

影像反白與漸層文字之偵測與二值化

何彥廷、曾逸鴻

E-mail: 321985@mail.dyu.edu.tw

摘要

由於多媒體技術不斷成長，導致彩色文件影像的遽增，但彩色文件影像包含許多不同的物件，例如：不同顏色的字、圖形、風景或背景顏色等，甚至這些物件會有彼此交疊的現象，一般光學文字識別軟體已無法處理含有彩色背景的文件影像。基於此，本研究將針對文件影像中的文字物件與背景物件進行分析，進而利用了影像本身中不同的物件之間的不同特性，加以分類。除了可輕易處理一般單純文件影像之外，更可以處理彩色文件影像。

關鍵詞：光學文字識別、影像二值化、文字偵測

目錄

| | | | |
|-----------------|-----|--------------|-----|
| 中文摘要 | iii | 英文摘要 | iii |
| iv 誌謝辭 | | v 內容目錄 | |
| vi 表目錄 | | vii 圖目錄 | |
| viii 第一章 緒論 | 1 | 第一節 研究背景與動機 | 1 |
| 1 第二節 研究目的 | 3 | 第三節 系統流程 | 3 |
| 4 第四節 研究範圍與限制 | 6 | 第五節 論文架構 | 6 |
| 7 第二章 文獻探討 | 8 | 第一節 光學文字識別 | 8 |
| 8 第二節 影像二值化 | 12 | 第三節 文字偵測 | 12 |
| 15 第三章 反白文字判別方法 | 19 | 第一節 反白文字特性 | 19 |
| 19 第二節 反白文字偵測 | 20 | 第四章 漸層文字判別方法 | 20 |
| 27 第一節 漸層文字特性 | 27 | 第二節 漸層文字偵測 | 27 |
| 28 第五章 實驗結果 | 35 | 第六章 結論 | 35 |
| 40 參考文獻 | 41 | | |

參考文獻

- 英文部分 Agnihotri, L., & Dimitrova, N.(1999). Text detection for video analysis. IEEE Workshop on Content-Based Access of Image and Video Libraries, 109-113. Washington, DC, USA. Bernsen, J. (1986). Dynamic thresholding of grey-level images. Pattern Recognition, 2, 1251-1255. Blayvas, I., Bruckstein, A., & Kimmel, R. (2006). Efficient computation of adaptive threshold surfaces for image binarization. Pattern Recognition, 39, 89-101. Chun, B. T., Bae, T., & Kim, T. Y. (1999). Text Extraction in videos using topographical features of characters. Fuzzy Systems Conference Proceedings, 1999 IEEE International, 2, 1126-1130. Chow, C.K., & Kaneko, T. (1972). Automatic boundary detection of the left ventricle from cineangiograms. Computers and Biomedical Research, 5(4), 388-410. Chena, Q., Suna, Q., Hengb, P. A. & Xiaa, D. (2008). A double-threshold image binarization method based on edge detector. Pattern Recognition, 41(4), 1254-1267. Chen, S., & Li. D. (2006). Image binarization focusing on objects. Neurocomputing, 69(16), 2411 – 2415. Chaddha, N., Sharma, R., Agrawal, A., & Gupta, A. (1995). Text segmentation in mixed-mode images. 28th International Conference on Signals, Systems and Computers, 2, 1356-1361. Dawoud, A. (2007). Iterative cross section sequence graph for handwritten character segmentation. IEEE transactions on image processing, 16(8), 2150-2154. Eikvil, L., Taxt, T., & Moen, K. (1991). A fast adaptive method for binarization of document images. First International Conference on Document Analysis and Recognition, 3, 435-443. Garcia, C., & Apostolidis, X. (2000). Text detection and segmentation in complex color images. Acoustics, Speech, and Signal Processing, 5, 2326-2329. Gargi, U., Antani, S., & Kasturi, R. (1998). Indexing text events in digital video databases. 14th International Conference on Pattern Recognition (ICPR'98), 1, 916-918. Gatos, B., Pratikakis, I., & Perantonis, S. J. (2006). Adaptive degraded document image binarization. Pattern Recognition, 39(3), 317-327. Gupta, M. R., Jacobson, N. P., & Garcia, E. K. (2007). OCR binarization and image pre-processing for searching historical documents. Pattern Recognition, 40(2), 389-397. Goto, H. (2006). OCRGrid: A platform for distributed and cooperative OCR systems. The 18th International Conference on Pattern Recognition, 2, 982-985. Jain, A. K., & Bhattacharjee, S. (1992). Text segmentation using gabor filters for automatic document processing. Machine Vision and Application, 5, 169-184. Jain, A. K., & Yu, B. (1998). Automatic text location in images and video frames. Pattern Recognition, 31(12), 2055-2077. Jeong, K. Y., Jung, K., Kim, E. Y., & Kim, H. J. (1999). Neural network-based text location for news video indexing. 1999 International Conference on Image Processing, 4, 319-323. Kamel, M.,

& Zhao, A. (1993). Extraction of binary character/graphics images from gray scale document images. *Graphical Models and Image Processing*, 55(3), 203-217. Kurita, T., Otsu, N., & Abdelmalek, N. (1992). Maximum likelihood thresholding based on population mixture models. *Pattern Recognition*, 25(10), 1231-1240. Kwon, S. H. (2004). Threshold selection based on cluster analysis. *Pattern Recognition*, 25(9), 1045-1050. Liao, P. S., Chen, T. S., & Chung, P. C. (2001). A fast algorithm for multilevel thresholding. *Journal of Information Science and Engineering*, 17, 713-737. Long, T., & Jin, L. (2008). A novel orientation free method for online unconstrained cursive handwritten chinese word recognition. *The 19th International Conference on Pattern Recognition*, 28, 1-4. Cai, M., Song, J., & Lyu, M. R. (2002). A new approach for video text detection. *International Conference on Image Processing*, 19, 117-120. Miyao, H., Nakano, Y., Tani, A., Tabaru, H., & Hananoi, T. (2004). Printed Japanese character recognition using multiple commercial OCRs. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 8(2), 200-207. Nagy, G. (2000). Twenty years of document image analysis in PAMI. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1), 38-62. Niblack, W. (1986). *An introduction to digital image processing*. Newjersey: Prentice Hall, 115-116. Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE International Conference on System, Man, and Cybernetics*, 9(1), 62-66. Shin, C. S., Kim, K. I., Park, M. H., & Kim, H. J. (2001). Support vector machine-based text detection in digital video. *Pattern Recognition*, 34, 527-529. Zhao, S. L., & Leem, H. J. (2001). High-precision two-kernel Chinese character recognition in general document processing systems. *Sixth International Conference on Document Analysis and Recognition*, 15, 617-621. Sezgin, M., & Sankur, B. (2004). A survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13(1), 46-168. Shijian, L., & Tan, C. L. (2008). Script and language identification in noisy and degraded document images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(1), 14-24. Chou, T. R., & Chang, F. (1997). Optical chinese character recognition for low-quality document images. *Proceedings of the Fourth International Conference on Document Analysis and Recognition*, 3, 608-611. Mao, W., Chung, F. L., Lam, K. K. M., & Siu, W. C. (2002). Hybrid Chinese/English text detection in images and video frames. *International Conference on Pattern Recognition*, 310-315. Wu, V., Manmatha, R., & Riseman, E. M. (1997). Finding text in images. *Proceedings of the second ACM international conference on Digital libraries*, 7, 3-12. Tang, X., Luo, B., Pissaloux, E., & Zhang, H. (2002). Video text extraction using temporal feature vectors. *IEEE International Conference on Multimedia and Expo*, 1, 85-88. Hua, X. S., Yin, P., & Zhang, H. J. (2002). Efficient video text recognition using multiple frame integration. *2002 International Conference on Image Processing*, 13, 71-75. Yan, C., Sang, N., & Zhang, T. (2003). Local entropy-based transition region extraction and thresholding. *Pattern Recognition*, 24(16), 2935 – 2941. Yi Lu, & Shridhar, M. (1996). Character segmentation in handwritten words - An overview. *Pattern Recognition*, 29(1), 77-96. Zhong, Y., Karu, K., & Jain, A. K. (1995). Locating text in complex color images. *Pattern Recognition*, 28(10), 1523-1536.