

# An Image-Text Retrieving Intelligent Agent For World Wide Web

陳孟君、陳鴻文

E-mail: 8804781@mail.dyu.edu.tw

## ABSTRACT

Due to the rapid growth of applications in the Internet and WWW, graphic and image files are widely used in multimedia webpages. Thus, the way of effectively managing and retrieving image files has become an important issue today. Although there are a few multimedia search engines available, e.g., Yahoo's image search engine, Webseek, ImageRover, they cannot completely match users' demands. That is, a manual update of webpages usually makes so small amount of images accessible in the search database. Thus, a quick way of retrieving relevant image files is pursued and reported in the thesis. Utilizing our constructed topic knowledge database, the search keyword typed by user will be automatically transformed to another group of keywords. Then, topic related webpages will be reported through full-text search engines and collected by our webpages retrieval module efficiently. The features of HTML, including title, keywords within the document and image filenames, are first analyzed by a quick character string match. Following that, a fuzzy inference module, which uses those HTML features as inputs, can measure the degree of topic relevance for all image files contained in the collected webpages. Experiments show that the novel image searching technique can illustrate a promising performance in searching time and correctness in retrieving automobile webpages and image files on Internet. Keywords: Information Retrieval Internet Intelligent Agent

Keywords : Information Retrieval ; Internet ; Intelligent Agent

## Table of Contents

第一章 緒論.....	1	第一節 研究背景.....	1	第二節 研究動機.....	3	第三節 研究目的.....	6	第四節 研究範圍.....	6	第五節 研究流程.....	7
第二章 文獻與基本理論探討.....	9	第一節 網路資訊檢索.....	9	第二節 智慧型代理器.....	18	第三節 模糊邏輯理論.....	23				
第三章 研究方法.....	28	第一節 網頁的觀察與探討.....	28	第二節 研究方法.....	33	第三節 系統功能架構.....	37				
第四章 系統實作與評估.....	43	第一節 系統發展的環境與工具.....	43	第二節 系統人機介面.....	44	第三節 ImageFinder圖形檢索實驗.....	46	第四節 實驗結果分析與問題討論.....	63		
第五章 結論及研究方向.....	69	第一節 研究結論.....	69	第二節 未來研究方向.....	73						

## REFERENCES

- 一、中文方面 [1] 陳秀美, WWW檢索工具的選擇與比較, 資訊傳播 與圖書館學, 1996-6 [2] Dan and Judith wesley原著, 江永祥譯, 深入 Intranet, 松格資訊有限公司, 1997-6 [3] 資策會推廣服務處, 1999-1, <http://www.find.org.tw> [4] 卜小蝶, Internet資源搜尋系統與圖書館資訊服務, 中國圖書館學會會報, 83年12月 [5] 楊維忠, 泛論「模糊邏輯」Fuzzy Logic, CAD與自動化, 1991-6 [6] 蔡聯澄, 網路資源應用, 資訊與教育61期p.30-37.
- [7] 卜小蝶, 圖書資訊檢索技術, 文華圖書館管理資訊股份有限公司, 85年.
- [8] 簡立峰, 易尋系統(Csmart)與中文智慧型資訊系統, 21世紀資訊科學與技術的展望國際學術研討會論文集, 85年11月.
- [9] 簡立峰, 智慧型中文資訊檢索計畫成果與展望, 第一屆海峽兩岸科技資訊研討會, 1998年3月.
- 二、英文部份 [10] David C. Lee and Scott F. Midkiff, "A Sample Statistical Characterization of the World-Wide Web," Southeastcon '97. Engineering new New Century., Proceedings. IEEE , Page: 174 -178 .
- [11] <http://www.ipix.com> [12] Sougata Mukherjea, Kyoji Hirata, Yoshinori Hara. "Toward a multimedia World-Wide Web information retrieval engine". Computer Network and ISDN system 29 (1997) 1181-1191.
- [13] Stan Sclaroff, Leonid Taycher, and Marco La Cascia, "ImageRover: A Content-Based Image Browser for the World Wide Web." Content-Based Access of Image and Video Libraries, 1997.Proceedings IEEE Workshop on , Page(s): 2-9.
- [14] John R. Smoth and Shin-Fu Chang, "Visual Search the Web for Content," IEEE Multimedia Volume: 43 , Page(s): 12-20.

- [15] L.A.Zadeh, "Fuzzy sets," Inform. Contr., Vol8, pp.338-353, 1965.
- [16] Charlottle Jenkins, Mike Jackson, Peter Burden and Jon Wallis, "Searching the world wide web: an evaluation of Available tools and methodologies ",Information and SoftWare Technology 39 (1998) pp.985-994.
- [17] F. W. Lancaster, Information Retrieval Today. Arlington, VA : Information Resources Press, 1993.
- [18] Gerard Salton, An Introduction to Modern Information Retrieval , New York: McGraw-Hill, 1983 [19] L.Sauders, "The Virtual Library Revisited," Computers in Libraries 12 (Nov. 1992):51-54 [20] Katia Obraczka et al., "Internet Resource Discovery Services," Computer 26 (Sept. 1993): 8-22.
- [21] Chris Sherman, "Search engine help: documentation and resources on the web," Online. 22 (6) November 1998.
- [22] M.Marchiori, "The quest for correct information on the Web: hyper search engines,"In: Proc. Of the 6th International WWW Conference (WWW 97). Santa Clara. USA, April 7-11, 1997.
- [23] Sergey Brin and Lawrence Page, "The anatomy of a large-scale hypertextual Web search engine," Computer Network and ISDN Sytems 30(1998) 107-117.
- [24] Nicholas J. Belkin and W. Bruce Croft, "Information Filtering and Information Retrieval: Two Sides of the Same Coin?," Communication of the ACM 35(Dec.1992):29-38.
- [25] Shoshana Loeb, "Architecting Personalized Delivery of Multimedia Information," Communication Of The ACM,December 1992 , Vol.35 , No.12 , P.39-50.
- [26] Peter W. Foltz and Susan T. Dumais , "Personalized Information Delivery : An Analysis of Information Filtering Methods," Communication Of The ACM, December 1992,Vol.35 , No.12 , P.51-60 [27] The IBM Agent, <http://activist.gpl.ibm.com:81/WhitePaper/ptc2.htm> [28] Russell, J., and P. Norving. Artificial Intelligence: A Modern Approach. Englewood Cliffs, NJ:Prentice Hall.1995 [29] Maes.P, "Artificial Life Meets Entertainment : Life-Like Autonomous Agents," Communication of the ACM, Vol.37, No.7.1995 [30] Smith, D. D., et al."KidSim: Programming Agents Without a Programming Language."Communication of the ACM, Vol.37, No.7.1994.
- [31] Hayes-Roth B, "An Architecture for Adaptive Intelligent Systems." Artificial Intelligence: Special Issue on Agents and Interactivity, Vol.72, Nos.1-2.
- [32] Eferaim Turban and Jaye Aronson, Decision Support Systems and Intelligent System. Phippe,Prentice Hall.1997 [33] Pattie Maes, "Agents that Reduce Work and Information Overkiad," Communications Of The ACM, July 1994, Vol.37,No.7, p.30-40 [34] O"Leary, D. "AI and Navigation on the Internet and Intranet "IEEE Expert.1996,April [35] Etzioni,O.,and D.S. Weld."Intelligent Agents on the Internet: Fact, Fiction, and Forecast." IEEE Expert, Vol.10, No.4.
- [36] Lee,J.K.,and W.Lee. "Intelligent Agent Based Contract Process in Electronic Commerce : Unik\_Agent Approach."Proceedings of the Thirtieth Hawaii International Conference on Systems Sciences. Wailea, HI; Los Alamitos,CA:IEEE Computer Society Press.
- [37] David W. Cheung, Ben Kao, Joseph Lee. "Discovering user access patterns on the World Wide Web". Knowledge- Based System 10 (1998) 463-470.
- [38] J.S.R.Jang, C.T.Sun, E.Mizutani. Neuro-Fuzzy AND Soft Computing. Prentice-Hall International, Inc.1998.
- [39] Kiyoshi.C. "Auto Magazine" 1998-11.