

Research on License Plate Character Recognition Technology

Liyuan Zhang^{1, a, *}, Ruxia Hong^{2, b}

¹Computer Application, Yu Zhang Normal University, Nanchang 330031, China.

²Artificial Intelligence, Yu Zhang Normal University, Nanchang 330031, China.

^{a, *} 592738185@qq.com, ^b545798076@qq.com

Abstract. In recent years, with the rapid development of China's economy and technology, social life has been progressing and urbanization has become more and more large. Many large and medium-sized cities have built highways, large parking lots, and high-end residential areas. At the same time, the rapidly growing number of cars is increasing the burden of cities, and the traffic problem is becoming more and more serious. People have put forward higher requirements for road use and motor vehicle management. In this situation, automatic license plate recognition system is produced. License plate character recognition technology is one of the key technologies in license plate recognition system. It covers many aspects including artificial intelligence, pattern recognition, computer vision and digital image processing and etc. In this paper, a character recognition algorithm based on two-level combination classification is proposed.

Keywords: License plate character recognition, two-level combination classification.

1. Overview

1.1 License Plate Recognition System and the Key Technologies

License plate recognition system refers to a technology that can automatically recognize the license plate of a vehicle and realize intelligent and automatic management of the vehicle in the process of driving without installing other devices and maintaining the original state of the vehicle. It is the computer image recognition technology used in the vehicle license plate in a technical means, the entire system realization process mainly includes: image acquisition, image preprocessing, license plate location, license plate segmentation and license plate recognition and other key technology.

Image acquisition: when the system captures the passing of the vehicle, it adopts the ground induction coil or the illumination detection vehicle to obtain the image information of the vehicle through the camera.

Image preprocessing: when the camera obtains vehicle image information, it needs to conduct image preprocessing, such as enhancement, recovery, transformation and other operations, so that the license plate information can better highlight the characteristics of the license plate after processing.

License plate location: according to the characteristics of the license plate, the license plate is located through scientific technical methods, and then captured the license plate image.

Character segmentation: to segment the characters in the license plate image acquired in the process of license plate location, mainly dividing the Chinese characters, Numbers and letters in the license plate into independent single characters.

Character recognition: according to the preset standard, unified processing of the characters obtained by segmentation and convert them into text for recognition.

The principle of license plate recognition system is as follows:

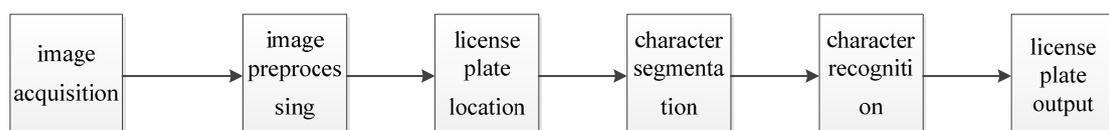


Figure 1. Principle of license plate recognition system

1.2 Character Recognition Technique

Character recognition is a technology that identifies and confirms all characters on the license plate, including Chinese characters, Numbers and letters, it is completed after license plate location and segmentation. License plate recognition technology is a key and complex technology in the whole license plate automatic recognition system. The most commonly used license plate recognition technologies are mainly based on template matching, the identification methods based on neural network and supporting vector machine[1][2].

(1) Based on template matching method: The method is to convert the characters to be recognized into the size of the template, and refer to the template set, and then calculate the matching degree of the input characters and the template set.

(2) Based on neural network method: Neural network simulates the structure of biological neurons and processes information in distributed parallel mode. Neural network method has strong fault tolerance and adaptive ability, but at the same time the algorithm convergence speed is slow, and the algorithm depends on the number of samples to a large extent.

(3) Based on supporting vector machine method: supporting vector machine(SVM) is a new character recognition method after neural network. It is a new machine learning method based on the VC dimension theory of statistical learning theory and the principle of structural risk minimization. The goal is to find an optimal hyperplane in the training sample set.

2. License Plate Recognition Algorithm based on Two-Level Combination Classification

In this paper, a license plate recognition algorithm based on two-level combinatorial classification is proposed. First, the characters are roughly classified. The result of rough classification is that some easily confused characters are tentatively classified as a class, for example, characters (0,D), (1,L,J), (2,Z), (5,S), (8,B) and (E,F) can be regarded as a class. If the character is not the kind of easily confusable, the recognition result is given directly. After the easily confused characters are sorted, it is proceed to the next fine classification. The fine-classifier designs different classification methods according to the local characteristics of these characters. The main task of fine-classification is to design different classification methods according to their local characteristics.

2.1 Rough Classification Method Idea

In this paper, the rough classification method is realized by combining KPCA and SVM. Since SVM algorithm will consume more computing time in problem optimization, KPCA is added to reduce the dimension of character image. The implementation steps are as follows: firstly, KPCA method is used to train the preprocessed images and construct the feature space; then, SVM training samples are projected on the constructed feature space to obtain the feature vector, which is used as the data input for SVM training. After the end of SVM training, the character image to be recognized obtain the feature vector by KPCA method, and finally SVM classifies the characters through the feature vector.

Some characters in the license plate are quite similar. If the whole character image is recognized and then classified, it is not good to distinguish some similar characters. We found that the identification method based on KPCA and SVM has the poor recognition ability to the characters such as (0, D), (1, L, J), (2, Z), (5 S), (8, B), (E, F) etc, and often can't distinguish out these characters. Therefore, this paper puts forward that in the process of coarse classification the similar characters are made training recognition as the same category, such as character 0, 1 character, L and D and J, etc. Rough classification results are: 0 and D are identified as 0; 1, L and J are identified as 1; 2 and Z are identified as 2; 5 and S are identified as 5; 8 and B are identified as 8; E and F are identified as E; and other characters are directly given the classification results. It shows below:

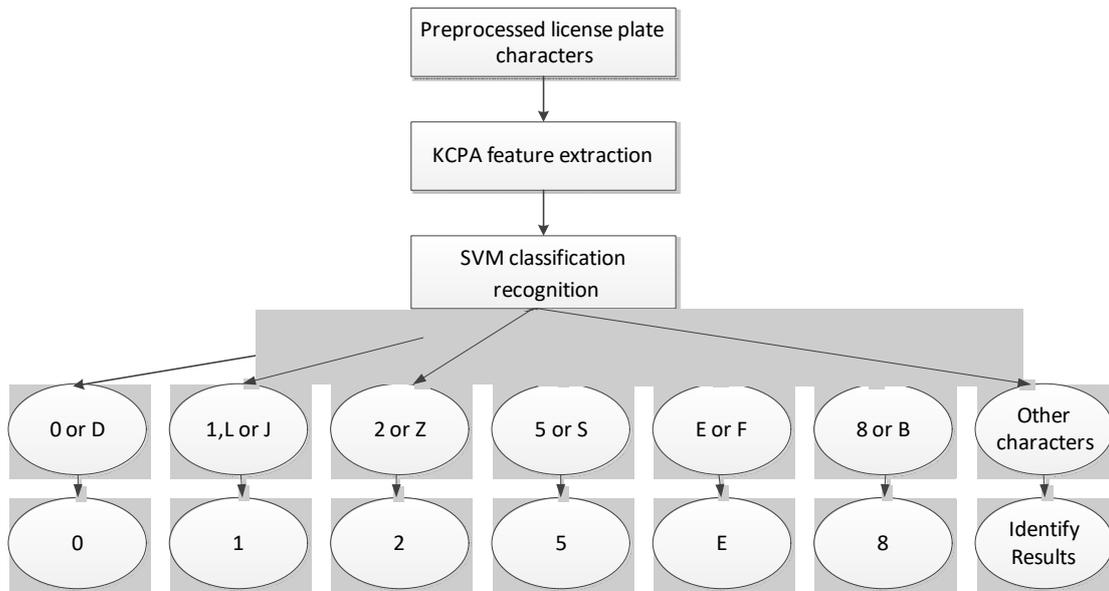


Figure 2. Rough classification result

2.2 Fine Classification Method Idea

After rough classification, confusable characters are classified into one category, and other characters get result directly. Next, confusable characters need to be further classified. The main idea of fine-classification is to design different classification methods for different local features of confusable characters. In this paper, the process of fine-classification are used the skeleton characteristic method and gray level statistical method to distinguish and identify.

(1) Skeleton characteristic

Skeleton is an effective form of shape representation, which is very important for object representation and recognition. It has been applied in image retrieval, computer graphics, font recognition, medical image processing and other fields. As skeleton plays an important role in shape description, many skeleton extraction algorithms[5] are used to measure and represent different shapes.

For example, skeleton characteristic method[3][4] is used to identify 2 and Z as follows:

Step 1: enlarge the preprocessed image by several times, and then use the skeleton characteristic method to extract the image character skeleton;

Step 2: scan the skeleton point from left to right and also top to bottom. The x-coordinate of the first skeleton appearing on the left is denoted as x_0 , and the y-coordinate of the skeleton is denoted as y_0 . The area of the region formed by the skeleton point and the line $y=y_0$, denoted as S_0 ;

Step 3: judge the license plate character according to the size of S_0 area in the previous step, if $S_0 \geq 10$, judge it as character 2; If $S_0 < 10$, the character Z is determined;



Figure 3. Character skeleton: (a) the green curve is the skeleton of 2 and Z;(b) the yellow area is S_0

The confusable characters 5 and S are classified according to the above steps. The difference is that the scanning direction of the step 2 is changed from right to left, from top to bottom.

(2) Grayscale statistical

Grayscale statistical method refers to distinguishing confusable characters by calculating the number of grayscale pixels in a certain feature area of an image. For binary images, only the number of black and white pixels is needed to count. In this paper, after the preprocessing of license plate characters, the value of black and white pixels in a feature area of confusable characters is calculated, and then the proportion of black pixels in the feature area is calculated. Finally, the character is determined and recognized by the given threshold. For example, the recognition steps of characters 0 and D through grayscale statistics are as follows:

Step 1: Set the width and height of character pictures are w and h in turn, scanning the character image: choose the starting point $(0, 2h/3)$ of the image and scanning right, record the position of the first black pixel $(w_0, 2h/3)$, at the same time, choose the starting point for $(w/2, 2h/3)$ to scan images to left, also to record the location of the first black pixels in $(w_1, 2h/3)$;

Step 2: select the starting point of the character image $(w/2, h)$ and scan upward, record the position of the first black pixel $(w/2, h_0)$;

Step 3: select the rectangular region P with height $(h_0 - 2h/3)$ and width $(w_1 - w_0)$, and calculate the proportion R of black pixels in the P region;

Step 4: given a threshold value R_0 , if $R \geq R_0$, the character is D; if $R < R_0$, the character is 0.

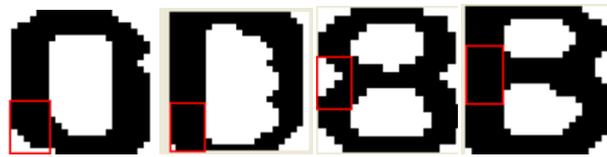


Figure 4. Distinguishing parts of confusable characters, the area inside the red box is the rectangular area P .

8 and character B also apply to grayscale statistics.

Some other methods can be used in fine-classification method, such as boundary distance method, improved threading method, etc. These methods can distinguish some confusing characters, and the effect is better.

The experiment shows that the two-level combinatorial character recognition algorithm of rough classification first and then fine classification improves the recognition rate of license plate recognition and has certain practical value.

3. Summarize

In this paper, a license plate character recognition algorithm based on two-level combinatorial classification is proposed. First, confusable characters are classified into one category by rough classification, which is carried out by combining KPCA and SVM. After rough classification, characters that are not easy to be confused can be recognized directly, while characters that are easy to be confused need to be further classified by fine-classification step. In this paper, skeleton characteristic method and grayscale statistical method are used to identify different confusable characters in step fine-classification. Experimental results show that the algorithm improves the recognition rate of license plate characters and has certain practical value.

Acknowledgments

Fund project: 2017 Science and Technology Research Project of Jiangxi Education Department "Research on Character Recognition Technology in License Plate Recognition System", No. GJJ171194.

References

- [1]. Chih-Wei H, Chih-Jen L. A comparison of methods for multiclass support vector machines[J]. *Neural Networks, IEEE Transactions on*. 2002, 13(2): 415-425.
- [2]. Cristianini N, Shawe-Taylor J. *An Introduction to Support Vector Machines*[Z]. Cambridge Cambridge University Press, 2000.
- [3]. Pedro F. Felzenszwalb, Daniel P. Huttenlocher. Distance Transforms of Sampled Functions[J]. 2004, 23(5): 59-63.
- [4]. Franky Y., Christopher. A skeletonization algorithm by maximal tracking on Euclidean distance transform[J]. *Pattern Recognition*. 1995, 28(5): 331-341.
- [5]. Sebastian, T.B, Klein, P.N, Kimia, B.B. Recognition of shapes by editing their shock graphs[J]. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. 2004, 26(5): 550-571.