

行政院國家科學委員會專題研究計畫成果報告

適用於蝴蝶的內容與影像查詢技術之研究(1/2)

Content and image based retrieval on Butterflies

計畫編號：NSC 90-2213-E-002 -141-

執行期限：90年08月01日至91年07月31日

主持人：項潔教授 國立台灣大學資訊工程學系

共同主持人：洪政欣副教授 國立暨南國際大學資訊工程學系

計畫參與人員：陳鍾誠，劉文俊博士生 國立台灣大學資訊工程學系

一、中文摘要

本年度的成果主要是發展了一個適用於蝴蝶的影像擷取模型，這個模型清楚地描述了蝴蝶影像的查詢問題。我們也設計了一些實驗，並從實驗結果中歸納出可做為自動化資料建構的方式。

關鍵詞：蝴蝶、外觀查詢、影像檢索、影像擷取

Abstract

In this year, we developed a methodology for fine-grained image retrieval, and suggest a way for automatic construction of needed data from experimental findings. Furthermore, we also discussed the interaction between user behavior and this methodology.

Keywords: Butterfly, image indexing, image retrieval

二、緣由與目的

在第一年(89年)的計畫成果中，我們設計了一個視覺化的資料建構工具來輔助資料建構員。因為人的參與，這個資料建構工具可以有效地建構出符合人視覺感受的結果。比起完全的人工建構而言，這個工具的確提供適當的幫助，減少了建構員的工作。但是，我們仍然覺得這個方式耗費許多人力，因此需要更進一步的改善。並且，在實作之外，我們認為應該把蝴蝶影像的查詢問題抽象化成為一個模型。當應用到其他類似的問題時，才能方便地對相關問題進行討論。

在第二年的研究中，我們先設計了一個模型來方便討論蝴蝶影像的查詢問題。這個模型是針對蝴蝶影像的特性而設計出來的，因此有別於一般用途的影像查詢系統。在一般用途的影像查詢系統裡，各種類型的影像都可能出現在資料庫中，因此這類的影像資料庫包含的有用的資訊可能非常豐富，也可能非常分歧。這些一般用途的影像查詢系統的目的在於分辨出影像間的粗略(coarse-grained)差異，當要處理像蝴蝶資料庫這類的影像時，一般用途的查詢系統便無法勝任。因為對同一類的影像而言，使用者關心的是影像間的

細微差異。

相對來說，我們的問題專注在查詢同一個領域的影像資料庫(例如：蝴蝶影像)。除了辨別影像間的細微差異外，可能遭遇的挑戰還有使用者對同一張影像的描述不一致，以及在一開始查詢時，使用者較難用文字描述一隻蝴蝶的問題。我們稱這類的問題為 fine-grained image retrieval。在這類資料庫的情形下，影像所傳達出的資訊是比較可以掌握的，使得我們可以分析影像上的資訊，並且整理分類而得到結構化的資訊。這些結構化的資訊便可以有效地幫助查詢系統的設計。

在我們的方法論裡，我們首先提出了 dimension 的觀念。蝴蝶影像上的資訊可以簡單地被分為 color, shape, pattern 三個 dimensions。Dimension 的觀念對影像資訊做了一個簡單的分類，而這個分類對於目標使用者(target user)來說是直覺而容易接受的。接著對於每個 dimension，可以分析出適當數量的 features。例如我們對蝴蝶影像資料庫做分析，而得到 18 個 color features, 7 個 shape features, 以及 17 個 pattern features。

Dimensions 與 features 所形成的結構化的資訊，不僅可以用來設計一個直覺而有效的互動式影像查詢系統，也可以用來幫助資料建構的自動化。下個章節會陸續介紹這兩個部份的研究成果。

三、研究成果

在第一個部份的研究成果中，會介紹我們的方法論，並且同時呈現蝴蝶查詢介面作為實例說明。而在第二個部份的成果裡，會介紹自動化資料建構的做法以及實驗結果，並且歸納出一個可行的方向。

我們提出的方法論分為三個部份：資料模型，查詢模型，以及互動模型。資料模型描述了各種在蝴蝶影像查詢上所必需的資料。查詢模型描述了我們提出的互動式查詢系統。而互動模型則是描述使用者與系統之間的互動過程。

(1) 資料模型

對蝴蝶影像資料庫而言，可以先分析出三個 dimensions 以及所對應的 features。

• 蝴蝶影像集合: A

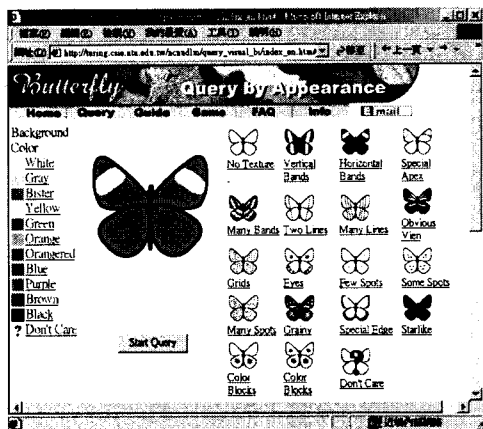
- dimensions 所成集合: $D = \{c, s, p\}$
(也就是 color, shape, pattern)
- features 所成集合: $F = F_c \cup F_s \cup F_p$,
其中 F_d 是 dimension d 的 features 所成集合。

另外，我們提出了兩個觀念：proximity 與 similarity。Proximity 是用來測量一張影像與一個 feature 之間的關係。而 similarity 是用來測量兩張影像之間的關係。這兩個觀念可以經由以下兩個函數得到。

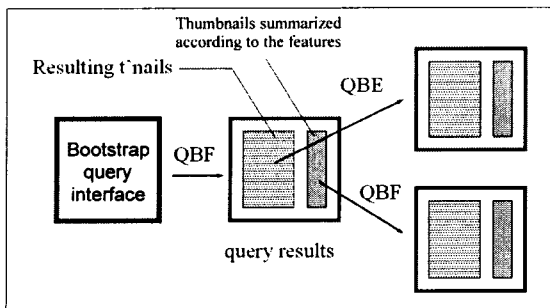
- $P(a, f) \geq 0$, where $a \in A$ and $f \in F$
- $S(a, b) \geq 0$, where $a, b \in A$

(2) 查詢模型

影像查詢系統的流程如下：(1)一開始，查詢系統的畫面包含了各個 dimensions 所對應的 features 的選單，使用者可以選擇不同的 features 來下查詢，處理系統則根據 proximity 回傳符合的影像，並且根據 features 對查詢結果做後置分類，這個查詢方式稱為 query-by-features (QBF)；(2)使用者可以在 QBF 的結果裡選擇影像來下查詢，處理系統則根據 similarity 回傳相似的影像，同時也呈現所選影像符合的 features，這個查詢方式稱為 query-by-example (QBE)。在 QBF 與 QBE 中，後置分類或後置分析的結果，便形成了使用者及系統互動的機制。如果把 q 定義成一個 QBF 或 QBE 的查詢，則查詢函數 $QBF(q, A)$ 及 $QBE(q, A)$ 會回傳相對的結果給使用者。圖一是我們所設計的查詢模型一開始的直覺式查詢介面，圖二則是查詢模型的概念圖。



圖一：直覺式查詢介面



圖二：查詢系統概念圖

(3) 互動模型

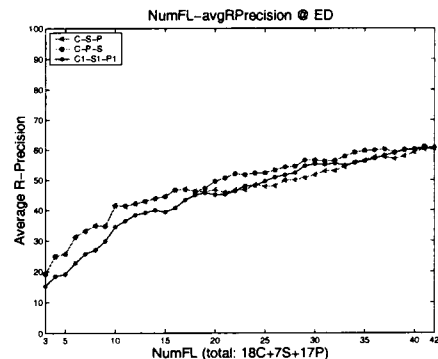
互動模型是討論使用者與查詢系統之間互動的過程。模型如下：

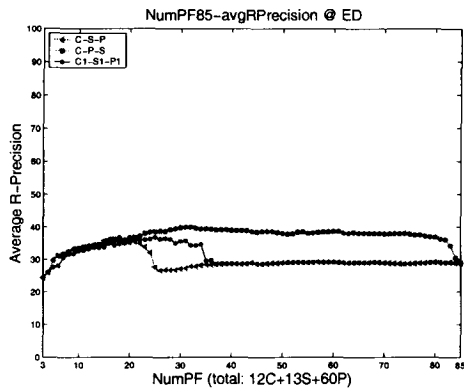
- $\Delta_1 = (\Phi_1, \Psi_1)$, result of the first iteration, where
 $\Phi_1 = QBF(q, A)$ and $\Psi_1 = F(q, \Phi_1)$
- $\Delta_n = (q, \Phi_n, \Psi_n)$, result of the n th iteration
 - If q a QBF query, $\Phi_n = QBF(q, \Omega)$ and
 $\Psi_n = F(q, \Phi_n)$, where $\Omega \in \{A, \Phi_{n-1}\}$
 - If q is a QBE query, $\Phi_n = QBE(q, A)$ and
 $\Psi_n = G(q)$

其中， Ψ_1 及 Ψ_n 是一些 interactive query options。

從 data model 中可以了解，dimensions、features、proximity、similarity 這四種資料建構得好壞會非常影響查詢系統的有效性 (effectiveness)。在影像擷取的系統中，如何萃取影像上的重要資訊，是一個困難的問題。Dimension 的觀念可以將這個問題切割成為較小的問題。當系統設計者訂定好 dimensions 之後，對所有的影像在某個 dimension 上的資訊 (例如蝴蝶的顏色資訊)，可以嘗試用 clustering 的技術來做叢聚分析。並且提供一個方便的工具讓使用者選擇出適當數量的 features。對 proximity 及 similarity 而言，我們希望可以盡量自動地建置出符合人視覺的結果。以下便簡要地描述我們的實驗以及實驗結果。

另外，image similarity 也是一個困難的問題。在我們所設計的實驗裡，實驗結果發現，用我們所定義的 features 以及影像處理技術常用的 primitive features 來建構 similarity，前者的準確率比後者好許多。因此可以使用 features 來自動計算出影像間的 similarity。我們所用的是 MPEG-7 裡所建議的一些影像描述子 (descriptor)。這些技術是用來記錄影像上不同 dimension 的資訊。例如我們採用了 color histogram 及 moments 來記錄 color 的資訊，CSS 及 ART 來記錄 shape 的資訊，以及 Gabor wavelet 來記錄 texture 的資訊。圖三是兩個實驗的結果，曲線越往上則越好。可以觀察得到，FL (使用我們所定義的 features) 比 PFR (primitive-features with real values) 的結果好許多。





圖三：features vs. primitive-feature with Euclidean distance

我們接下來的研究方向是，採用 clustering 技術，發展一個半自動化的 feature extraction 工具，讓使用者方便地建構出 features。另外，如果將我們的 fine-grained 影像擷取的方法論，資訊萃取工具 (feature extraction)，以及 proximity 和 similarity 的建構工具，結合為一個整合性的發展環境時，就可以很方便地應用到另一個特定領域的影像資料庫 (例如：樹葉，花，人臉，汽車等)，而得到一個 user-friendly, effective 的視覺化查詢系統。

四、參考文獻

[1] MPEG-7 Committee Draft – Visual Part, *ISO/IEC JTC1/SC29/WG11/N3701*, LaBaule, October 2000.

[2] Farzin. M and Sadegh A. et al., Robust and Efficient Shape Indexing through Curvature Scale Space, *British Machine Vision Conference*, 1996.

[3] Farzin M., Sadegh A., and Josef K., Efficient and Robust Retrieval by Shape Content through Curvature Scale Space, *Int. Workshop on Image Databases and Multimedia Search*, 1996

[4] Rafael C. Gonzalez and Richard E. Woods, Digital Image Processing, *Addison-Wesley Publishing Company*, Reprinted in September, 1993.

[5] D. Patel and T.J. Stonham, Low level image segmentation via texture segmentation, *Proc. SPIE Visual Comms. and Image Processing*, vol 1606, p621, 1991.

[6] Y. Ishikawa, R. Subramanya, and C. Faloutsos, "Mindreader: Query databases through multipole

examples," in *Proc. Of the 24th VLDB Conference*, (New York), 1998.

[7] L. Hepplewhite and T.J. Stonham and R.J. Glover, Automated visual inspection of magnetic disk media, *Proc. Of 3rd ICECS*, vol2, p732-735, 1996

[8] Yong Rui, Thomas S. Huang, and Sharad Mehrotra, "Content-based image retrieval with relevance feedback in MARS", *IEEE Int. Conf. on Image Processing*, 1997.

[9] A. K. Jain and F. Farrokhnia, Unsupervised texture segmentation using Gabor filters, *Pattern Recognition*, 24(12):1167-1186, 1991.

[10] H. Tamura and S. Mori and Y. Yamawaki, Textural features corresponding to visual perception, *IEEE Trans. On SMC*, vol 6 no6, p460-473, 1978

[11] Hongjiang Zhang, Liu Wenyin, Chunhui Hu, "iFind—A System for Semantics and Feature Based Image Retrieval over Internet", *ACM MULTIMEDIA 2000--The 8th ACM International Multimedia Conference*, Los Angeles, California, October 30 - November 3, 2000