

行政院國家科學委員會補助專題研究計畫成果報告

類神經網路影像檢索之研究(IV):
網格計算影像與視訊資料檢索系統之建構

計畫類別：個別型計畫

計畫編號：NSC 96-2221-E-009-167

執行期間：2007/08/01 ~ 2008/07/31

計畫主持人：傅心家

共同主持人：包曉天

成果報告類型(依經費核定清單規定繳交)：精簡報告

本成果報告包括以下應繳交之附件：

出席國際學術會議心得報告及發表之論文各一份

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、
列管計畫及下列情形者外，得立即公開查詢

執行單位：交通大學資訊工程系

中 華 民 國 97 年 10 月 23 日

中文摘要

隨著數位影像技術的提升、大容量儲存媒介價格的下降與網際網路的普及。如何有效且快速準確地從網路獲得想要的資訊，已是網路搜尋服務亟欲解決的問題。本實驗室首先針對 Corel 影像資料庫，研發建立了以 Visual keyword 作為影像索引的影像搜尋引擎。隨後，本實驗室將此技術拓展於網際網路上的影像搜尋上。本年度將延伸前三年所完成之影像搜尋引擎的技術，研究開發出以視訊內容為基礎的網際網路視訊搜尋引擎，完成一個能在網際網路上搜尋影像與視訊的系統。基於網格(Grid)的資源分享與網格的無縫(seamless)存取服務的特性，預期將可有效的減少重複的運算負載並動態平衡網格的資源，進而提升系統的搜尋速度，以達到在可接受的時間內快速找到使用者所要找尋的影像或視訊。本計畫建構一個包括 20 台電腦以網格為基礎的分散式影像與視訊搜尋引擎，初步的實驗結果顯示在上萬張影像中其搜尋命中率可達六成以上，每次搜尋約需時 1~5 秒。

關鍵詞：類神經網路，分散式影像資料檢索，影像分群，分散式視訊資料檢索，網格計算

Abstract.

The growth of the Internet and services has caused a corresponding explosion in the amount of media data that needs to be archived. The most important issue of the web-based search service is how to retrieve the desired data from Web efficiently and correctly. In the early projects, we proposed a novel image index called Visual keyword and built an image search engine to retrieve the Corel gallery images. Then, we built a web-based image search engine. In this project, we plan to exploit a novel video index based on Visual keyword and develop a content-based video search engine on Web. Based on the concepts of resource sharing and seamless accessing of the Grid computing, a web-based distributed image and video search engine will be developed on Grid to decrease duplicated computing, balance network resource, and speed up data retrieval time. This project develops a Grid based distributed system with 20 servers. The experiment results show that the hit rate of the retrieved images is about 60%, and the response time is about 1~5 seconds.

Keywords: Neural Network, distributed image retrieval, image clustering, distributed video retrieval, Grid computing

一、前言

由於數位影像技術的提升與大容量儲存媒介價格的下降，大量的影像及視訊資料的儲存愈來愈普遍。面對這些愈來愈大的影像或視訊資料，如何有效且快速地在搜尋所需的影像或視訊已是急欲解決的課題。此外，隨著網路品質的提升與網際網路的普及，有愈來愈多的數位資料，如文字、影像、視訊等被存放於網際網路上。如何有效且快速準確地從網路獲得想要的資訊，已是網路搜尋服務亟欲解決的問題。目前已有許多網路搜尋引擎以文字做為關鍵字來幫助使用者於網際網路上搜尋其所要的資料。若所要搜尋的為影像或視訊資料時，由於目前並沒有很成熟完美的自動影像或視訊文字索引技術，以文字做為關鍵字的搜尋技術可能無法精確地找到使用者所期望的影像或視訊資料。因此，目前有越來越多關於這方面的研究出現，如以影像內容為基礎的搜尋方式(Content-Based Image Retrieval)研究 [1-9] 與以視訊內容為基礎的搜尋方式(Content-Based Video Retrieval)的研究 [12-19]。

目前雖已有許多以文字為基礎的影像及/或視訊搜尋引擎[10][11]。然而，用影像或視訊內容為基礎的大量資料的搜尋引擎系統則還未看到。因此，本實驗室首先針對 Corel 影像資料庫[21]，於前年研發建立了以 Visual keyword 作為影像索引的影像搜尋引擎。隨後，本實驗室於是將此技術拓展於網際網路上的影像搜尋上。

為了驗證所提之索引技術在大型資料庫之運作情形，本年度擬將延伸前三年所完成之影像搜尋引擎的技術，研究開發出以視訊內容為基礎的網際網路視訊搜尋引擎，完成一個能搜尋網際網路上影像與視訊的搜尋系統。為了讓系統能在運作時能快速處理使用者來自各方的詢問，同時加快電腦的搜尋處理速度，本計畫準備建構一個以網格(Grid)為基礎的分散式影像與視訊搜尋引擎。

二、研究目的

由於分享於網際網路上的影像數量相當龐大，本實驗室將研究開發以網格運算(Grid computing)為基礎來設計影像的搜尋系統。由於本系統是一個包含了多個伺服器的網際網路上影像搜尋引擎，如何能在可接受的時間內快速找到使用者所要找尋的影像是研發此系統的最重要關鍵技術。為了讓系統能在運作時能快速處理使用者來自各方的詢問，同時加快電腦的搜尋處理速度，本計畫準備建構一個以網格(Grid)為基礎的分散式影像搜尋引擎。期望使用者詢問影像時，使用者可以不用費心 Grid 內部如何運做，Grid 會於內部動態的挑選適宜的電腦作運算以達到有效率的 Grid 的資源分配。

三、研究方法

本計畫以網格運算(Grid computing)為基礎來設計影像的搜尋系統，並驗證所設計的影像蒐尋方法在網格運算下的運作情形。我們所要完成的搜尋系統包含了七種伺服器，茲說明如后：

1. **Web Server** 提供使用者透過瀏覽器下達影像訊搜尋的介面。
2. **Web Spider** 定期自動地到網際網路上搜集影像，並將發現的新影像傳送到 Media Indexing Server 來建立可供搜尋的索引。
3. **Media Indexing Server** 抽取由 Web spider 所傳送過來的影像索引，並根據影像的特性將此索引存入適當的 Metadata Server 中。由於目前已有許多網站允許使用者上傳分享視訊片段，為了找尋所想要的視訊片段，本年度將計畫提出了以 Visual keyword 為基礎的視訊索引。我們預定用 Visual keyword 來描述視訊片段上的動態特徵，並以此建構動態視訊索引。相較於在視訊片段中

的每一個畫面取 Visual keyword，所提的方法將只對影像中的關鍵畫面取 Visual keyword 作關鍵畫面索引。當搜尋視訊時，根據使用者所選的 Visual keyword，系統會找尋與其相似的關鍵畫面。然後系統會以此關鍵畫面中對應的 Visual keyword 往前或往後追蹤此 Visual keyword，直到於畫面中的 Visual keyword 消失為止。如此我們即可得到視訊中包含所要詢問的 Visual keyword 的片段。

4. **Metadata Server:** 存放影像與視訊的索引。由於網路影像資料非常龐大，為了要加快影像的搜尋速度，我們利用分群(clustering)的概念將影像分成多個叢集(cluster)。每部 Metadata Server 將針對某個特定語意的叢集來做儲存。這些儲存的索引將提供 Metadata Server 作運算使用。此部分將根據影像的 Visual keyword 將影像做分群。
5. **Matching Server:** 根據使用者所提供的影像範例的特性與當時網格的運作情形，動態的選擇網格上適宜的 Metadata Server，計算其儲存之影像與視訊與使用者所提供的影像範例的相似程度。在建構 Matching Server 時所應解決的問題是如何在網格中，選擇適宜的 Metadata Server 來作相似度運算。於是 Matching Server 只須詢問 Metadata Server 即可知是否需要比較此 Metadata Server 上的影像或視訊的索引。
6. **Dispatch and Aggregate Server:** 分析使用者所提供的影像範例(sample)特徵，將其影像與視訊的詢問要求分派到網格上適合的 Matching Server 上執行。當 Matching Server 完成相似度計算後，則整合其不同 Matching

Server 上的計算結果。我們預計將根據前述 Metadata Server 之分群概念，把使用者所提供的影像範例分到有相同的叢集做運算。在設計 Dispatch and Aggregate Server 時所要考慮的問題是如何有效整合這些伺服器的輸出結果。

7. **Proxy Server:** 存放已經有計算過相似度的影像與視訊資料。由於在 Grid 上同屬於一個虛擬組織的使用者其所要尋找的影像或視訊是相似的。因此，當使用者在提供的影像範例作詢問時，有可能不久前已有另一使用者已經詢問過相同或相似的影像範例。此時，系統將無須重新計算影像或視訊相似度，直接將儲存於 Proxy Server 中的結果回傳給使用者即可。如此可以加快系統處理速度。如何決定哪些使用者是屬於同一個虛擬組織是在設計 Proxy Server 時所要考慮的問題。我們預計將根據前述 Metadata Server 之分群概念，將使用者所提供之影像範例作分群，屬於同一群之使用者即可認定為同一虛擬組織。

四、結果與討論

我們採用 Globus® Toolkit [20]來當做我們開發系統的基礎。為了完成此系統，我們在網格上開發一能提高檔案傳輸速率與可靠度的檔案傳輸服務來完成網格上繁複的檔案傳輸。同時，為了減少系統中 Matching Server 到 Metadata Server 的流量，我們建立一複製品位置服務。此服務會將 Matching Server 已經處理過某個 Metadata Server 中的影像或視訊的索引儲存一份備份於 Matching Server 上。因此，同個影像或視訊的索引將會分散儲存於網格上。我們也建立快速搜尋法則，根據 Matching Server 於網格的位置與網格

的運作狀況，快速的找尋其所要找的影像或視訊的索引。同時，我們也建立監視與探索服務來監看網際網路影像與視訊搜尋系統上的狀況，動態調整參與影像或視訊搜尋計算的伺服器與資源，進而達成網格上的動態平衡。我們建立了一個由 20 台電腦組成以網格運算的網路影像搜尋系統 (<http://140.113.216.66/WebImageSearch>)

。系統包括：

1. Web Server-- 建立一個能提供使用者透過瀏覽器下達影像與視訊搜尋的友善介面。

2. Media Indexing Server-- 為網際網路上的影像與視訊建立索引。

3. Metadata Server-- 用於存放影像與視訊的索引。

4. Matching Server --能根據使用者所提供的影像與範例的特性，並審酌網格當時的運作情形，動態選擇網格上適宜的 Metadata Server 之影像與視訊來作相似程度比對。

5. Dispatch and Aggregate Server --根據前述 Metadata Server 之分群概念，分析使用者所提供的影像與視訊範例的特徵，將其影像與視訊的詢問要求分派到網格上適合的 Matching Server 上執行。並於完成相似度計算後，整合其不同 Matching Server 上的計算結果。

6. Proxy Server 存放已經有計算過相似度的影像與視訊資料。

五、計畫成果自評

本計畫建構一個包括了 20 台電腦以網格(Grid)為基礎的分散式影像搜尋引擎。根據網格資源分享無縫(seamless)存取服務的特性，本系統可有效的減少重複的運算負載並動態平衡網格的資源，進而提升系統的搜尋速度。初步的實驗結果顯示在上萬張影像中其搜尋命中率可達六成以上，每次搜尋約需時 1~5 秒。本研究成果可推廣到其他多媒體資料及系統上，除可大幅降低在多媒體上以人為標記所耗費的大量金錢與人力外，更可助益國家在數

位化所欲促進之教育、學術研究，以及民生育樂上的利用。

參考文獻

- [1] W. Niblack, R. Barber, and et al. "The QBIC project: Query images by content using color, texture and shape," *Proc. SPIE Conf. On Vis. Commun. And Image Proc.*, Feb 1994.
- [2] J. R. Bach, C. Fuller, A. Gupta. A. Hampapur, B. Horowitz, R. Humphrey, R. Jain, and C. fe Shu, "The Virage image search engine: An open framework for image management," *Proc. SPIE Conf. On Vis. Commun. And Image Proc.*, Feb 1994
- [3] A. Pentland, R. Picard, and S. Sclaroff, "Photobook: Content-based manipulation of image databases," *Storage and Retrieval for Image, Video Databases II*, San Jose, CA, Feb. 1994, SPIE, vol. 2185.
- [4] John R. Smith and Shih-Fu Chang, "Visualeek: A fully automated content-based image query system," *ACM Multimedia*, 1996, pp. 87-98.
- [5] W. Y. Ma and B. S. Manjunath, "Netra: A toolbox for navigating large image databases," *Proc. IEEE Int'l Conf. Image Processing*, 1997, pp. 568-571.
- [6] T. S. Huang, S. Mehrotra, and K. Ramchandran, "Multimedia analysis and retrieval system (MARS) project," *Proc. Of 33rd Annual Clinic on Library Application of Data Processing- Digital Image Access and Retrieval*, 1996.
- [7] C. Frankel, M. J. Swain, and V. Athitsos, "Webseer: An image search engine for the world wide web," *Tech. Rep. TR-96-14*, Computer Science Department, University of Chicago, 1996.
- [8] F. Rabitti, P. Stanchev, "GRIM_DBMS: a graphical image

- database mangement system”, *Visual Database System*, Elsevier Science Publishers, B. V., IFIP, 1989, P415-430.
- [9] Chad Carson, Megan Thomas, Serge Belongie, Joseph M. Hellerstein, and Jitendra Malik, “Blobworld: A system for region-based image indexing and retrieval,” *Third International Conference on Visual Information Systems*, June 1999, pp. 509–516
 - [10] Google, <http://www.google.com>
 - [11] YaHoo, <http://tw.yahoo.com>
 - [12] M. Flinker, H. Samhey, W. Niblack et al., “Query by Image and Video Content: The QBIC System”, *IEEE Computer*, 28, (Sept. 1995), pp. 23-32.
 - [13] A. Hampapur, A. Gupta, B. Horowitz, C-F. Shu, C. Fuller, J. Bach, M. Gorkani, and R. Jain, “Virage Video Engine”, *SPIE* Vol. 3022, 1997.
 - [14] D. Ponceleon, S. Srinivasan, A. Amir, D. Petkovic, and D. Diklic, “Key to Effective Video Retrieval: Effevtive Cataloging and Browsing”, *ACM Multimedia*, '98, pp. 99-107.
 - [15] S-F. Chang, W. Chen, H. Meng, H. Sundaram, and D. Zhong, “A Fully Automated Content Based Video Search Engine Supporting Spatio-Temporal Queries”, *IEEE Transaction on Circuits and Systems for Video Tecnology*, Vol. 8, No. 5, Sept., 1998.
 - [16] S. Adali, K. S. Candan, S-S. Chen, K. Erol, and V. S. Subrahmanian, “Advanced Video Information System: Data Structure and Query Processing”, *Multimedia System* Vol. 4, No. 4, Aug. 1996, pp. 172-86.
 - [17] C. Decleir, M-S. Hacid, and J. Kouloumdjian, “A Database Approach for Modelling and Querying Video data”, *LTCs-Report* 99-03, 1999.
 - [18] H. Jiang and A. Elmagarmid, “Spatial and temporal content-based access to hypervideo databases” *VLDB Journal*, 1998, No. 7, pp. 226-238.
 - [19] J. Z. Li, M. T. Ozsu, and D. Szafron, “Modeling of Video Spatial Relationships in an Object Database Management System”, *Proc. of Int. Workshop on Multi-media Database Management Systems*, 1996, pp. 124-132.
 - [20] The Globus Alliance, <http://www.globus.org> .
 - [21] Corel Gallery 1000000, <http://www.corel.com> .