Robust Segmentation Method for Doorplate Recognition System[1)]

HONG Yi-Ping YI Jian-Qiang ZHAO Dong-Bin

(*Laboratory of Complex Systems and Intelligence Science, Institute of Automation, Chinese Academy of Sciences, Beijing* 100080)

(E-mail: {yiping.hong, jianqiang.yi, dongbin.zhao}@mail.ia.ac.cn)

Abstract Robot navigation based on character recognition is an effective vision method for compensating the disadvantage of ultrasonic and infrared sensors. A typical example of character recognition for mobile robot navigation is the doorplate recognition system. The captured doorplate images contain unexpected noise from irregular illumination conditions, various imaging angles, different imaging distances, *etc.* The unexpected noise may still exist after segmentation step. In this paper, a robust segmentation method based on speculating the candidates of the characters and feeding back the classification result to the segmentation process is presented. If the candidates of doorplate characters cannot be determined at the segmentation step, a speculation according to known knowledge is executed. The threshold for character extraction from candidates is adjusted when the corresponding character is rejected after classification. The experimental results indicate that the recognition results are effectively improved with the proposed segmentation method.

Key words Robot navigation, doorplate recognition, segmentation

1 Introduction

Generally, a character recognition system consists of the following steps: preprocessing, segmentation, feature extraction, classification, and post processing[1~3]. The input to the character recognition system is usually raw data that cannot be directly recognized. The preprocessing should remove as much as possible noise that may affect latter processes. However, it is difficult for a doorplate recognition system to eliminate noise after preprocessing, and noise may still exist even after segmentation step[1,4,5]. This makes the purpose of the segmentation of a doorplate recognition system different from that of a general character recognition system.

The captured doorplate images often contain unexpected noise from irregular illumination conditions, various imaging angles, different imaging distances, *etc.* It is usual that only a part of doorplate characters can be extracted, with a global threshold. In car license recognition systems[2,6,7], the existing models of doorplate recognition systems almost belong to the sequence structure, in which the segmentation step will directly influence the final recognition result. To overcome these disadvantages, a robust segmentation method based on speculating the character candidates and feeding back the recognition result to the segmentation process is proposed in this paper. Thus, the model of the corresponding doorplate recognition system becomes a feedback structure.

In Section 2, the new model of doorplate recognition system is described. Section 3 is the preprocessing of the doorplate recognition system. The segmentation, feature extraction and classification steps are discussed from Section 4 to Section 6. In Section 7, experiments are conducted to verify the proposed segmentation method. The conclusion is finally derived.

2 New model of doorplate recognition system

In a character recognition system, the final recognition result is greatly affected by the segmentation step. To restrain noise, we construct a new model of doorplate recognition system based on speculating the digit candidates and feeding back the recognition result to the segmentation process (*i.e.*, a robust segmentation method), as shown in Fig. 1. The new model also includes preprocessing,

1) Supported by Hundred Talents Project of Chinese Academy of Sciences
Received December 25, 2003; in revised form June 23, 2004

segmentation, feature extraction, classification, and post processing. However, the actual segmentation step and the system structure are quite different from those of the existing models.

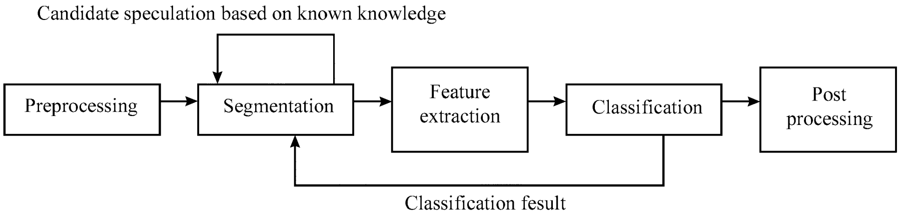

Fig. 1 A new model of a doorplate recognition system based on robust segmentation method

The new model assumes that the output of preprocessing is a parallelogram region only comprising concerned characters and background. After the segmentation process the characters should be normalized to a size-predefined binary character image.

3 Preprocessing

Here, the doorplate is supposed to consist of two regions. The upper region indicates the doorplate number of three digits. The lower region shows the corresponding room name. In the doorplate recognition system we only recognize the doorplate number.

The preprocessing can be categorized into three steps: edge detection, doorplate position determination, and character region extraction. The detail of the former two steps can be found in [8]. The third step is shown in Fig. 2. The region in the outside rectangle is the result of the doorplate position determination. The middle rectangle is the deduced rational region including doorplate numbers according to the known knowledge. The character region only containing concerned digits and background is closed by the inside parallelogram.

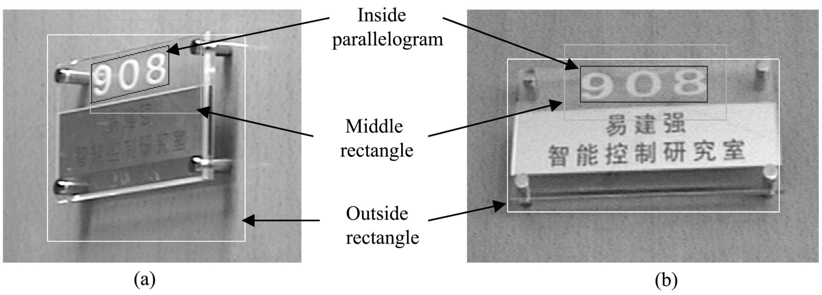

Fig. 2 Extraction of character region

4 Segmentation

4.1 Determination of digit candidates

The first task of segmentation is to select a proper threshold to determine the digit candidates according to the histogram of the character region achieved by the preprocessing. In the character region, the ratio of the pixel number of digits to that of background is uncertain. If the average gray value is selected as a threshold, the result of segmentation will be instable. Here we present an approach to choose the threshold. This approach assumes that the digits with high gray level at least occupy 10% pixels in the whole character region, while the background with low gray level at least have more than 30% pixels in the same region. We label the average gray value of the low 30% pixels in the region as LOWAVERAGE, and label the average gray value of the high 10% pixels as HIGHAVERAGE. Thus, the threshold is selected as follows.

$$Threshold = LOWAVERAGE + k^{*}(HIGHAVEREAGE - LOWAVERAGE) \quad 0 < k < 1 \qquad (1)$$

By default, we select the mean value of LOWAVERAGE and HIGHAVERAGE as the threshold to extract the digits.

After processing with the selected threshold, there are four possible results: 1) No digit candidate is determined; 2) Only one digit candidate is determined; 3) Two digit candidates are determined; 4) three digit candidates are all determined. In the first case, another threshold would be chosen for further processing. In the fourth case, all the digit candidates are found and the recognition-based segmentation in candidates is executed in succession. However, in the other two cases, we should speculate the other digit candidates according to some known knowledge. Examples illustrating the speculation of the candidates are shown in Fig. 3. The above line in Fig. 3 correspond to the second case that only one digit candidate is determined and the other two digit candidates should be speculated. The second line is the third case that two digit candidates are determined and the other one candidate should be speculated.

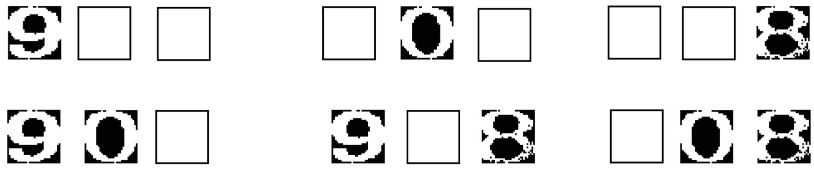

Fig. 3 Examples of digit candidate speculation: the binary images denote the candidates found; the squares denote the speculated candidates according to known knowledge

The known knowledge for speculation refers to the actual doorplate shapes and the font of the characters. In our experiments, the known knowledge mainly includes the following rules.

1) The ratio of width to height of each candidate must be between 0.6 and 1.4;

2) The size of each of the three digit candidates should be proportional to the size of the whole doorplate;

3) Geometrical relations between each candidate must be the same as those of the actual doorplate characters;

4) The position of the digit character region is relatively fixed in the doorplate.

If all the speculated candidates are rational, the next step would be processed, otherwise another threshold is selected to determine the candidates according to (1). To judge the rationality of the speculated candidates, we also depend on the above four known rules to verify them. If a speculated digit candidate satisfies all the known rules, it is considered rational, otherwise it is not rational.

Every speculated candidate should be marked. Meanwhile, a confidence value would be assigned to the final recognition result of the doorplate according to the number of the speculated candidates. If the number of the speculated candidates is large, a low confidence value is given.

The size and shape of characters in images vary with the imaging distance and angle. Usually the affine rectification should be employed before character segmentation[9]. In this paper we have not adopted the affine rectification because of the following two reasons: 1) With illumination disturbance the affine parameters are difficult to estimate precisely. A bad estimation of parameters will introduce some new error for segmentation; 2) The size and shape of the characters in images only change a little when the doorplate slants at a limited extent. However, for compensating the absence of the affine transform, we will extract different character samples slanted at some extent for training the classifier in the recognition step.

4.2 Recognition-based segmentation in candidates

If three digit candidates are rational according to the known knowledge, then the extraction of three digits from the corresponding candidates is executed. Here we also assume that each candidate only contains a digit with high gray level and background with low gray level, and that the selection of the threshold follows equation (1). By default, we choose $k = 0.6$. Fig. 4 shows the results of extracted digits from the same doorplate image when $k = 0.5, 0.6, 0.65$, and 0.8, respectively. The extracted characters in Figs. 4 (b) and (c) are relatively easy to be recognized, but the other cases may be difficult.

(a) k=0.55 (b) k=0.6 (c) k=0.65 (d) k=0.8

Fig. 4 Extracted digits from the same doorplate image when $k = 0.55, 0.6, 0.65$, and 0.8, respectively

However, a default threshold in candidates is not always the best one. A suitable threshold would result in a good classification result. Thus, to find a good threshold becomes how to measure the extraction results of digits in the candidates. We choose the classification result of the extracted digits as a measurement of the digit extraction. If the extracted digit is rejected, another threshold would be selected for further recognition. If the extracted digit is classified to some digit model, the threshold would be viewed as a good one. With the feedback of classification information to the digit segmentation (*i.e.*, a recognition-based segmentation), a better recognition result of doorplate would be achieved. The captured doorplate images are various in size. So for further classifying the extracted digits should be normalized to a standard size. In this paper, we use the standard size with 32*32 pixels.

5 Feature extraction and classification

Under noise conditions, there is almost no feature that is completely insensitive to the changing illumination[10]. To restrict noise, optimizing the performance of every step in doorplate recognition system is needed. A strategy of the recognition-based segmentation can help us get a robust recognition result.

Here we choose the structure features for combining classification. We scan the samples from four sides in vertical and horizontal directions. On each side, a group of features is obtained. There are many different approaches of classifier combination developed[11~12]. In our experiment, four neural network classifiers combined by a fuzzy model are used. Each classifier corresponds to a group of features.

Let $D = \{D_1, D_2, D_3, D_4\}$ be a set of classifiers, and let $\Omega = \{\omega_1, \omega_2, \cdots, \omega_{10}\}$ be a set of class labels. The input to each classifier is a feature vector $\boldsymbol{X} = [x_1, x_2, \cdots, x_n]^{\mathrm{T}}$. Each classifier assigns an input x to a 10-dimentional vector with supports to the classes, i.e.,

$$\boldsymbol{D}_i(\boldsymbol{X}) = [d_{i,1}(\boldsymbol{X}), d_{i,2}(\boldsymbol{X}), \cdots, d_{i,10}(\boldsymbol{X})]^{\mathrm{T}} \tag{2}$$

We restrict $d_{i,j}(\boldsymbol{X}),\ i = 1, 2, 3, 4,\ j = 1, 2, \cdots, 10$, within the interval [0,1] in network training. Thus $d_{i,j}(\boldsymbol{X})$ is the degree of support given by classifier $\boldsymbol{D}_i$ to the hypothesis that $\boldsymbol{X}$ comes from class ω_j. The aim of combining the classifiers is to find a class label for $\boldsymbol{X}$ based on the four classifier outputs $\boldsymbol{D}_1, \boldsymbol{D}_2, \boldsymbol{D}_3, \boldsymbol{D}_4$. Finally, we can find a vector of supporting the classes, denoted as

$$\mu(\boldsymbol{X}) = [\mu_1(\boldsymbol{X}), \mu_2(\boldsymbol{X}), \cdots, \mu_{10}(\boldsymbol{X})]^{\mathrm{T}} \tag{3}$$

where $\mu_t(\boldsymbol{X}), t = 1, 2, \cdots, 10$ is worked out according to the average rule from $d_{i,j}(x)$, *i.e.*,

$$\mu_t(x) = (d_{1,t}(x) + d_{2,t}(x) + d_{3,t}(x) + d_{4,t}(x))/4, \quad t = 1, 2, \cdots, 10 \tag{4}$$

Based on $\mu(\boldsymbol{X})$, we can use the maximum membership rule to assign $\boldsymbol{X}$ to class ω_s or reject:

$$combiningclassifier(\boldsymbol{X}) = \begin{cases} \omega_s, & if\ \mu_s(\boldsymbol{X}) = \max \mu_t(\boldsymbol{X}) > \alpha, \forall t = 1, 2, \cdots, 10 \\ rejection, & else \end{cases} \tag{5}$$

where $combiningclassifier(\boldsymbol{X})$ is the final result of classification. α is a threshold to $\mu_t(\boldsymbol{X})$. If $\mu_s(\boldsymbol{X}) < \alpha$, the input sample x will be rejected.

The output of the four classifiers can be seen as four 10-dimentional fuzzy vectors supporting the classes, so the processing completed by the four classifiers is an obfuscation step. The combining algorithm based on average rule from $d_{i,j}(\boldsymbol{X})$ acts as a fuzzy reasoning process. Finally, defuzzification is achieved by the maximum membership rule.

The selection of threshold α for rejection is conducted by experiments in this paper. If one digit character is rejected, then another threshold for extracting the digit would be chosen from the corresponding candidate.

6 Experiments

In order to test the performance of the segmentation method for the doorplate recognition system, 369 pieces of doorplate images captured from real situations at daytime or night with greatly changing illuminations are employed in experiments. When snapping these images the angle between the camera optical axis and the normal direction of doorplate is approximately limited within 20 degrees. The training of the four neural network classifiers is based on a data set with 400 samples extracted under noise conditions.

We experiment the recognition system on a computer with 2.0GHz CPU and 256M RAM. The doorplate images are 768∗576 pixels in gray level. The calculation time of the whole recognition system is between 50ms and 70ms.

Table 1 to Table 3 show the test results of the doorplate recognition system. The doorplate can be extracted successfully after preprocessing in all 369 images. Table 1 shows that in the 369 tested images the sum of images speculated is 22. In these 22 images there are 12 images that have been recognized correctly. The recognition results have been improved by 3.25%. Table 2 shows the improved results in extracting digits from candidates. The improvement of recognition ratio with 7.86% indicates that the recognition-based segmentation is clearly effective in the doorplate recognition system. Table 3 is the recognition ratio of the doorplate images.

Table 1 Improved results after speculation

Total images speculated	Total images that the recognition results are correct	The ratio of improved images to all test images
22	12	3.25%

Table 2 Improved results after recognition-based segmentation

Total images that the classification result is fed back	Total images correctly recognized	The ratio of improved images to all test images
48	29	7.86%

Table 3 Recognition ratio of doorplate images

Total images correctly recognized	Total images rejected	Total images recognized falsely	The correct recognition ratio
302	21	46	81.84%

7 Conclusion

In this paper, a new model based on robust segmentation method is described first. Then, the algorithm for preprocessing, segmentation, feature extraction, and classification is presented in succession. The experiment results show that the improved recognition ratio after introducing speculation of the digit candidates is 3.25% and a 7.86% improved recognition ratio is achieved by the recognition-based segmentation. The above results indicate that the segmentation method for doorplate recognition system in this paper is robust in improving digit character recognition with unexpected noise.

References

1 Tomono M, Yuta S. Mobile robot navigation in indoor environments using object and character recognition. In: Proceedings of IEEE International Conference on Robotics and Automation, 2000. **1**: 313∼320

2 Naito T, Tsukada T, Yamada K, Kozuka K, Yamamoto S. Robust license-plate recognition method for passing vehicles under outside environment. *IEEE Transactions on Vehicular Technology*, 2000, **49**(6): 2309~2319
3 Steinherz T, Rivlin E, Intrator N. Off-line cursive script word recognition—A survey. *International Journal on Document Analysis and Recognition*, 1999, **2**: 90~110
4 Liu Y, Yamamura T, Tanaka T, Ohnishi N. Character-based mobile robot navigation. In: Proceedings of IEEE International Conference on Intelligent Robots and Systems. Kyongju, Sonth Korea: 1999. **2**: 610~616
5 Meng Q, Rainer B. Rapid door-number recognition by a humanoid mobile robot. In: Proceedings of the 2003 IEEE International Conference on Robotics, Intelligent System and Signal Processing. Changsha, China: 2003. 71~76
6 Cornet B, Rothkrantz L. Recognition of car license plates using a neocognitron type of artificial neural network. *Neural Network World*, 2003, **13**(2): 115~131
7 Hsieh J, Yu S, Chen Y. Morphology-based license plate detection in images of differently illuminated and oriented cars. *Journal of Electronic Imaging*, 2002, **11**(4): 507~516
8 Hong Y, Yi J, Zhao D. An Extraction Method of Doorplate under Various Illumination Conditions. In: Proceedings of the 2nd International Conference on Computational Intelligence, Robotics and Autonomous Systems. Singapore: 2003.
9 Chen X, Yang J, Zhang J, Waibel A. Automatic Detection and Recognition of Signs from Natural Scenes. *IEEE Transactions on Image Processing*, 2004, **13**(1): 87~99
10 Trier O, Jain A, Taxt T. Feature extraction methods for character recognition—A survey. *Pattern Recognition*, 1996, **29**(4): 641~662
11 Xu L, Krzyzak A, Suen C Y. Methods of combining multiple classifiers and their applications to handwriting recognition. *IEEE Transactions on Systems, Man, and Cybernetics*, 1992, **22**(3): 418~435
12 Grim J, Kittler J, Pudil P, Somol P. Multiple Classifier Fusion in Probabilistic Neural Networks. *Pattern Analysis and Applications*, 2002, **5**(2): 221~233

HONG Yi-Ping Received his bachelor and master degrees from Nanchang University in 1997 and Dalian University of Technology in 2002, respectively. Now he is a Ph. D. candidate at Institute of Automation, Chinese Academy of Sciences. His research interests include computer vision and image processing.

YI Jian-Qiang Received his Ph. D. degree in 1992 from Kyushu Institute of Technology, Japan, and currently he is a professor at Institute of Automation, Chinese Academy of Sciences. His research interests include intelligent control, robotics, and mechatronics.

ZHAO Dong-Bin Received his Ph. D. degree from Harbin University of Technology in 2000. He is now an associate professor at Institute of Automation, Chinese Academy of Sciences. His research interests include intelligent control, robotics, and mechatronics.