

Study on Key Techniques of Dynamic Text Recognition System

Chengxin Yan

*College of Mechanical Engineering, China University of Petroleum, Qingdao 266555, China
yanjack@163.com*

Keywords: Dynamic text, feature extraction, text recognition.

Abstract: Key techniques on dynamic text recognition system are studied in this paper. Rectangularization algorithm is adopted for text localization. Vertical projection based on a priori knowledge is used to achieve twice character segmentation. Coarse mesh feature and 13-point features are used and support vector machine is adopted as the final classifier. A dynamic text recognition system is developed on the basis of those key techniques and can be used to logistics, banking, postal and other similar areas.

1. Introduction

Dynamic text recognition is a research hot spot in the field of text recognition, involving image processing, data mining, pattern recognition and other interdisciplinary subjects. Dynamic text recognition is applied more and more widely in many fields, such as logistics, transportation, postal service, bank and automated office. In this paper, the key technology in the process of dynamic text recognition is studied, and a dynamic text recognition software system is developed, which can be used in the fields of logistics, banking, post and other fields.

2. Key Techniques on Dynamic Text Recognition

In the process of dynamic text recognition, the text location, character segmentation, feature extraction and recognition algorithm are four important aspects. The selection and research of the four aspects play a very important role in improving the system recognition effect.

2.1 Text Location

Commonly used grayscale image text location methods include Hough transform [1], mathematical morphology [2], edge detection [3] and so on. In order to increase the reliability of the detection, this paper uses the Sobel operator to get the edge image. Then the redundant edge is removed by the morphological operation, and the text location is obtained by using the rectangularization algorithm [4].

In order to find the largest rectangle of a connected domain, the rectangularization transformation is to transform the irregular connected domain into the rectangular connected domain. Contour tracking is used to find rectangles, and then we judge whether the rectangle area

meets the requirements based on prior knowledge. The steps of the rectangularization algorithm are as follows:

- (1) Read a binary image;
- (2) Scanning each pixel of the image, if the current pixel grayscale value is 255 and all the grayscale values of the four points on the top and bottom of the pixel are 0, the gray value of the pixel is 0; if the current pixel grayscale value is 0, and the sum of the grayscale value of the upper, lower, left and right four pixels is at least 510, the grayscale value of the pixel is 255.
- (3) Loop step (2), so that all irregular connected domains can be transformed into rectangular connected domains.

2.2 Character Segmentation

The first task for character segmentation is to revise incline, then remove the edges and denoise. Finally, the single character is segmented.

2.2.1 Incline Revise

The purpose of incline revise is to adjust the position and posture of the characters to a reasonable state so that the text can be effectively segmented. The revise of a character image can be divided into three steps: incline judgment, incline information extraction and rotation revise.

- (1) Get the text rectangle by using the text localization algorithm mentioned above.
- (2) Contour detection is used to detect the contour of the four sides and store the detected contour.
- (3) The quadrangle most matched with the character area is selected from the quadrangles detected based on the prior knowledge, and the oblique angle of the quadrangle is extracted, and this angle is the oblique angle of the character area.
- (4) Rotate the character image with the extracted oblique angle to complete the correction.

2.2.2 Edge Remove and Denoise

The corrected characters contain external edge and noise which should be removed. In order to improve the speed of the algorithm, line scan method is used to remove the edges. The basis of line scanning is that there will be a large number of gray jumps in the character block area because of the existence of the character sequence. In the binary image, the exact edge can be judged by the statistics of whether the number of edges is reached to a certain threshold or not. The detail steps of this method are as follows:

- (1) Scan the image line from up and down row by row. When the frequency of the scan to a line is higher than a certain threshold, it is considered to find the upper edge of the block area, stop scanning, and record the top row number $R1$.
- (2) Scan the image from bottom to top as step 1, and record the bottom row number $R2$. At the same time we obtain the height of the image $H=R2-R1$;
- (3) Use the top and bottom rows obtained in the first two steps to obtain an image which edges outside those two rows removed.

In order to improve the accuracy of segmentation, morphological operation and median filter should be performed to made the contour of the characters' smooth.

2.2.3 Single Character Segmentation

The characters can be segmented after edge remove and denoise. The common character segmentation methods include vertical projection method [5], template matching method [6],

connected area method [7], cluster analysis [8] and so on. In practice, a single segmentation method is generally not enough. Several suitable improved algorithms are often selected for different applications. The system to be developed in this paper is mainly used in the indoor environment of logistics, bank, post and so on. The background is relatively simple. In this paper we will first make rough positioning through the vertical projection method, and then judge the rough location according to the prior knowledge. Finally, the twice segmentation of characters is realized by judging, dividing, merging and discarding respectively. The detailed steps are as follows:

(1) Perform vertical projection on binary images, the number of white (white characters on the black background) pixels in each column is counted and be put into the variable $nCount[i]$. The length of the variable is number of column of the image.

(2) Set a threshold of T , generally selecting a small integer (such as 1, 2, etc.), starting from the first data of the array, and comparing the value of $nCount[i]$ with T . Getting the first pixel greater than or equal to the T position, corresponding to the subscript of the array $nCount[i]$, is recorded as $Start1=i1$; Go on comparing till the value of $nCount[i]$ is less than or equal to T , the subscript of the array is recorded as $End1=i2$. Till now the first character area is found.

(3) Let $Width1=End1-Start1$, and let $Ratio(i) = Width(i) / H$ be the width to height ratio of the i th segmented character. From prior knowledge we know the standard width to height ratio is 1:2;

(4) After the width to height ratio of the first character is obtained, the judgment can be made as follows: (1) If $Ratio(1) < 0.35$, the obtained character area maybe too narrow (for example, 1, I, J, etc.) or there maybe exist some polluted area not removed. Judgment should be performed again. When $nCount[(Start1+End1) / 2] > T_x$ then this area is considered as a character, otherwise it is considered as pollution area and being abandonment; (2) If $0.35 < Ratio(1) < 0.65$ then it is considered successful to get characters; (3) If $0.65 < Ratio(1) < 1.25$, then it is considered that the characters of adhesion are obtained, and the intermediate values of $Start1$ and $End1$ are taken directly to separate characters; (4) If $1.25 < Ratio(1)$, it is considered to be lost and give up segmentation;

(5) Repeat 2-4 steps, traverse the array and divide all characters.

(6) Confirm the number of segmented characters, if the number is equal to the number obtained by prior knowledge, then the segmenting process could be finished, otherwise the threshold should be adjusted and the segmenting process should be repeated.

2.3 Feature Selection

Feature selection is also one of the key steps in the dynamic text recognition process, which requires that the selected features have good stability and easy to be separated. The commonly used feature extraction methods include 13-points feature extraction method [9], coarse grid feature extraction method [10] and PCA dimension reduction method [11].

Coarse grid feature extraction method [10] is proved have good stability and separability and being adopted in our dynamic character recognition system [12].

2.4 Recognition Algorithm

The commonly used text recognition algorithms include template matching classifier, probability statistics based Bias classifier, geometric classifier, neural network classifier and classifier based on support vector machine (SVM) [13] [14].

SVM has good performance on classifying either characters or numbers [12], especially the number of sample is limited. So in our recognition system SVM is adopted as classifier.

3. Implementation of Dynamic Text Recognition System

3.1 The Hardware Composition of the System

The hardware of the recognition system designed in this paper includes the image acquisition card, the CCD industrial camera, the industrial level security protection computer, and the software platform for the Windows operating system.

3.2 Developing Environment

OpenCV is an excellent computer vision open source library issued by Intel [15]. In this paper the system is developed under Windows operating system, VS2005 compiler and C/C++ language, and based on OpenCV2.0 open source library.

3.3 System Module Composition

The system mainly includes preprocessing, image manipulation, image transformation, recognition process, character recognition, and help module. The preprocessing part can perform image enhancement, and denoising. Image manipulation includes edge detection, converting to binary image and morphological operation. The image transformation is mainly homomorphic filtering, and the recognition process includes single character feature extraction, character segmentation, location matting and so on; character recognition modules include template matching recognition, support vector machine recognition and other algorithms. Through the processing of text location, character segmentation, feature extraction, recognition algorithm and so on, the character recognition data in text format is finally obtained for the operator to further manage the product information.

The system designed in this paper can be used in the field of dynamic text recognition, such as logistics, banking, postal service, oil pipe pole identification, etc., and has broad application prospects.

References

- [1] Xiao Zhitao, et al. *Detection of oblique text images based on Hough transform. Infrared and Laser Engineering*, 2002, 31 (4): 315-317,350.
- [2] Li Wei, et al. *A new license plate location algorithm based on mathematical morphology and color features. Bulletin of Science and Technology*, 2009 (2): 214-21.
- [3] Li Zhaohui, Yu Yinglin. *Automatic detection of video text based on edge information and LH. Computer Application Research*, 2004, 12:166-167.
- [4] Feng Guojin, Gu Guohua. *License plate automatic location and fuzzy recognition algorithm. Opto-electronic Laser*, 2003, 7: 750-752.
- [5] Hirsch. J. W, et al. *Shadow elimination for effective moving object detection by Gaussian shadow modeling. International Journal of Image and Vision Computing*.2003, 21(6):505-516.
- [6] Fan Weiqi, Mu Changjiang. *A character segmentation method for vehicle license plates based on the Chinese character structure. Journal of Instrument and Meter*, 2003,24 (4): 472-474.
- [7] Yang.F, et al. *A novel approach for license plate character segmentation. The 1st IEEE Conference on Industrial Electronics and Applications, Singapore,2006:1-6.*
- [8] Xie Weisheng. *Research and implementation of license plate location and character segmentation algorithm. Southwest Jiao Tong University*, 2010.6:32.
- [9] Huang Xiangxing. *Research and implementation of key technologies of license plate recognition system. Beijing Jiaotong University*, 2010.6: 40.
- [10] Lee Liaoliao. *Research on character recognition method in industrial field. Hefei University of Technology*, 2003.3:39.
- [11] Lin Bo. *Research on key technologies of automatic license plate recognition. Chongqing University*, 2010.4:39-40.

- [12] Wen Bin, *Research on key technologies of dynamic text recognition. Master's Thesis. Qingdao, China University of Petroleum. 2011*
- [13] Xu Hailan, et al. *Three methods of character recognition in pattern recognition. Journal of Beijing Broadcasting Institute, 2005,12 (4): 34-38.*
- [14] Lee S. W., et al. *A new methodology for gray-scale character Segmentation recognition. IEEE Trans. Pattern Analysis and Machine Intelligence, 1996, 18(10): 1045-1050.*
- [15] Jiang Yongcheng, et al. *OpenCV based mobile robot video streaming acquisition and processing. Machine Tools and Hydraulics, 2010, 38 (15): 40-43.*