫、及人像圖庫等龐大資料的搜尋，皆可透過影像檢索系統的搜尋功能取代耗時的傳統人力查詢。

二、研究目的

本研究提出一個基於影像向量量化(VQ)指標與色彩質心空間位置的影像檢索系統，VQ是將原影像色彩長方圖再量化成低解析度，資料庫內影像根據VQ編碼指標分成多個群組，先根據影像向量量化指標找出影像屬於哪個群組，以避免對全部影像進行比對，加快檢索速度，本系統再利用四個色彩質心的相對位置來描述一個影像的色彩空間分布特徵，一個像素的質量即是它的色彩灰階，我們將一個影像分成四個象限，每個象限有一個色彩質心，利用四個色彩質心彼此之間的歐基里得距離作為搜尋影像的索引，此方法的影像辨識能力受影像的平移、翻轉、旋轉、及縮放等變化的影響不明顯。

三、文獻探討

早期所發展的影像檢索系統，大多以文字特徵(textual features)來搜尋影像資訊，例如使用影像的標題、檔名、或關鍵字來描述影像的內容，再將這些文字特徵當作索引，和圖像一起儲存到資料庫以便往後查詢。因為必須以人工方式為資料庫裡的每張圖像標示關鍵字，以每個資料庫動輒數萬張的圖檔來看，這種方法是非常的耗費人力與時間，另外受人類主觀認知的影響，不同的人對於同一張影像的內容會有不同的描述，而且同一張影像中物體的大小、顏色、形狀、以及物體與物體之間的關係皆須以文字描述，使用關鍵字的方法將愈形複雜與不適宜。目前的影像檢索系統均朝向以影像內容為基礎(content-based)的檢索技術發展，亦即抽取能夠代表影像內容的特徵來作索引的條件[1][2]，這些特徵包括：影像色彩(color)、物件的細部紋路結構(texture)、物件的形狀(shape)、以及物件間的空間關係(spatial relationship)。另為解決龐大影像資料在儲存與傳輸上的空間與時間問題，資料壓縮技術被廣泛應用，一般分成在空間域(Spatial domain)或轉換域(Transform domain)來處理，例如空間域的向量量化(VQ; Vector Quantization)[3]與碎形(Fractals)技術[4]，轉換域的離散餘弦轉換(DCT; Discrete Cosine Transform)[5]與小波轉換(DWT; Discrete Wavelete Transform)[6]，影像檢索

系統亦朝向以該領域獲得之參數特徵來作索引的條件[7][8]，這些方法皆需有大量的計算。

本研究提出兩階段檢索系統，首先基於影像向量量化(VQ)指標檢索，先找出類似的影像群組，因此無需對資料庫內全部影像進行比對，再根據色彩質心空間位置特徵在已找出的影像群組內檢索影像。VQ 是將原影像色彩長方圖再量化成低解析度，資料庫內影像根據 VQ 編碼指標分成多個群組，本系統再利用四個色彩質心的相對位置來描述一個影像的色彩空間分布特徵，一個像素的質量即是它的色彩灰階，我們將一個影像分成四個象限，每個象限有一個色彩質心，利用四個色彩質心彼此之間的歐基里得距離作為搜尋影像的索引，實驗結果顯示此方法的影像辨識能力受影像的平移、翻轉、旋轉、及縮放等變化的影響不明顯。

四、研究方法

所有的顏色都是由紅(R)、綠(G)、藍(B)這三種原色依比例混合而成的，而大多數影像的色彩分布特徵則是使用色彩長方圖(color histogram)來描述：

$$h_c[l] = \sum_{i=0}^{W-1} \sum_{j=0}^{H-1} \delta(I^q(i, j), y_l), (l = 0, \dots, L-1) \quad (1)$$

W 和 H 為圖的長和寬， $\delta(., .)$ 是 delta function，我們將所有色彩向量量化成 L 個色彩(亦即長方圖 bin 的個數)， y_l 則是第 l 個向量量化的色彩， $I^q(i, j)$ 是在位置 (i, j) 的像素經量化的色彩， $h_c[l]$ 即是影像中色彩可量化成 y_l 的像素總數量。為考量不同大小影像的比較，將長方圖正規化：

$$h = \frac{h_c}{\left(\sum_{l=0}^{L-1} |h_c[l]|^2 \right)^{1/2}} \quad (2)$$

由於整體色彩特徵描述只能統計全圖的色彩數量，無法明確的表示各種顏色的空間分布，文獻[9][10][11][12]建議將圖分割成不重疊的區塊來做色彩特徵描述，將色彩與空間資訊結合，如此可以提高影像鑑別的精確度，根據色彩特徵進行相似度量測的方法有多種[13][14][15]，做法簡單而有效率的主要有兩種：

(1) Histogram Intersection

$$D_1 = (q, t) = \sum_{l=0}^{L-1} |h_q(l) - h_t(l)| \quad (3)$$

h_q 及 h_t 分別是查詢影像與目標影像的色彩長方圖，當運算總合越小的時候，就越符合所要尋找的目標。這樣的做法較為簡單，但是遇到顏色相近的圖像的時候，便會發生錯誤的搜尋。

(2) Histogram Euclidean Distance

$$D_2(q, t) = \sum_{l=0}^{L-1} [h_q(l) - h_t(l)]^2 \quad (4)$$

D_2 其實是 D_1 的二次方，總合仍然是越低越接近原圖，但當查詢影像與目標影像

的色彩長方圖落差不大的時候，二次方會將其落差明顯的表示出來，增加了搜尋的精確性。由於空間域色彩長方圖檢索方法做法簡單而有效率，因此，本研究提出一影像檢索系統，以顏色為特徵值的直方圖來描述影像，並藉著影像中的顏色空間分佈的狀況來找到最相似的圖。圖 1 是影像檢索系統架構圖

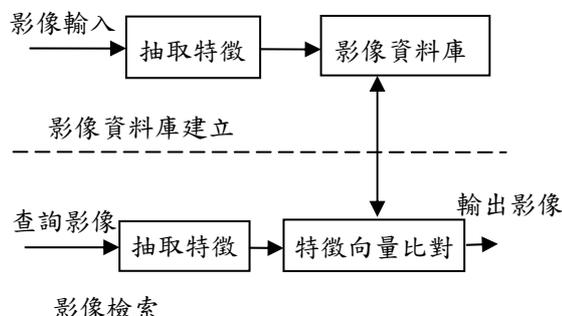


圖 1、系統架構圖

影像檢索系統包含影像資料建立與影像檢索兩個部分。影像資料建立部分在於擷取出每一個影像資料的特徵，並將其儲存於資料庫之中，當作影像查詢系統之索引。影像檢索部分則是當使用者輸入一欲檢索之影像，並且抽取該影像之特徵，同時與影像資料庫所儲存之影像特徵做比對後，輸出特徵值最相像的影像並傳回給使用者。本論文所提出的影像檢索架構分為兩個部分：第一個部分為影像資料建立部分，第二部分是資料檢索部分。

(1) 影像資料建立部分

除影像資料本身外，我們需要建立碼簿(Codebook)、編碼指標(Code index)、及色彩質心(Center of color mass)，首先是建立碼簿 CB，碼簿 CB 內含 K 個碼字(Code words)，亦即找出 K 個長方圖，其可以代表所有的長方圖，每個碼字長度(Code length)為 L ， L 即是色彩長方圖的 Bin 個數， L 愈大則其色彩長方圖每一 Bin 所代表之顏色範圍愈小，建立碼簿的方法是利用圖庫訓練而得，首先從圖庫中挑出若干代表性的圖片或全部，計算出它們的長方圖，以這些長方圖為樣本空間，再根據 LBG (Linde - Buzo - Gray)演算法來產生向量量化(VQ)編碼碼簿，碼簿的重要參數有碼簿大小(Codebook size)及碼字長度(Codeword length)，每一碼字予以編號，稱之為編碼指標 k ，我們將圖庫內每一張影像之長方圖與碼簿內各碼字做距離計算，以最小距離之碼字編碼指標儲存於資料庫，在計算色彩質心方面，我們以一般物理求質心原理為之，僅將質量以灰階值取代，首先將圖片平均分成四個大區塊(四個象限， Q_1 、 Q_2 、 Q_3 、 Q_4)如圖 2 所示，各區塊分別以 R、G、B 三種顏色為色彩物質，求其四個質心，在位置 (x, y) 的像素的色彩質量(亦即 RGB 三灰階)為 $(m_{xy}^r, m_{xy}^g, m_{xy}^b)$

$$x_i^r = \frac{\sum_{x,y \in Q_i} m_{xy}^r x}{\sum_{x,y \in Q_i} m_{xy}^r}, \quad y_i^r = \frac{\sum_{x,y \in Q_i} m_{xy}^r y}{\sum_{x,y \in Q_i} m_{xy}^r}, \quad i = 1, 2, 3, 4 \quad (5)$$

$$x_i^g = \frac{\sum_{x,y \in Q_i} m_{xy}^g x}{\sum_{x,y \in Q_i} m_{xy}^g}, \quad y_i^g = \frac{\sum_{x,y \in Q_i} m_{xy}^g y}{\sum_{x,y \in Q_i} m_{xy}^g}, \quad i = 1,2,3,4 \quad (6)$$

$$x_i^b = \frac{\sum_{x,y \in Q_i} m_{xy}^b x}{\sum_{x,y \in Q_i} m_{xy}^b}, \quad y_i^b = \frac{\sum_{x,y \in Q_i} m_{xy}^b y}{\sum_{x,y \in Q_i} m_{xy}^b}, \quad i = 1,2,3,4 \quad (7)$$

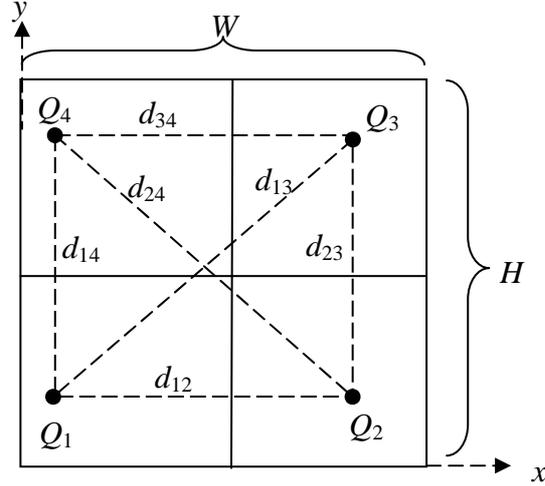


圖 2、質心與距離

再求各質心間之距離，並予正規化， W 與 H 分別是圖的寬與高，正規化之目的是避免因影像大小所造成的差異性，

$$d_{ij}^r = \sqrt{(x_i^r - x_j^r)^2 + (y_i^r - y_j^r)^2} / \sqrt{W^2 + H^2}, \quad i < j, \quad i, j = 1,2,3,4 \quad (8)$$

$$d_{ij}^g = \sqrt{(x_i^g - x_j^g)^2 + (y_i^g - y_j^g)^2} / \sqrt{W^2 + H^2}, \quad i < j, \quad i, j = 1,2,3,4 \quad (9)$$

$$d_{ij}^b = \sqrt{(x_i^b - x_j^b)^2 + (y_i^b - y_j^b)^2} / \sqrt{W^2 + H^2}, \quad i < j, \quad i, j = 1,2,3,4 \quad (10)$$

因此，每一張影像存於資料庫的特徵值為

$$(k, d_{12}^r, d_{13}^r, d_{14}^r, d_{23}^r, d_{24}^r, d_{34}^r, d_{12}^g, d_{13}^g, d_{14}^g, d_{23}^g, d_{24}^g, d_{34}^g, d_{12}^b, d_{13}^b, d_{14}^b, d_{23}^b, d_{24}^b, d_{34}^b) \quad (11)$$

以 100×100 之黑白圖為例(如圖 3 所示)，四個質心分別在座標(17, 20)、(65, 33)、(57, 88)、及(33, 64)，而彼此間正規化距離 $d_{12} = 0.351$ 、 $d_{13} = 0.557$ 、 $d_{14} = 0.331$ 、 $d_{23} = 0.393$ 、 $d_{24} = 0.315$ 、及 $d_{34} = 0.240$ 。

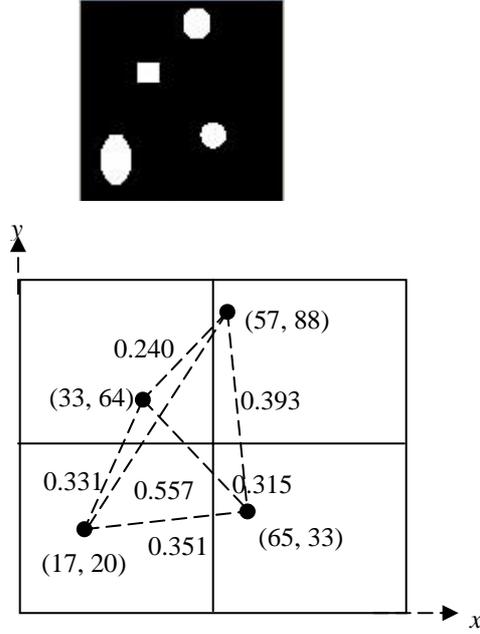


圖 3、質心與距離實例

(2) 資料檢索部分

當使用者提出查詢影像要求時，此一查詢影像經由資料庫建立時相同的前處理程序抽取出特徵後，在與影像資料庫中的特徵向量進行比對，我們經由兩階段比對，首先比對指標 k ，找出資料庫中具相同 k 值的群組，再計算該群組中的每一影像質心距離向量與查詢影像質心距離向量的距離(Euclidean Distance) D ，得到距離最小的一組排序的影像會被輸出到使用者介面。

目標影像特徵向量：

$$\begin{aligned} &(k, d_{12}^{rt}, d_{13}^{rt}, d_{14}^{rt}, d_{23}^{rt}, d_{24}^{rt}, d_{34}^{rt}, \\ &d_{12}^{gt}, d_{13}^{gt}, d_{14}^{gt}, d_{23}^{gt}, d_{24}^{gt}, d_{34}^{gt}, \\ &d_{12}^{bt}, d_{13}^{bt}, d_{14}^{bt}, d_{23}^{bt}, d_{24}^{bt}, d_{34}^{bt}) \end{aligned} \quad (12)$$

查詢影像特徵向量：

$$\begin{aligned} &(k, d_{12}^{rq}, d_{13}^{rq}, d_{14}^{rq}, d_{23}^{rq}, d_{24}^{rq}, d_{34}^{rq}, \\ &d_{12}^{gq}, d_{13}^{gq}, d_{14}^{gq}, d_{23}^{gq}, d_{24}^{gq}, d_{34}^{gq}, \\ &d_{12}^{bq}, d_{13}^{bq}, d_{14}^{bq}, d_{23}^{bq}, d_{24}^{bq}, d_{34}^{bq}) \end{aligned} \quad (13)$$

相似度(Euclidean Distance)：

$$D = \frac{1}{\sqrt{18}} \left[\sum_{\substack{i < j \\ i, j = 1, 2, 3, 4}} (d_{ij}^{rq} - d_{ij}^{rt})^2 + \sum_{\substack{i < j \\ i, j = 1, 2, 3, 4}} (d_{ij}^{gq} - d_{ij}^{gt})^2 + \sum_{\substack{i < j \\ i, j = 1, 2, 3, 4}} (d_{ij}^{bq} - d_{ij}^{bt})^2 \right]^{1/2} \quad (14)$$

五、實驗結果與討論

(一) 實驗結果

我們可依需要建立不同種類的碼簿，本研究建立了三個碼簿，碼簿 CB₁ 之參數為： $L_1=1024$ 與 $K_1=256$ ，碼簿 CB₂ 之參數為： $L_2=512$ 與 $K_2=256$ ，碼簿 CB₃ 的參數為： $L_3=256$ 與 $K_3=256$ ，碼簿愈大表示圖庫內依顏色特徵被分的群組數愈多，搜尋所需時間較少，碼字長度愈長表示顏色被分的愈多，搜尋結果較精確，我們選用 1000 張測試圖片(來源：<http://wang.ist.psu.edu/~jwang/test1.zip>)，訓練出上述三種碼簿，無論根據何種碼簿，圖 4 之待查詢影像，經二階段檢索結果輸出最相似的前 20 張影像，如圖 5 所示，



圖 4、查詢影像



圖 5、最相似的前 20 張影像

表 1 列出此 20 張影像與查詢影像之相似度。我們將(13)式之各質心距離重新組合，可以檢索到查詢影像旋轉 90 度、180 度、及 270 度的影像，或上下顛倒及左右翻轉的影像。

表 1、對應各輸出圖之相似度

0	0.05	0.07	0.07	0.08
0.08	0.08	0.08	0.10	0.10
0.10	0.12	0.12	0.13	0.13
0.13	0.13	0.13	0.14	0.14

(二)討論

本系統採兩階段檢索，第一階段先用 VQ 技術將資料庫中差異度較大的影像先排除，因此，與比對資料庫內所有影像之傳統方法相較，可以節省大量時間，第二階段再利用色彩質心的幾何相對位置作剩下的影像比對，由於是質心的幾何相對位置，因此，可以克服影像的平移、翻轉、和旋轉，而正規化的距離則可以克服影像的縮放，經實驗證明本論文所提影像檢索方法可行而有效率，另參考文

獻[3]所提出的以 VQ 方式先將影像壓縮後，再以編碼指標為長方圖的 BIN，將影像內相同編碼指標的區塊數作為 BIN 的高度，此方法須對圖庫內所有影像作比對，而本研究所提方法僅對圖庫部分影像作比對，顯然較為有效率。

六、參考文獻

- [1] T. S. Hung and Y. Rui, "Image retrieval: Past, present, and future," Proc. Of Int. Symposium on Multimedia Information Processing, Dec. 1997.
- [2] V. Castelli and L. Bergman, IMAGE DATABASES: Search and Retrieval of Digital Imagery, John Wiley & Sons, Inc., 2002.
- [3] F. Idris and S. Panchanathan, "Image indexing using vector quantization," Proc. SPIE: Storage & Retrieval for image and Databases III, 2420, 373-380, 1995.
- [4] M.F. Barnsley and L.P. Hurd, Fractal Image Compression, AK Peters Ltd., Wellesley, Mass., 1993.
- [5] K.R. Rao and Y. Yip, Discrete Cosine Transforms – Algorithms, Advantages, Applications, Academic Press, New York, 1990.
- [6] M. Antonini et al., "Image coding using wavelete transform," IEEE Trans. Image Pcess. Vol.1(2), 205-220, 1992.
- [7] F. Idris and S. Panchanathan, "Image indexing using wavelete vector quantization," Proc. SPIE: Digital Image Storage Archiving Systems, 2606, 369-275, 1995.
- [8] M.Shneier and M.A. Mottaleb, "Exploiting the JPEG compression scheme for image retrieval," IEEE Trans. Pattern Anal. Machine. Intell. 18(8), 849-853, 1996.
- [9] M. Stricker and A. Dimai, Color indexing with weak spatial constrains, in Symposium on electronic imaging: Science and Technology – Storage & Retrieval for image and Databases IV, Proc. SPIE 2670 IS&T/SPIE, 1996.
- [10] W. Hsu, T.S. Chua, and H.K. Pung, "An integrated color-spatial approach to content-based image retrieval, Proc. ACM Intern. Conf. Multimedia (ACMMM), 1995.
- [11] C.E. Jacobs, A. Finkelstein, and D.H. Salesin, "Fast multiresolution image querying," ACM SIGGRAPH, Computer Graphics Proceedings, Annual Conference Series, Los Angles, CA, 1995, pp. 277-286.
- [12] 以顏色複雜度與顏色空間分佈特徵為基礎的影像查詢系統；陳志雅；2003/1/15；朝陽科技大學資訊管理學系碩士論文。
- [13] M.J. Swain and D.H. Ballard, "Color indexing, Int. J. Comput. Vis., 1991.
- [14] W. Niblack, et al., "The QBIC project: Querying image by content using color, texture, and shape," Storage and Retrieval for Image and Image and Video Databases, SPIE, 1993.
- [15] J.R. Smith, "Query vector projection access method," IS&T/SPIE Symposium on Electronic Imaging: Science and Technology-Storage & Retrieval for Image and Video Databases VII, San Jose, Calif., January 1999.