

# THE RESEARCH AND REALIZATION OF VEHICLE LICENSE PLATE CHARACTER SEGMENTATION AND RECOGNITION TECHNOLOGY

JIAN-XIA WANG, WAN-ZHEN ZHOU, JING-FU XUE, XIN-XIN LIU

College of Information Science & Engineering, Hebei University of Science and Technology,  
Hebei Shijiazhuang 050018, China  
E-MAIL: wang\_jianxia@126.com

## Abstract:

On the basis of the vehicle license plate location, an image grey vertical projection segmentation approach based on the distribution character segmentation is proposed in this paper. A two-stage approach consisting of coarse and accurate segmentation is adopted. It can increase the accuracy of the segmentation and has good segmentation speed. And in recognition process, character features are extracted from character segmentation results, in order to identify character exactly, an improved template matching method is used to character recognition. Experimental results show that character segmentation method is efficient and quick and recognition algorithm is applicable.

## Keywords:

License Plate Location; Character segmentation; Template Matching; Character recognition; Improved template matching

## 1. Introduction

Automobile License Plate Recognition System is an important link of realizing intelligent traffic management. The vehicle license plate recognition system focuses on the key technologies, which include the license plate region location, license plate character segmentation and license plate character recognition [1]. License Plate Character Segmentation plays the role to segment character from the located region of license plate, based on vehicle license location results. The segmentation result of license plate character is the precondition and preparation for character recognition. License plate character recognition is mainly to process, extract character feature and identify the license plate characters segmented. Character recognition is the critical step in the system, which directly determines the performance of the recognition system. In this paper, on the basis of the domestic and foreign researchers' work [2], we have analyzed and explored the existing license plate character segmentation and recognition algorithms, a license

plate character segmentation method and recognition algorithm is presented and described in detail, and the corresponding experimental results are also given.

## 2. License Plate Character Segmentation

The first character in China's current vehicle license is the abbreviation of the Chinese province name, the second character is an English letter, other five letters are number or English letter, and the English letter is no more than two. In order to identify the Chinese character, English letter and number on the license plate, each character must be segmented. Character segmentation will affect the accuracy of character recognition.

### 2.1. Comparison of Character Segmentation methods

The license plate character segmentation will divide license plate region into single character area, each character area must be the smallest rectangle area that contains a single character. The more precise the segmentation is, the better the recognition effect is. The following compares several common license plate character segmentation methods:

(1) *The vertical projection method* The character segmentation method using vertical projection to divide characters is rather simple and accurate, but segmentation effect could be influenced by noisy and the slope degree of the plate easily[3].

(2) *Connected domain method* This license plate character segmentation method has smaller influence by the slope degree of the plate, but the character segmentation is not precise enough.

(3) *Clustering analysis method* Clustering analysis method has solved the issue of disconnected character better, but the program's logic design is complex, more loop nesting, and execution time being slower than the former two.

Through the comparison of mentioned above several ways, considering the license plate recognition system need

for real-time performance and accuracy, an image grey vertical projection segmentation approach based on the distribution character segmentation is applied here. In order to avoid effect of noise and license plate slope degree to segmentation, noise is reduced in ahead image pre-processing, the slope image is corrected in the vehicle license plate location. Character segmentation includes coarse segmentation and precision segmentation.

## 2.2. Character coarse segmentation method

Character coarse segmentation algorithm can complete characters segmentation basically, the special condition can be completed by the precise segmentation method. Character coarse segmentation process is as follows.

(1) Remove the possible upper and lower plate border by searching the license plate from top to bottom to find a row whose projection value less than half of the average projection value. This row will be set to license plates top edge. This row will be empty if it is found in the image. Then searching the plate from bottom to top, if we find a row whose projection value less than half of the average projection value, this row will be set to the plate bottom. This row will be empty if it is found in the license plate image.

(2) Remove possible left and right borders of the license plate by searching the license plate from left to right to find these column whose projection value less than 1/5 of license plate height and those continuous n columns whose projection value larger than 3/5 of vehicles plate height, and n is half of vehicle height, then the new left and right edge distance are set up.

(3) If the license plate after the above treatment does not meet the size requirement, its size is too small or too large, this license plate will be deleted from the candidate license plate, and the next candidate license plate will be processed.

(4) The empty column will be cleared from the most left of license plate. Firstly, 0 is assumed the width of the empty column, the segmented character width is over the scope of the normal character width if this empty column is cleared, width of the empty column will be increased by one, and then segmentation will restart.

(5) Check the segmented character, it will be thought as interferential points if the segmented character is too small or the character strokes are too few, then this segmentation result will be discarded, the next character segmentation will be continued.

(6) Eliminate the empty row of upper and lower boundary of the segmented character. The possible existing license plate frame will be searched within 1/5 of height from the upper and lower boundary. The new upper and lower edge will be set up if the new license plate frame is found.

(7) Check the upper and lower boundary of the segmented character, if they are within the normal range, this segmentation will be preserved, and then jump to Step (4) to segment the next character until all characters are segmented completely.

(8) All the Segmented characters are checked, and too small segmentation characters are united.

(9) If the number of the segmented characters is less than 5, this character segmentation is failed, then this license plate is removed from the candidate license plate.

(10) The next candidate license plate will be continued, until all candidate license plates have been processed.

License plate image is processed by the mentioned above method, the obtained result is shown as Figure 1 (b).

## 2.3. Character precision segmentation process

There may be long block, ultra-short block and other interference block in coarse segmentation process, the approach of segmenting long block, merging ultra-short block and removing the interference block will be described as follows in detail.

(1) The segmented character is selected to process word for word from the candidate license plates.

(2) Remove the empty row of the upper and lower boundary of segmented character. The possible license plate frame will be searched within 1/4 of height from the upper and lower boundary of character. The new upper and lower edge of the character will be set up if a new license plate frame is found.

(3) Check if there is the character with two blank columns, this character may be the Chinese word "||", otherwise, it can be the left border or the right border of license plate frame.

(4) And then check if the character stroke of left blank column is greater than 2 pixels, and then check whether it is the “丷” radical structure.

(5) Check whether the amount of character segmented is larger than 7, the amount more than 7 means there are ultra-short partitions. And then the possible segment needs to be merged.

(6) Continue with the next candidate license plate, until all candidates have been processed.

A large number of license plate images are processed using the mentioned above method, the obtained results are in accord with the criteria, the segmented results are shown as figure 1 (c).



a) License plate image after pretreatment

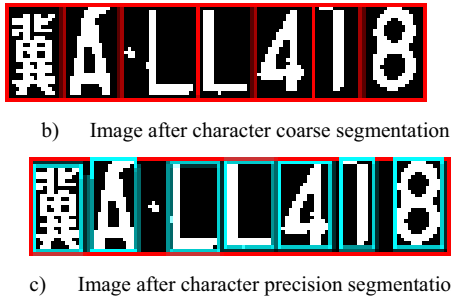


Figure1. License Plate Character Segmentation Chart

### 3. License Plate characters recognition

The template matching is a classical pattern recognition method, and it is the most direct character recognition method. Firstly, the sample is classified and the sample databases are established. Then the recognized characters are binarization processed and normalized by the template size in the character database, it will match with all templates and calculate their similarity, finally the best match will be chosen as the result.

#### 3.1. The implement principle of template matching[5]

There is M kinds of classifications:  $w_1, w_2, \dots, w_M$ , there are several vector identifications in each classification's characteristic vector, for example  $X_i$  is characteristic vector of  $w_i$  classification, then

$$X_i = \begin{Bmatrix} x_{i1} \\ x_{i2} \\ x_{i3} \\ \vdots \\ x_{in} \end{Bmatrix} \quad (1)$$

To the arbitrary character recognized  $X$ ,  $X = \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{Bmatrix}$  (2)

Calculate the distance of  $d(X_i, X)$ , if existing  $i$ , let  $d(X_i, X) < d(X_j, X), j = 1, 2, \dots, M, i \neq j$  (3)

Then  $X \in w_i$ . The distance of  $X, Y$  two points can be expressed

$$\begin{aligned} & |X - Y|^2 \\ & d(X, X_i) \\ & = |X - X_i|^2 \\ & = (X - X_i)^T (X - X_i) \\ & = X^T X - X^T X_i - X_i^T X + X_i^T X_i \\ & = X^T X - (X^T X_i + X_i^T X - X_i^T X_i) \end{aligned} \quad (4)$$

$X^T X_i + X_i^T X - X_i^T X_i$  in the formula(4) is the linear function of the characteristic, it can be regarded as discriminant function.

$$d_i(X) = X^T X_i + X_i^T X - X_i^T X_i \quad (5)$$

If  $d(X, X_i) = \min\{d_i(X)\}$ , then  $X \in w_i$

The distance of tested sample and all known sample in training concentrate is circulation calculated to find out the smallest distance between tested sample and known sample. The classification of known sample is classification of the tested sample.

#### 3.2. Improved template matching method

The traditional template matching classifier treats each sample of the training sample concentration as a standard template, using the tested sample compared with the standard template, to identify the most similar and the nearest standard template, the standard template classification is as its own category. For example, class A has eight training samples, so it has eight templates. Class B has eighteen training samples, so it has eighteen templates. In classifying, similarity is calculated between any tested sample and the twenty six templates, and the most similar template is found, if the template is class B, we determine the tested sample is as one of B class, otherwise it belongs to class A. But a large amount of calculation is an obvious disadvantage, in this paper, license plate characters arrangement is as  $X_1 X_2 \cdot X_3 X_4 X_5 X_6 X_7$ , where  $X_1$  is Chinese characters, it is shortened forms of the Chinese provinces and cities,  $X_2$  is the English letter, it is the licensing authority department code,  $X_3 X_4 X_5 X_6 X_7$  is the combination of English letters and numbers. In order to enhance recognition speed and recognition rate, an improved template matching method is presented. Its recognition process is as follows.

(1) Firstly, the  $X_1$  template is used to match. It is only necessary to compare the standard template for each Chinese character with the Chinese character of objective license plate,

and calculate their similarity, because X1 is Chinese character.

(2) Finding two Chinese characters of the greatest similarity, and then all of their templates are compared to find the most similar characters.

(3) Then X2 template is used to match. The English letter of the greatest similarity with object letter is searched from the standard letter template, and then it is subdivided.

(4) The same operation is also finished with X3, X4, X5, X6 and X7, their templates are not only the standard letter

templates but also the standard number templates.

This method can reduce the times of comparisons and improve recognition efficiency.

#### 4. Test result and analysis

The following table compares recognition result of improved template matching method and traditional template matching method, the test result is shown as table.1

TABLE.1 CHARACTERS RECOGNITION TEST TABLE

Network	Image numbers	recognized images	recognition rate improved/tradition	average response time improved / tradition(ms)
Chinese word	500	453	91.7 % / 88.7 %	289 / 315
Letter/Number	500	460	92.8 % / 90.7 %	192 / 230
Letter	500	469	93.6% / 91.0 %	263 / 287
Number	500	478	94.3 % / 92.4%	154 / 195

As seen from the table 1, using the improved template matching method, the average recognition rate and recognition speed of Chinese characters, letters and numeric characters have been enhanced, it is more than the recognition rate and recognition speed before improved. Furthermore, in this paper, the more than 1000 samples are trained, the more than 2,000 samples are tested, and these samples were from the actual violation vehicle images of traffic intersection. Therefore, the test result is more general and universal.

#### 5 Conclusion

In this paper, the process of license plate character segmentation and character recognition are described in detail. The license plate characters are segmented by a method of the connection of coarse segmentation and accurate segmentation, and character recognition adopts an improved template matching recognition technology. A fast and efficient license plate recognition system is designed and implemented. Through testing to the 500 pieces different backgrounds and lighting conditions images, the

test results show that this method can achieve good character segmentation and high recognition rate.

#### Acknowledgements

This work is sponsored by doctor science foundation project of Hebei University of Science and Technology.

#### References:

- [1] LI Meng-xin, WU Cheng-dong. Pattern recognition algorithm for license plate character based on multilevel network .Application Research of Computers, 2009,5(26),pp.1703-1705.
- [2] DENG Hong-yao,GUAN Shu-an,SONG Xiu-li. Unite projection and template match to segment license plate characters Computer Engineering and Design, 2008 ,29 (6), pp.1568-1570
- [3] Gu Chenqin,Ge Wancheng. Character Recognition Based on Template Matching Method Communications Technology, 2009, 42 (3) ,pp.220-222
- [4] ZouMingming,LuDi.Recognition algorithm of car license plate characters based on modified template match. Research and exploitation, 20 10,29(1),pp.59-61
- [5] Luo Sanding,Yang Jinping,Li. Improved algorithm of template matching .Computer Engineering and Application, 2010,46(3),pp.169-171
- [6] Li Xingyang,Wang Wenqing.Research of license plate character recognition algorithm based on template matching.Journal of Shanxi Normal University(Natural Science Ediction), 2008, 36(6),pp.176-178