

# 行政院國家科學委員會專題研究計畫成果報告

## 數位圖書館中影像內容擷取的相關回饋

### Relevance Feedback of Content-Based Image Retrieval in Digital Library

計畫編號：NSC 89-2213-E-004-005

執行期限：88年8月1日至89年7月31日

主持人：沈錕坤 政治大學資訊科學系

#### 一、中文摘要

數位圖書館中儲存大量的文字、影像、視訊與音訊資訊，並提供使用者方便且有效的方法擷取資訊。影像擷取是數位圖書館系統設計的主要關鍵性技術之一。影像資訊擷取的主要特性在於相似擷取與內容擷取。

相關回饋是傳統以文字為主的資訊擷取領域所發展的技術。相關回饋利用使用者對先前查詢結果的回饋，自動修正查詢產生新的查詢。因此，修正過的查詢將會更接近使用者所要資訊。

在本計劃中，我們研究影像內容擷取中相關回饋技術。在處理影像的相關回饋時，最直接的方法就是以向量的方式來表示影像的特徵。因此，影像相關回饋的處理類似於以向量空間表示文件的資訊擷取中的處理。但是，並非所有的影像皆以向量的形式表示。因此，我們將分別針對向量形式與非向量形式的影像特徵，發展相關回饋中修正查詢的方法。為了評估所發展的相關回饋技術的效果，我們也發展一全球資訊網上的影像搜尋引擎。

**關鍵詞：**數位圖書館、影像內容擷取、相關回饋

#### Abstract

The digital library stores and manages much of text, image video and audio information. It provides easy and efficient way for users to access multimedia information. Image retrieval is one of the main design issues of digital library. The distinguished features of image retrieval lie in the similarity measure and content-based retrieval.

Relevance feedback is a process of automatically adjusting an existing query using the information fed back by the user about the relevance of previously query result. The adjusted query is expected to be a better approximation to the user's information needed.

In this project, feedback techniques of content-based image retrieval are investigated. The straightforward approach to support relevance feedback of image retrieval is to model the image as the vector of image features. Query refinement is therefore performed similar to the traditional text-based relevance feedback technique based on the vector space model. However, not all image features are represented as vectors. Therefore, we develop the relevance feedback algorithms for image features with and without vector representation.

We also develop a WWW image search engine to measure the effectiveness of the developed technique.

**Keywords:** Digital Library, Content-based Image Retrieval, Relevance Feedback

#### 二、緣由與目的

數位圖書館中儲存大量文字、影像、視訊與音訊資訊，並提供使用者方便且有效的方法擷取資訊。數位圖書館的相關技術包括資料的建立儲存與管理搜尋與擷取資料的傳播資料安全與著作權管理等。

在資料的搜尋與擷取的相關技術中，包括文字、影像、視訊及音訊的擷取。其中，文字擷取的部分，傳統的資訊擷取 (Information Retrieval) 領域的學者已發展出成熟的技術。目前圖書館中的圖書查詢系統以及全球資訊網上的搜尋引擎都是利

用資訊擷取的技術。至於影像、視訊、音訊資料的擷取技術，則是近年來多媒體、資料庫、影像處理領域學者的主要研究課題之一。

影像擷取的研究主要包括下列四項課題：影像特徵的萃取(Extraction of Image Features)，相似度的衡量(Similarity Measure)，增進擷取速度的方法與影像查詢的方式。

在影像特徵的萃取方面，常見的影像特徵包括顏色、紋路、形狀及空間關係。

最常表示顏色特徵的表示法是顏色分佈(Color Histogram)。顏色分佈記錄了影像中每一個顏色的圖素(Pixel)數量。如果共有 M 個顏色，顏色分佈就可表示成 M 維向量。

常見的紋路特徵表示法包括 Tamura 特徵、SAR 模型、MRSAR (MultiResolution Simultaneous AutoRegressive) 模型。Tamura 紋路特徵以明暗對比度(Contrast)、緊密度(Coarseness)、方向(Directionality)表示紋路特徵。MRSAR 紋路特徵源自於 SAR 模型。SAR 模型以代數式(Algebraic Equation)記錄了影像中每個圖素之相鄰圖素(Neighboring Pixels)的資訊。而 MRSAR 模型也是將紋路特徵表示成多維的向量。向量中的每一元素分別代表 MRSAR 模式中代數式的參數。

形狀特徵的表示可由面積、中心點、周長、方位(Orientation)等度量組成。這些表示法不受位移、旋轉的影響、另外，傅立葉描述(Fourier Descriptors)、連鎖碼(Chain Code)也可表示形狀。IBM QBIC 以面積、循環性(Circularity)、離心性(Eccentricity)、主軸方位(Major Axis Orientation)等度量組成的向量表示形狀。其中，循環性定義成(周長)<sup>2</sup>/面積，主軸方位則先由邊線的圖素計算出二階共變矩陣(Second Order Covariance Matrix)，而由中求出最大的特徵向量(Eigenvector)。至於離心性則定義成其最小與最大特徵值的比值。

影像中物體的空間關係是人類最直覺的查詢方式之一。二維字串(2D String)

是最早有關於這方面的研究。簡言之，二維字串以 X 軸方向、Y 軸方向表示此二維表示物體的空間關係。二維字串的優點在於其以簡單的方式記錄空間關係，此外查詢處理更可轉換成二維部份字串比對(2D Subsequence Matching)的處理。

在影像的查詢方式方面，系統必須提供方便的方式供使用者表示視覺查詢。例如以數值表示顏色並不符合人類的直覺。目前可行的方式包括瀏覽(Browsing)、以樣本影響查詢(Query-By-Image-Example)、以影像特徵的描繪查詢(Query-By-Sketch)及相關回饋(Relevance Feedback)。

相關回饋是資訊擷取領域所發展的方法。系統根據使用者對查詢結果的滿意度，自動作查詢修正(Query Refinement)，重新產生新的查詢。目前全球資訊網上的搜尋引擎，如 Alta Vista 都有提供相關回饋的功能。在影像擷取中，使用者可以從查詢結果中，點選相似度高與相似度低的影像，系統根據使用者點選所提中的資訊，自動修正產生新的查詢影像。

由於影像資訊的豐富性以及人類對影像資訊的認知差異，在上述四種查詢方式中，相關回饋以互動的方式，更有助使用者查詢影像。但是，在已發展的影像擷取系統與研究中，只有少數相關回饋的研究。而且，其技術多基值於傳統資訊擷取的方法。因此，本計畫的主要目的在研究有效的方法提供並處理影像內容擷取的相關回饋。本計畫研究的範圍包括相關回饋的方式、相關回饋的處理及其效果評估。

### 三、結果與討論

#### (1) 相關回饋的方式

針對相關回饋的方式，無論是資訊擷取或影像擷取的系統，目前皆由使用者從查詢結果中，以點選相關(相似)與不相關(不相似)的方式回饋給系統。但是，由於影像資訊的豐富性，使用者在做相似判斷，往往面對是與非的抉擇。因此，另一種可行的方式是允許使用者對查詢的結果以排名(Ranking)的方式回饋給系統。

#### (2) 相關回饋的處理

針對相關回饋的處理，目前影像擷取的研究皆利用資訊擷取處理相關回饋的方法。資訊擷取中表示文件特徵的模型包括布林模型(Boolean Model)、向量模型(Vector Model)、機率模型(Probabilistic Model)等。

其中，向量模型以向量的方式表示文件特徵。例如，如果整個數位圖書館的文件經分析後共有  $N$  各關鍵字(keywords)，每份文件的特徵因此表示成一  $N$  維向量。此一  $N$  維向量表示此  $N$  個關鍵字在這份文件中的加權值。最直覺的加權值就是關鍵字出現的頻率。而使用者的查詢也以向量的方式表示，其中加權值代表使用者對此關鍵字的重視程度。查詢  $Q$  與文件  $I$  之間的相似度因此可以定義成兩向量的內積，

$$Sim(I, Q) = \frac{\sum_1^N (I_i \cdot Q_i)}{\sqrt{\sum_1^N I_i^2 \cdot \sum_1^N Q_i^2}}$$

在處理相關回饋時，修正查詢的產生方式常利用如下的 Rocchio 公式，

$$Q_{k+1} = r \cdot Q_k + s \cdot \sum_1^{n1} \frac{R_j}{n1} - \lambda \cdot \sum_1^{n2} \frac{S_j}{n2}$$

其中， $Q_k$  代表第  $k$  次查詢向量， $R_i$  是被使用者點選為相關的文件  $j$  的向量， $S_i$  是被使用者點選為不相關的文件  $j$  的向量， $n1$  是相關文件的數目， $n2$  是不相關文件的數目。、 $r$ 、 $s$ 、 $\lambda$  是調整用的係數。

由於影像的顏色、紋路、形狀特徵皆可以表示成多維向量，因此，目前影像擷取的研究也都以 Rocchio 的方式處理影像擷取的相關回饋。但是，當使用者對查詢的結果以排名的方式回饋給系統時，Rocchio 的方式就無法適用。而且有些影像特徵並非以向量的形式表示。例如空間關係多以二維字串表示，又如形狀以循環碼(chain code)的方式表示。

針對上列問題，我們提出新的相關回饋演算法。此演算法針對使用者的排名回饋，重新產生符合回饋的影像。此演算法分為三大步驟，

**步驟一：**根據影像間的相似度，將數位圖書館中所有影像表示成  $M$  維空間中的

點。我們利用 FastMap[4]的演算法。已知  $N$  個 object 間的相似度，FastMap 將這  $N$  個 object 表示成  $M$  維空間中的點，使得任三個 object  $a$ 、 $b$ 、 $c$ ，如果  $a$  與  $b$  的相似度大於  $a$  與  $c$  的相似度， $M$  維空間中的點  $P_a$  與  $P_b$  間的幾何距離也會小於  $P_a$  與  $P_c$  間的幾何距離。

**步驟二：**根據使用者的排名回饋，找出符合排名的影像的集合。例如當使用者第一次查詢影像後，系統傳出的答案有  $a$ 、 $b$ 、 $c$  三張影像。此三張影像都不是使用者所要的影像，因此使用者以排名的方式提供相關回饋，其相似程度的排名是  $a > b > c$ 。

以圖一為例，其中  $P_a$ 、 $P_b$ 、 $P_c$  分別就是  $a$ 、 $b$ 、 $c$  三張影像經過步驟一處理後表示成二維空間中的點座標。 $L_{ab}$ 、 $L_{bc}$ 、 $L_{ac}$  分別代表點  $P_a$  與  $P_b$ 、點  $P_b$  與  $P_c$ 、點  $P_a$  與  $P_c$  的垂直平分線。 $L_{ab}$  左邊區域中的任何點  $P_x$ ，到  $P_a$  的距離小於到  $P_b$  的距離。同理  $L_{bc}$  左上方區域中的任何點  $P_x$ ，到  $P_b$  的距離小於到  $P_c$  的距離。 $L_{ac}$  左邊區域中的任何點  $P_x$ ，到  $P_a$  的距離小於到  $P_c$  的距離。

因此符合使用者回饋排名的影像，就是落在  $L_{ab}$  左邊、 $L_{bc}$  左上方、 $L_{ac}$  左邊所交集的區域，也就是圖一中的灰色區域。

在實際處理的時候，我們並不需要找出區域出來。我們只要對回饋排名中每一對點，利用向量內積的方式就可以判斷。也就是說，對於任何兩點  $a$ 、 $b$ ，如果使用者回饋的相似度排名是  $a > b$ ，符合這條件的點  $P_x$ ，其條件為

$$(P_x - \frac{P_a - P_b}{2})(P_a - \frac{P_a - P_b}{2}) > 0$$

以圖一為例，落在灰色區域中的點  $P_x$  符合

$$(P_x - \frac{P_a - P_b}{2})(P_a - \frac{P_a - P_b}{2}) > 0$$

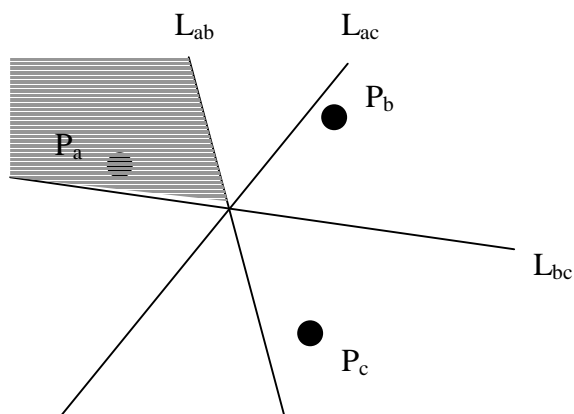
$$(P_x - \frac{P_b - P_c}{2})(P_b - \frac{P_b - P_c}{2}) > 0$$

$$(P_x - \frac{P_a - P_c}{2})(P_a - \frac{P_a - P_c}{2}) > 0$$

**步驟三：**將這些符合排名條件的影像，依照與原來查詢影像的相似度依序列出。

以上三個步驟中，步驟一是 preprocessing 的部份。步驟二與三則是在

run time 時處理使用者的相關回饋。因此，如果數位圖書館中有  $N$  張影像，而查詢系統每次傳出  $R$  張影像，共需要  $C_2^R * N$  的 computation time，其中  $C$  代表組合函數。



圖一：二維空間上三個影像。

我們發展了一個全球資訊網上的多媒體搜尋引擎。利用這搜尋引擎，我們評估相關回饋的效果。我們共收集了 3219 張影像，我們提供使用者以顏色查詢影像，雖然顏色多以向量的方式表示，但是我們提供使用者以排名的方式做相關回饋。顏色特徵以 HSV model, 128 色。系統每次傳出 10 張影像。我們請測試者查詢並依系統傳出結果做相關回饋。我們以使用者查詢到所要的影像所經過的回饋次數來衡量回饋機制的效果。由於步驟一的 FastMap 的維數會影響影像間相似度關係的精確度，我們分別測試了不同的維度。結果如下表所示。

表一：相關回饋效果評估。

M	10	20	30	40	50
平均次數	2.27	2.18	2.20	1.97	1.65

#### 四、計畫結果自評

本計畫成果提供新的相關回饋方式幫助使用者查詢影像。其應用範圍除了影像之外，本計畫所提出的相關回饋機制也可應用在視訊、音訊等有時間先後關係的資料。本計畫的實作搜尋引擎部份曾獲得國科會八十九年度大專生專題研究計畫大專生研究創作獎。

#### 參考文獻

- [1] C. Buckley, G. Salton and J. Allan, The Effect of Adding Relevance Information in a Relevance Feedback Environment, in Proceedings of ACM SIGIR'94 International Conference on Research and Develop in Information Retrieval, pp. 292-300, 1994.
- [2] S. F. Chang, A. Eleftheriadis and R. McClintock, Next Generation Content Representation, Creation and Searching for New Media Applications in Education, Proceedings of IEEE, Special Issue on Multimedia Signal Processing, Vol. 86, No. 6, pp.884-890, 1998.
- [3] M. D. Dunlop, The Effect of Accessing Nonmatching Documents on Relevance Feedback, ACM Transactions on Information Systems, Vol. 15, No. 2, pp. 137-153, 1997.
- [4] C. Faloutsos and K. I. Lin, FastMap: A fast Algorithm for Indexing, Data Mining and Visualization of Traditional and Multimedia Datasets, in Proceedings of ACM SIGMOD'95 International Conference on Management of Data, pp. 163-174, 1995.
- [5] A. Gupta and R. Jain, Visual Information Retrieval, Communications of ACM, Vol. 40, No. 5, pp. 71-79, 1997.
- [6] P. Mitra and B. B. Chaudhuri, Efficiently Computing the Closest Point to a Query Line, Pattern Recognition Letters, Vol. 19, No. 9, pp. 1027-1035, 1998.
- [7] J. J. Rocchio, Relevance Feedback in Information Retrieval, in G. Salton ed. The Smart Retrieval System-Experiments in Automatic Document Processing, Prentice Hall, 1971.
- [8] Y. Rui, T. S. Huang and S. Mehrotra, Content-based Image Retrieval with Relevance Feedback in Mars, in Proceedings of IEEE ICIP'97 International Conference on Image Processing, 1997.
- [9] Y. Rui, T. S. Huang and S. Mehrotra, Human Perception Subjectivity and Relevance Feedback in Multimedia Information Retrieval, in Proceedings of IS&T/SPIE Storage and Retrieval for Image and Video Databases VI, 1998.
- [10] Y. Rui, T. S. Huang, M. Ortega and S. Mehrotra, relevance feedback: A Power Tool for Interactive Content-based Image Retrieval, IEEE Transactions on Circuit and Systems for Video Technology, Vol.8, no. 5, pp. 644-655, 1998.
- [11] M. E. J. Wood, N. W. Campbell and B. T. Thomas, Interactive Refinement by Relevance Feedback in Content-based Digital Image Retrieval, in Electronic Proceedings of ACM Multimedia'98, Bristol, UK, 1998.