

視訊畫面之場景文字偵測與辨識

謝明遠、曾逸鴻

E-mail: 360553@mail.dyu.edu.tw

摘要

人眼有時需專注特定事務，無法兼顧週遭環境，例如：開車；人們在開車的過程中，必須專注前方交通狀況，注意行車安全與行人的動向，容易忽略標示牌的指示訊息。因此，配合電腦視覺技術，辨識與尋找場景中的標示牌，可輔助人眼進行搜尋與觀看標示牌內容。本研究中有兩個重點，包括傾斜文字的學習與訓練以及在連續畫面中的文字辨識，由於拍攝角度的不同，文字常常出現傾斜的情況，導致傳統的文字辨識使用在場景文字的辨識中效果並不好，本研究利用傾斜文字的學習與訓練，讓文字辨識核心學習傾斜文字特徵，對場景中的文字進行辨識，而以往的影像辨識幾乎都是在單張的影像上進行尋找目標以及辨識分析，從視訊設備拍攝的連續影像中可以找出相對比較清晰的影像，可以降低影像處理的難度以及同時提升影像辨識的效果，因此搭配視訊之連續畫面進行文字辨識，可提高文字辨識的正確率。本研究測試部分，使用行車記錄器預錄畫面來模擬測試實際情況，在系統偵測標示牌擷取的正確率達到92.43%，在傾斜文字中的文字辨識率也達到85.63%，本研究解決了以往場景文字辨識上的不足部份。

關鍵詞：傾斜文字、場景文字辨識、電腦視覺

目錄

中文摘要	ii	英文摘要	ii
iii 致謝辭		iv 內容目錄	
v 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	2	第三節 系統流程	2
2 第四節 研究範圍與限制	3	第五節 論文架構	3
4 第二章 文獻探討	5	第一節 文字區塊	5
5 偵測與文字擷取	5	第二節 傾斜文字校正	6
9 校正	9	第三節 文字辨識與	6
12 畫面之前處理	12	第一節 視訊畫	6
18 取	18	面之擷取	6
22 與特徵學習	22	第二節 畫面場景標示牌之定位	16
27 與檢索	27	第三節 標示牌文字之擷	16
35 參考文獻	35	取	16
		第一節 傾斜文字樣型之分類	22
		第二節 標示牌傾斜類型判定	26
		第三節 連續畫面之標示文字辨識	26
		第六章 結論	36

參考文獻

- Bugeau, A., & Perez, P. (2009). Detection and segmentation of moving objects in complex scenes. *Computer Vision and Image Understanding*, 113(4), 459-476.
- Chang Rong-Chi (2011). Intelligent text detection and extraction from natural scene images. *Nano, Information Technology and Reliability (NASNIT)*, 2011 15th North-East Asia Symposium, 23-28
- Christos Conomis(2006). Conics-Based Homography Estimation from Invariant Points and Pole-Polar Relationships. *Proceedings of the Third International Symposium on 3D Data Processing, Visualization, and Transmission (2006)*, 908 - 915
- Gao Yu-peng, Li Yi-ming and Hu Zhong-yi (2011). Skewed Text Correction Based on the Improved Hough Transform. *Image Analysis and Signal Processing (IASP)*, 2011 International Conference, 368-372
- George Nagy(2000). Twenty Years of Document Image Analysis in PAMI. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, NO. 1
- Jirasuwankul N.(2011).Effect of Text Orientation to OCR Error and Anti-Skew of Text using Projective Transform Technique. *Advanced Intelligent Mechatronics (AIM)*, 2011 IEEE/ASME International Conference, 856-861.
- Katherine L. Bouman, Golnaz Abdollahian, Mireille Boutin, and Edward J. Delp(2011). A Low Complexity Sign Detection and Text Localization Method for Mobile Applications. *Multimedia*, IEEE Transactions Vol.13, 922 - 934 .
- Keechul Junga, Kwang In Kimb, Anil K. Jainc (2004).Text information extraction in images and video: a survey. *Pattern Recognition* 37, 977 – 997.
- Kim Egyul, Lee SeongHun, Kim JinHyung(2009). Scene Text Extraction using Focus of Mobile Camera. *Document Analysis and Recognition*, 2009. ICDAR '09. 10th International Conference on, 166 - 170
- Lee,C.M. A. Kankanhalli (1995). Automatic extraction of characters in complex images, *Int. J. Pattern Recognition Artif. Intell.* 9(1) (1995) 67 – 82.
- Paul D. Gader , Mohamed Ali Khabou(1996).

Automatic Feature Generation for Handwritten Digit Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 18, NO 12, Pawan Kumar M., Jawahar C.V., Narayanan P. J. (2004). Building blocks for autonomous navigation using contour correspondences. Image Processing, 2004. ICIP '04. 2004 International Conference Vol.2 ,1381 - 1384 . Qing Laiyun, Wang Weiqiang, Gao Wen (2002). Automatic Text Extraction and Recognition for Video Indexing and Retrieval. National Natural Science Foundation of China under Grant No. 69789301. Sarfraz M., Mahmoud S.A. and Rasheed Z. (2007). On Skew Estimation and Correction of Text. Computer Graphics. Imaging and Visualisation, 2007. CGIV '07 Sriganesh Madhvanath, Venu Govindaraju (2001) The Role of Holistic Paradigms in Handwritten Word Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, VOL. 23, NO. 2, FEBRUARY 2001 Tsai L.-W., Hsieh J.-W., Chuang C.-H., Tseng Y.-J., Fan K.-C., Lee C.-C.(2008) .Road sign detection using eigen colour. Published in IET Computer Vision Received on 18th September 2007. Vol.2,164-177 Wang Yuming, Naoki Tanaka(2008).Text String Extraction from Scene Image Based on Edge Feature and Morphology. The Eighth IAPR Workshop on Document Analysis Systems. 323 – 328 Yang Haojin, Maria Siebert, Patrick Luhne, Harald Sack, Christoph Meinel(2011). Automatic Lecture Video Indexing and Analysis Using Video OCR Technology. Multimedia (ISM), 2011 IEEE International Symposium. 111 – 116. Yamaguchi, T.; Nakano, Y.; Maruyama, M.; Miyao, H.; Hananoi, T.(2003). Digit classification on signboards for telephone number recognition. Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003). Vol.1 359 – 363. Zhong Y., Karu K., Jain A.K. (1995). Locating text in complex color images, Document Analysis and Recognition, 1995., Proceedings of the Third International Conference, Vol.1 ,146-149. Oivind Due Trier , Anil K Jain , Torfinn Taxt(1996).Feature Extraction Methods For Character Recognition - A Survey. Pattern Recognition Vol.29 , No4 ,641-662.