

基於內涵樣式比對之概念導向式影像擷取系統

Concept-Oriented Image Retrieval System Based on Content-Pattern Matching Techniques

蘇家輝¹、翁欲盛³、曾新穆²

¹開南大學資訊管理學系

²成功大學資訊工程學系

³財團法人資訊工業策進會創新應用服務研究所

¹bb0820@ms22.hinet.net

²tsengsm@mail.ncku.edu.tw

³bow4989@iii.org.tw

摘要

近年來，隨著數位影像攝影器材的普及化，加上記憶體儲存裝置之容量與日俱增，數位資料的取得和存取已經變得極為普遍，只要隨手一部數位相機或攝影機，人們就可以將身邊所有的一景一物予以蒐集紀錄，且這些數位科技產品都以簡單、方便操作為設計目標，相對的使得人們能擁有的數位資料也就跟著變得越來越多了。然而讓人煩惱的事情也因此發生了，人們開始思考：「在大量數位資料被記錄下來的同時，我要如何有效的運用這些資料，並能從眾多的資料下，迅速找到目前所需要的一筆或多筆我需要的資料呢？」

過去，人們運用主題式瀏覽或關鍵字來查詢數位資料，然而這些方法對於影像的搜尋卻存在了無可避免的缺點：無法讓使用者即時上傳所喜好之圖片完成搜尋。於是，近年來已發展出一種以內涵式為基礎之影像擷取(Content-Based Image Retrieval, CBIR)技術為基礎的影像搜尋系統，然而此種內涵式影像擷取方法亦面臨到一嚴重的瓶頸：需耗費較長的比對時間，且違反人類自然語意存取方式。所以，針對上述缺點，本研究結合影像內涵與語意存取的概念，並運用樣式比對的技術，成功地將視覺特徵與人類語意連結起來，並實作出一架構在網路平台上的「概念導向式影像擷取系統」，讓使用者可以隨時隨地透過網路，透過上傳圖片或語意，有效地執行概念式之影像存取。

關鍵詞：概念導向式影像擷取技術、內涵式影像擷取技術、樣式比對技術

1. 前言

隨著電腦與網路技術的快速發展，數位影像的技術越來越成熟，儲存裝置的容量也因應人們的需求而與日俱增，如何讓使用者能從龐大的資料量中迅速的取得自己想要的資訊，已成為現今熱門的研究方向之一。追溯傳統的關鍵字索引影像的搜尋方式，例如使用影像的檔名或相關文字敘述來描述影像的內容，然後將這些文字當作索引值與影像一併儲存下來。由於每張影像都必須用人工的方式建立相關的索引值，因此對於現今資訊爆炸的時代，此種方法已經不敷使用了。另外，對於一些比較特殊的影像，我們很難用非常合適的文字來對圖片下註解，使用者也很難正確的提出關鍵字作索引，因此，往往造成其搜尋結果並無法滿足使用者真正之查詢需求。

日常生活中，由於人類的需求使然，數位相機與儲存裝置技術得以迅速的發展，數位攝影器材使得人們不必再擔心的底片不足或是拍攝出錯的問題，任何時間都可以刪除重來，畢竟時間不能停留，能保存的就惟有那最真實、最具價值的畫面，人們都希望透過影像的保存，留下值得回憶的畫面，大量的影像蒐集已經變成一種很簡單的事情，這也直接推動了數位相機技術的發展，使得無論在質量或是可儲存量上，都有很大的突破。將高畫質影像保存下來的代價莫過於記憶體容量的占據，然而值得慶幸的是，人們對記憶體容量需求的提升，也造就記憶體技術的蓬勃發展。如今，我們可以不必花費太多金錢就能取得一個龐大可觀的記憶體容量，因此對於現今社會來說，儲存空間已經不是問題了，但高畫質影像豐富的內容訊息雖然能夠提供一個很好的搜尋資訊，但是礙於在這龐大的資料量運算的處理，往往就造成了搜尋時間的延長而使得系統效率降低。因此提供一個快速的影像搜尋平台，是當今社會所迫切需要的。

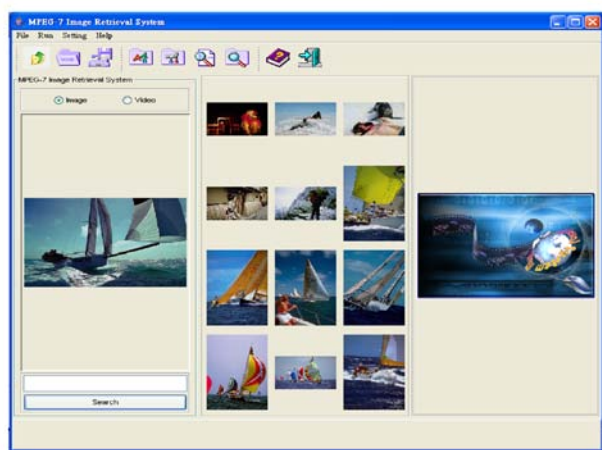


圖 1：內涵式影像擷取系統範例。

近年來，國內外發展出一種以影像內涵為基礎(Content-Based Image Retrieval，如圖 1 所示，簡稱CBIR)[2][5][9][10]的搜尋系統。此種系統讓使用者能夠提供一張影像進行相關性的搜尋，而這種搜尋方式較常用到的技術，就是直接以低階特徵值作相似度的運算，藉此找出類似圖片。但這種技術最大的缺點是需耗費較大的運算時間成本，嚴重的

降低了搜尋效率，而這種方法的搜尋結果也往往不盡人意。此外，事實上此種內涵式的影像搜尋系統並不符合人類語意存取方式，故現行之影像搜尋系統皆以關鍵字為基礎而執行影像搜尋，例如：Google Image Search Engine 及 Yahoo Image Search Engine 等。如圖 2 所示，此類的影像存取雖較符合人性，但無法滿足使用者希望釐清在視覺與概念上的差異性。例如：在同樣是交通工具-車的搜尋情況下，某些人希望找黑色的車，有些人喜歡白色的車。當這種情形發生時，最有效的解決方式便是結合影像之視覺特徵與人類之語意概念做影像搜尋。於是，本論文係針對此問題提出一套結合影像內涵與語意存取的概念，將使用者上傳的影像轉化成語意，使得搜尋方式更人性，搜尋效果更完善。如此一來，使用者不必再為想不出恰當或確切的關鍵字而煩惱，又能夠迅速的從龐大的資料群中，搜尋出自己所需要的資訊。本研究已實際架構在網路上¹，並能實際運作。使用者可透過各種通訊工具，藉由網路，隨時隨地進行影像搜尋，實驗亦證明我們的方法能快速找到使用者想要的影像。



圖 2：以文字為基礎之影像擷取系統範例。

2. 文獻探討

在此資訊蓬勃發展的時代，因為網路的蓬勃發展跟多媒體資料的取得容易，多媒體資料的管理就相對顯地重要。傳統的影像搜尋方式大部分是以文字為基礎來進行搜尋。如 Google、Alta Vista，使用者輸入文字當作搜尋索引的目標，進而找到相關的影像，這樣的方式，稱為 Text-Based Image Retrieval System。雖然以文字為基礎的搜尋方法所得的圖片，其內容與人類概念比較相近，但是這種方法除了需要耗費大量的人工來描述外，現實生活中存在著大量的影像，並不是每張影像都會有其相對應的文字說明。於是，針對關鍵字之影像搜尋，有些研究便致力於自動的影像註解，如：[6][7][8]。但是，每個人對同一張圖片的感受不同，也因此不同的使用者會使用不同的文字描述，但他們所尋找的都是同一張影像，這會造成搜尋結果誤差。

¹ <http://140.116.247.190:9833/cbir/>

有鑑於此，另一種影像搜尋系統其發展也開始跳脫純粹由關鍵字搜尋的限制，而是往影像的內涵(Content)著手。至目前為止，愈來愈多將 CBIR 技術結合其它演算法或不同概念的影像搜尋系統紛紛推出。許多專家學者希望能發展自動化且以內容為基礎的影像搜尋系統。但這樣的查詢方式，當資料量巨大時，查詢的結果便可能變成不穩定。為此，許多研究提出更多的想法，如將影像樣本查詢方式，再細分的以實際影像為樣本，或是由使用者主動描繪的意涵影像來查詢。或是將使用者的習性加入考量的查詢方式，或主動要求使用者反饋(Feedback)等方法。甚至，結合兩種查詢方式的系統也孕育而生。

QBIC(Query By Image Content) [4]是由 IBM 公司的研究中心所發展出來的，可以說是影像內容檢索的代表，可用來查詢的條件依據包括關鍵字和影像的顏色、紋理、形狀的等特徵。QBIC 具有檢索大型影像資料庫的技術，而且 QBIC 提供了顏色及形狀兩種查詢介面，讓使用者可以在色盤中直接挑選顏色，或是在繪圖區畫出期望影像中所需要顏色物件的方塊，並設定其代表的形狀。

CBIRD (Content-Based Image Retrieval from Digital libraries) [1]是由加拿大 Simon Fraser 大學 Vision and Media 實驗室所開發的系統，可由關鍵字或挑選顏色、紋理特徵與光度的不變性來進行查詢。並且有提供工具與繪圖區(1x1,2x2,4x4,8x8)可讓使用者自行規畫影像的顏色配置、紋理的配置以及顏色比例控制、模組設定做為查詢的依據。

Djeraba 在[3]提出將關聯式規則探勘法套入 CBIR 系統中。該篇論文所提出之方法為，將圖片分割成數個區域，擷取這些區域之圖片顏色與紋理兩特徵。先將這些特徵利用分群方法分成數群，藉以縮小不同特徵值之數目，並應用關聯式規則探勘法探勘出顏色與紋理兩特徵之關聯法則。然而，該論文影像分割與關聯式法則探勘之計算較為複雜。

總結來看，以上內涵式影像搜尋系統，多會用到顏色、紋理、形狀等影像特徵，此外都會結合更多的人為控制資訊，讓使用者可以調整多種參數，或進行繪畫。但是在實際應用上，對使用者來說，可能反而是種負擔。絕大多數的使用者只想用簡單的方法，快速的就得到想要的結果，想想平日使用各大網路搜尋引擎，雖然都俱備進階搜尋的動作，但是使用的次數卻是少之又少。因此，我們的系統將著重在適度的語意功能與搜尋速度的追求，並利用快速樣式比對的方式，將最快、最正確的搜尋結果交給使用者。

3. 概念導向式影像擷取系統建置

3.1 系統架構

傳統的影像搜尋系統，多為以文字為基礎之影像搜尋系統，其重點皆著重於文字與文字之間的比對，如：Google Image Search Engine 及 Yahoo Image Search Engine 等。當使用者上傳一查詢關鍵字，系統便至圖片資料集裡找出那些具有與使用者查詢關鍵字相同註解的圖片。使用者無法上傳圖片，亦無法與系統互動，常造成搜尋結果的不正確。針對這些衍生的問題，本論文提出一「概念導向式影像擷取系統」，如圖 3 所示，其主要功能可分為下列數項：

- A. 第一部份為離線前置作業(offline preprocessing phase)，主要工作是搜集圖片資料與建立特徵樣式資料庫。首先，我們將影像資料庫中的影像特徵值擷取出來，

接著透過分群演算法將圖片轉成一字串，稱為樣式字串(pattern string)，最後將每一字串存入樣式字串資料庫(pattern-string database)。

- B. 第二部分為圖片搜尋作業(online image retrieval phase)，負責分析樣本圖片並從圖片資料集中搜尋相似之圖片。這主要是透過快速樣式比對(pattern matching)技術，結合使用者所選擇的圖片特徵與低階特徵值分群資訊所達成的。
- C. 第三部分為使用者回饋作業(user feedback phase)，主要是與使用者進行互動，提供搜尋結果及圖片相關語意。當圖片蒐尋完後，系統會分析搜尋結果，提供使用者相關的語意概念，以做下一步更精確的語意圖片蒐尋。

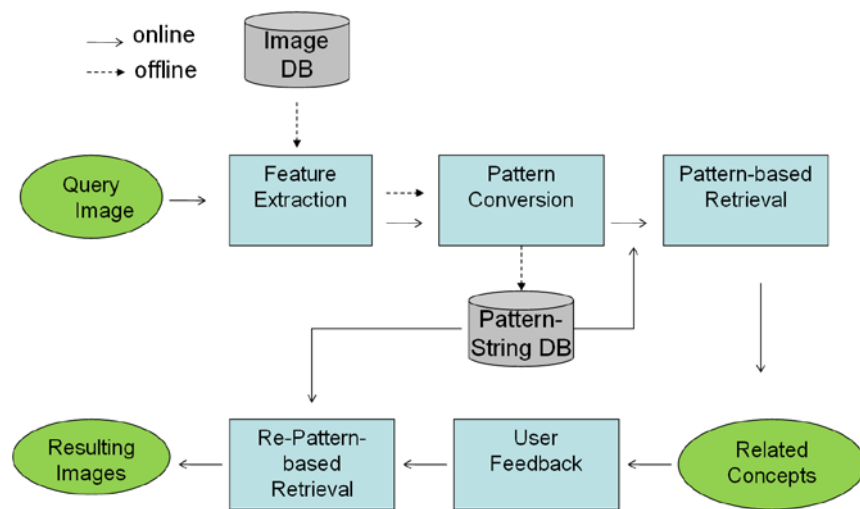


圖 3：系統架構。

透過上述的功能，本系統可彌補內涵式及文字式影像擷取系統的缺點，其主要特色茲敘述如下：

- A. 高效率：透過樣式比對技術，影像之間的比對複雜度大幅降低，使得影像存取更具效率。
- B. 智慧性：使用者查詢之影像，透過擷取分析，被有效的轉成人類語意概念。
- C. 人性化：影像查詢不再侷限於影像低階特徵的比對，經由影像語意化的過程，影像查詢更貼近人類自然的語意。

3.2 離線前置處理階段

此階段主要目的是將資影像資料庫的影像轉換成樣式字串。其處理流程如圖 4 所示，首先將圖片之 4 種低階特徵值擷取出來，包括：Color Layout、Scalable Color、Homogeneous Texture、Edge Histogram。接著，將影像分群並編碼，於是影像被轉換成樣式字串。最後將轉換後的樣式字串存入樣式字串資料庫，供影像搜尋階段時計算相似度之用。

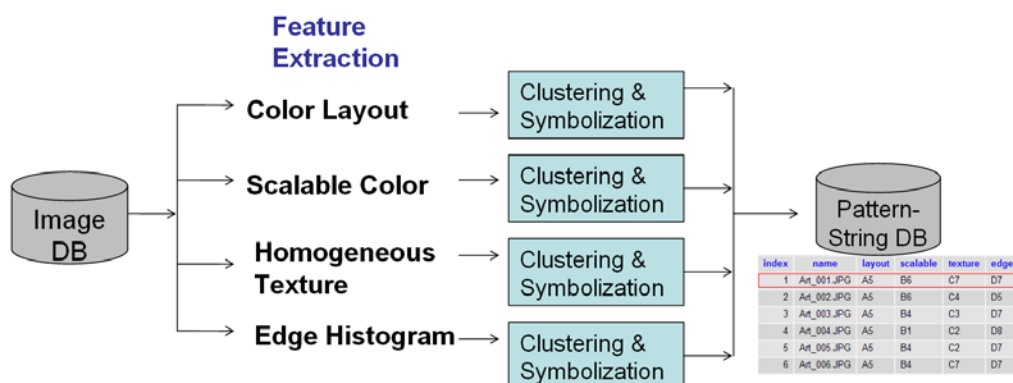


圖 4：離線前置處理流程圖。

3.2.1 擷取低階特徵值

因為搜尋影像時，影像與影像之間相似度比較的關鍵，就是利用影像的低階特徵值或其衍生出來的數據資料。所以，首先我們必須先擷取出影像資料庫中所有影像的低階特徵值。這部份是利用 MPEG-7 官方網站所提供之特徵值擷取軟體 XM Tool 來擷取，其即可將影像內容量化為高維度的向量數值集合。本系統所使用的影像特徵共有 Color Layout、Scalable Color、Homogeneous Texture 和 Edge Histogram 這四種。擷取出來部份結果如表 1、表 2、表 3、表 4 所示，之後還要再利用這些數值進行後續處理動作。

表 1：影像的低階特徵值[Color Layout]。

Name of Image	Value
Bluesky_001.jpg	31 52 13 16 6 15 16 15 16 19 16 16
Bluesky_002.jpg	40 58 6 8 18 7 22 19 22 21 15 8
Bluesky_003.jpg	26 50 20 16 7 13 15 10 18 23 14 12

表 2：影像的低階特徵值[Scalable Color]。

Name of Image	Value
Bluesky_001.jpg	-9 9 59 -2 -3 13 -9 14 3 24 1 12 -23.....
Bluesky_002.jpg	-13 0 37 7 -12 12 3 17 30 22 -2 20 -45.....
Bluesky_003.jpg	-63 51 85 -30 6 13 14 28 28 25 -10 27 -19.....

表 3：影像的低階特徵值[Homogeneous Texture]。

Name of Image	Value
Bluesky_001.jpg	227 0 159 211 213 166 211 215 144 175.....
Bluesky_002.jpg	184 140 173 207 207 193 225 219 176 180.....
Bluesky_003.jpg	69 200 243 244 212 204 206 208 218 235.....

表 4、影像的低階特徵值[Edge Histogram]

Name of Image	Value
Bluesky_001.jpg	0 0 0 0 0 0 0 0 0 0 0 0 2 0 1 3.....
Bluesky_002.jpg	0 1 0 0 0 0 0 0 0 0 0 0 1 0 0.....
Bluesky_003.jpg	4 1 0 0 0 1 0 0 1 0 1 0 0 4 4.....

以下我們針對 4 種特徵值加以詳細描述：

- A. Color Layout：用來描繪出各種顏色的空間分佈狀態。首先影像會被分割為 8x8 的區塊，每個區塊的 Dominant Color 會以 YCbCr 顏色格式儲存。每個色頻裡的 Dominant Color 再套用離散餘弦轉換，以其的系數做為特徵值。此特徵值有 12 個維度，數值介於 0 至 63 之間。
- B. Scalable Color：利用 Haar 轉換方程式，在 HSV 顏色系統下記錄顏色的分佈狀態。優點在於可適應圖片不同的呈現方式。此特徵值有 256 個維度，數值介於-190 至 255 之間。
- C. Homogeneous Texture：描繪區域性的紋路與材質特性。利用 Gabor 過濾函式，做出以紋理的走向趨勢(共 5 種)及大小規模(共 6 種)過濾器來過濾影像；較適合作為搜尋一大堆相似度很高的圖片之查詢依據。此特徵值有 62 個維度，數值介於 0 至 255 之間。
- D. Edge Histogram：表現空間中紋理的 5 種不同類型邊線(4 種有向邊、1 種無向邊)的分佈狀態，可以應用在非一致邊線的影像比對。將影像分割為 16 個子影像(sub-image)，分別計算出在 Null(無方向性)、0 度、45 度、90 度、135 度時的個數。此特徵值有 80 個維度，數值介於 0 至 7 之間。

3.2.2 資料分群

所謂的分群觀念是將一堆分散資料，藉由不同演算法及相似度公式運算，將相似內容之資料聚集為同一群聚，並找出各個群聚的代表中心點，此即分群。這裡我們採用 K-means 分群演算法，此演算法需預先給定分群個數，根據隨機且數量固定的初始中心點，分別以歐幾里德距離(Euclidean Distance)公式對影像的低階特徵值作相似度運算，將最接近的數值歸在同一群中，即可產生最後的分群結果。

所以，此階段就是將之前所擷取出來的所有低階特徵值進行分群的動作。首先，將圖片資料集中全部圖片的四種低階特徵值分別集合起來，例如將圖片編號 1 至圖片編號 n，共 n 筆的 Color Layout 特徵值歸為一堆。以此類推，總共可得到四大堆的分散資料(低階特徵值)。最後再經由 K-means 分群演算法，分別對四堆的資料進行分群，結果共獲得四堆的已經過分群的資料。

K-means 分群演算法：

1. 從收集的資料中隨機選定 k 個物件，當作個別群聚(cluster)的中心點。
2. 計算一筆的資料與所有群聚中心點的距離。

3. 把每一筆的資料分配到最接近的群聚內。
4. 計算每一個群聚的新中心點。
5. 重複步驟 2，直到物件分配情形恆定不變或已達設定循環次數上限為止。

3.2.3 資料編碼

將圖片資料集中之四種低階特徵值分群的結果分別予以編碼(以下稱為樣式)。以下方的表 5 所示為例，即是將 Color Layout 分為 12 群，予以編號 A0 至 A11；Scalable Color 分為 8 群，給予編號 B0 至 B7；Homogeneous Texture 分為 8 群，給予編號 C0 至 A7、Edge Histogram 則分為 10 群，給予編號 C0 至 C9。以此做為依據，每一張的圖片就可獲得與其四種影像低階特徵值所相對應的四個編碼。在未來，此編碼將成為圖片的特徵樣式，作為圖片之間進行比較的重要依據。

3.2.4 建立特徵樣式資料庫

將每一張圖片的相關資訊皆匯入到資料庫當中，而資訊儲存結構就如圖 2 所示。每一列包含一張圖片的所有資訊。其中，第一欄 index 為圖片在資料庫中的編號，第二欄 name 為圖片的檔案名稱，接下來的四個欄位 layout、scalable、texture、edge 依序存放此張圖片之前述四種特徵值所相對應的特徵樣式。就以表 5 中的第二列為例，此張圖片在資料庫中的編號為 6620、圖片的檔案名稱為 Ship_125.JPG、layout 低階特徵值所相對應的特徵樣式為 A1、scalable 低階特徵值所相對應的特徵樣式為 B4、texture 低階特徵值所相對應的特徵樣式為 C2、edge 低階特徵值所相對應的特徵樣式為 D1。

表 5：特徵樣式資料表範例。

index	name	layout	scalable	texture	edge
6620	Ship_125.JPG	A1	B4	C2	D1
6621	Ship_126.JPG	A7	B4	C7	D7
6622	Ship_127.JPG	A7	B4	C7	D7
6623	Ship_128.JPG	A7	B4	C2	D7
6624	Ship_129.JPG	A7	B4	C6	D1
6625	Ship_130.JPG	A7	B4	C7	D7

3.3 線上圖片搜尋階段

3.3.1 圖片搜尋與排序概念

- A. 像低階特徵值之間的相似程度作為排序的依據，讓最相似的圖片能排在最前面。
- B. 兩張具有一模一樣的樣特徵樣式視為最相似(the most relevant)。
- C. 當具有相同樣視之相似的圖片超過系統預定的張數時，便以低階特徵值之歐幾里德距離(Euclidean Distance)為排序依據。

3.3.2 圖片搜尋流程

輸入：一張圖片

輸出：一至十張的圖片

處理：見圖 4 所示

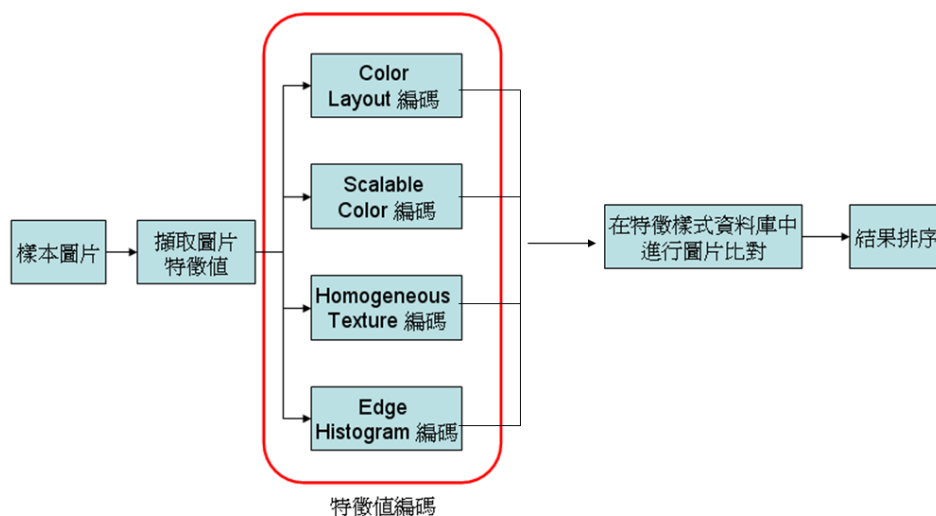


圖 5：圖片搜尋流程。

整個圖片搜尋的流程如圖 5 所示。一開始，將樣本圖片(query image)先進行四種低階特徵值的擷取。擷取完成之後，四組特徵值依序與原有的特徵值資料分群進行比對計算，以獲得四個樣式，接著利用使用者要用到的樣式，即所需要的圖片特徵到特徵樣式資料庫中進行比對以獲得結果。以下，我們將分別敘述”特徵值編碼”，”相似度比對”及”結果排序”逐一說明。

A. 特徵值編碼

針對四種圖片特徵，把樣本圖片(query image)與圖片資料集的特徵值資料分群進行比對。即透過歐幾里德距離公式計算樣本圖片與每一個群聚中心點之間的距離後，再將擁有最短距離的群集之編號視為樣式回傳，最後可得到全部所需要的特徵樣式作為樣本圖片之搜尋依據。

表 6：特徵樣式比對結果範例。

圖片名稱	特徵樣式
Images001.jpg	A1、B2、C3、D4
Images002.jpg	A1、B2、C3、D4
Images003.jpg	D4
Images004.jpg	A1、B2、C3、D4
Images005.jpg	B2、C3
Images006.jpg	A1、B2、D4

B. 相似度比對

接著，我們將比對的詳細說明如下。舉例來說，使用樣本圖片 Images001.jpg，若使

用所得到特徵值樣式 A1、B2、C3、D4 至特徵樣式資料庫中進行比對，結果就會如下頁的表 6 所示。從表中可知道，只有 Images001.jpg、Images002.jpg、Images004.jpg 完全符合四個樣式，所以就得到這三張圖片作為搜尋結果回傳並排序。

C. 結果排序

針對搜尋結果與樣本圖片的相似程度進行排序，目的是讓愈相似的圖片愈能排在搜尋結果的前面。至於這部分的相似度比較，是根據使用者所需要的影像特徵，利用歐幾里德距離(Euclidean Distance)公式對搜尋結果與樣本圖片的影像低階特徵值作相似度運算，所得數值愈小代表相似度愈高，就愈能排在搜尋結果的前面。在進行相似度運算之前，必須先將兩者的影像特徵低階特徵值進行正規化(Normalized)，以平衡使用者不同影像特徵的影響力。

以表 7 為例，若樣本圖片的特徵樣式為 A4、B4、C7、D7，那就可以得到搜尋結果 Ship_126.JPG、Ship_127.JPG、Ship_130.JPG 這三張圖。得到這三張搜尋結果後，分別將他們的四種影像低階特徵值進行正規化，如以表 8 的 Ship_126.JPG 作為例子，即把 Ship_126.JPG 的 Color Layout 特徵值：11、43、21、18、22、7、17、10、16、19、16、14 等十個數值分別除以 63，其餘的 Scalable Color、Homogeneous Texture、Edge Histogram 也依同理進行正規化，之後再利用歐幾里德距離(Euclidean Distance)公式進行運算與樣本圖片間相似度，然後將所得的數值做遞增排序。

表 7：特徵樣式資料庫。

index	name	layout	scalable	texture	edge
6620	Ship_125.JPG	A1	B4	C2	D1
6621	Ship_126.JPG	A7	B4	C7	D7
6622	Ship_127.JPG	A7	B4	C7	D7
6623	Ship_128.JPG	A7	B4	C2	D7
6624	Ship_129.JPG	A7	B4	C6	D1
6625	Ship_130.JPG	A7	B4	C7	D7

表 8：影像的低階特徵值[Color Layout]。

Name of Image	Value
Ship_126.JPG	11 43 21 18 22 7 17 10 16 19 16 14
Ship_127.JPG	14 38 24 22 17 10 15 12 14 25 18 9
Ship_128.JPG	16 45 25 15 25 13 12 18 17 23 16 8

3.3.3 使用者回饋

根據上述，在找到前 10 張最相似之圖片後，接下來要將圖片的自然語意概念自動顯示。其方法為：將前 10 名圖片類別分別於使用者介面的右方框架(frame)下自動打勾。此時使用者可針對自己有興趣的語意類別做勾選或取消的動作，最後再回饋給使用者，以達到更精準的搜尋品質。

4. 實驗測試

4.1 圖片資料搜集

使用自行建立的圖片資料集做為測試資料。其中一部份取自 Corel Gallery 專業圖庫之外，其它則是另外由各處搜集而來。我們從 Corel Gallery 專業圖庫中挑選了大約 4000 張，以人為主觀判斷有能明確分類特性的圖片，而另一個主要的來源，則是從 Google、百度(Baidu)之圖片搜尋引擎與 Yahoo!奇摩拍賣之搜尋引擎所搜尋出來的結果，經過篩選與過濾後，挑出大約 5000 張的圖片。最終，總計搜集了共 8645 張張的圖片。

4.2 實驗規劃

整個實驗是以 Java 在作業系統為 Windows Vista、資料庫為 MySQL、硬體環境為 Pentium-4 2.0 GHz、2.0 GB RAM 的 HP 筆記型電腦作為實驗準確率與效率之平台。而另外此系統已架設在 Pentium-4 3.0 GHz、1.0 GB Ram 的 PC 上。此實驗係取用所有的一般類圖片(8645 張)進行測試。使用的圖片特徵固定為 Color Layout、Scalable Color、Homogeneous Texture、Edge Histogram 這四種，不因圖片不同而做更改。其準確率為： $(\text{答對的圖片張數}/\text{回傳的圖片張數}) * 100\%$ 。

4.3 實驗結果

這個測試是將所有的圖片(8645 張)皆進行搜尋所得到的，整體的準確率之平均約 46.4%，其部份結果數據如表 9 所示。從表 9 中我們可以發現 Bw 此基礎類別圖片的準確率最高。Bw 為 BlackWhite 的縮寫，其中的圖片皆為灰階之人像圖案。Bw 圖片之準確率如此高，主要的原因是顏色比例的懸殊，基本上圖中的顏色不是接近黑色就是接近白色，所以與其它種的圖片在顏色部份差異相當大；至於最低的 Flora 則為花叢的圖片，彼此之間的顏色差異大，也無足以表率的形狀及紋路，所以準確率較低。此外，從表 9 中也可以知道到 Bw 雖然有最高的準確率，但因 Bw 圖片在形狀、紋路方面差異過大，而造成搜尋出來的結果數量相當的低，平均只有 3 張。

表 9：類別準確率。

準確率前 10 高			準確率前 10 低		
基礎類別	準確率	平均搜尋結果張數	基礎類別	準確率	平均搜尋結果張數
Bw	0.733105626	3	Flora	0.2739866	14
Decoy	0.652470042	8	PaintingD	0.2925437	11
DinoArt	0.636407307	13	Lion	0.3037077	9
Soccer	0.63298831	11	Mushroom	0.3272204	8
Tennis	0.631174327	12	Wolf	0.3357424	6
Satelliteimage	0.627832825	8	Penguin	0.3442206	6
Volleyball	0.621111724	9	Fox	0.3560404	7
Billiardsball	0.613642226	15	Leopard	0.3572303	8
F1	0.593653596	5	Toy	0.3581631	8
Bluesky	0.588928752	8	Couple	0.3624945	6

接著我們將此樣式比對方法(以下稱第一種方法)與直接用低階特徵值計算歐幾里德距離(以下稱第二種方法)作比較。使用第一種方法的準確率約為 46.4%(標準差：約 0.102)，優於第二種方法的準確率約為 31.5%(標準差：約 0.192)。此實驗結果顯示，我們提出的方法，的確能較有效的找到相關圖片。此外，從執行效率的角度來分析，此比較也是使用所有的圖片進行測試後再平均，其結果為在回傳 25 張的情況下，第一種方法約花費 0.0191480 秒，遠優於第二種方法之 0.4905379 秒。

4.4 系統介面

圖 6 為系統之示意畫面。當使用者上傳一張圖片後，系統會回傳此查詢圖片之可能之語意概念，並在系統畫面左側之語意概念樹中自動勾選相關的語意。使用者可根據所喜好的語意勾選做回饋，此時，系統會根據回傳的語意做精確的搜尋(如圖 7 所示)，以達到 100%的準確率。

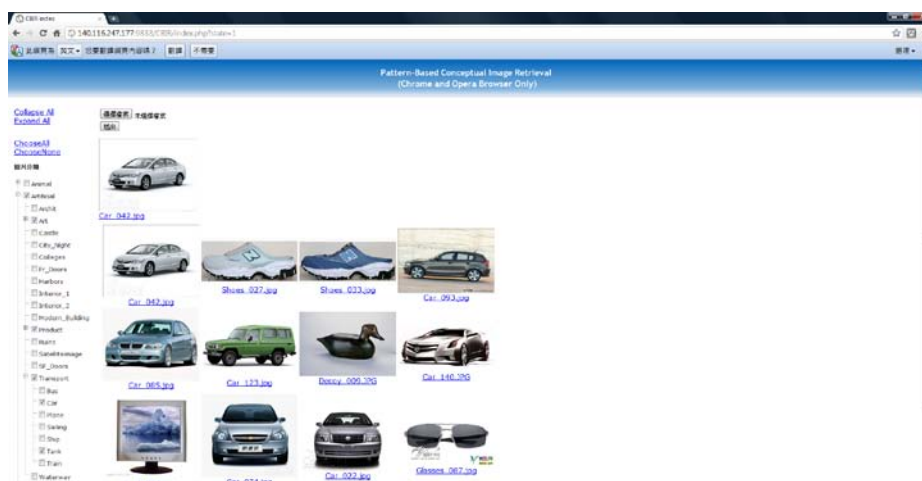


圖 6：以查詢圖片為基礎之系統蒐尋結果畫面。

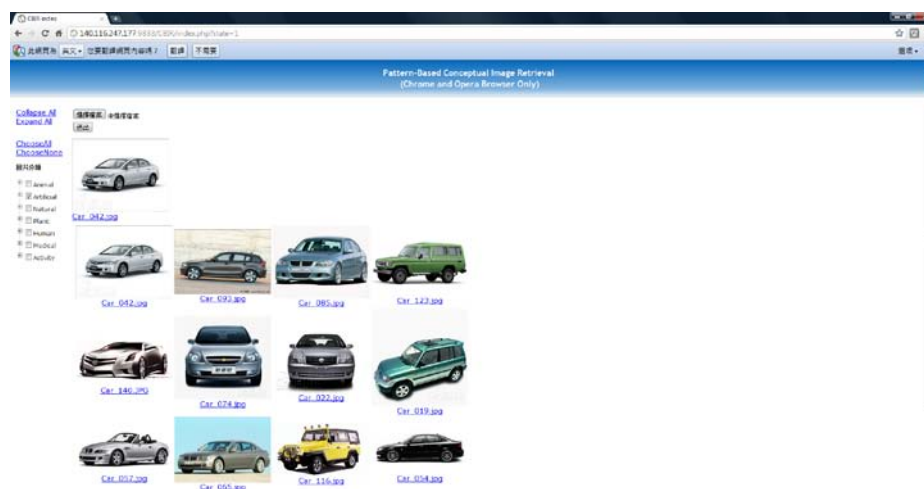


圖 7：以查詢圖片之語意為基礎之系統回饋蒐尋結果畫面。

5. 結論

在本篇論文中，我們提出了一套具高效率之智慧型概念導向式影像擷取系統。此系統雖然仍建立在影像內容為基礎的方法上，卻有別於一般傳統影像內容為基礎的搜尋系統。其目的不但要兼顧一般影像搜尋系統的特色，更要具有人類語意的概念，更有效率。為了達成此目的，我們結合了影像內涵與語意概念，並運用了樣式比對的技術，使得執行速度更快，更能貼近人類自然的查詢方式。

由於本技術現正處於初步完成階段，為使效果更加完美，許多子技術仍有改進空間；此外，我們將以更多類型之實際影像資料進行測試分析，並進一步地推廣到其他相關多媒體領域上，希望在未來，人類視覺與語意之間能更加接近。

誌謝

本研究部份由國科會數位典藏科技計畫所補助，計畫編號 NSC100-2631-H-006-002；部分依經濟部補助財團法人資訊工業策進會「100 年度資訊應用與整合技術開發第二期計畫」辦理。

參考文獻

1. Vision and Media Lab of CS at Simon Fraser University in Canada
<http://www.cs.sfu.ca/research/groups/VML>
2. J.Black J. A., Fahmy G., Panchanathan S., "A Method for Evaluating the Performance of Content-Based Image Retrieval Systems." In *Proc. of the Fifth IEEE Southwest Symposium on Image Analysis and Interpretation*, pp. 96-100, 2002.
3. Djeraba C., "Association and Content-Based Retrieval," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 15, No. 1, 2003.
4. Niblack W., Barber R., Equitz W., et al, "The QBIC project: Querying images by content using color, texture, and shape." In *Proc. of SPIE Electronic Imaging: Science and Technology*, San Jose, CA, 1993.
5. Rui Y., Huang T., and Mehrotra S., "Content-based image retrieval with relevance feedback in MARS." In *Proc. of IEEE International Conference on Image Processing*, pp. 815-818, New York 1997.
6. Su J.H., Chou C.L., Lin C.Y. and Tseng V.S., "Effective Semantic Annotation by Image-to-Concept Distribution Model." *IEEE Transactions on Multimedia (TMM)*, Vol. 13, no. 3, pp. 530-538, June 2011.
7. Tseng V.S., J.H. Su, Huang J.H. and Chen C.J., "Integrated Mining of Visual Features, Speech Features and Frequent Patterns for Semantic Video Annotation." *IEEE Transactions on Multimedia (TMM)*, vol. 10, no. 1, pp. 260-267, February 2008.
8. Tseng V.S., Su J.H., Wang B.W. and Lin Y.M., "Web Image Annotation by Fusing Visual Features and Textual Information." In *Proc. of the 22nd ACM Symposium on Applied*

Computing (SAC), Seoul, Korea, March 11 – 15, 2007.

9. Yin P.Y., Li S.H., “Content-based image retrieval using association rule mining with soft relevance feedback.” *Journal of Visual Communication and Image Representation*, vol.17(5), pp.1108-1125 ,2006.
10. Zhang H., Chen Z., Li M., and Su Z., “Relevance feedback and learning in content-based image search.” *World Wide Web*, Vol. 6, No. 2, pp. 131–155, 2003.