

# Chinese Traffic Panels Detection and Recognition From Street-Level Images

Yajie Chen<sup>1</sup> and Linlin Huang<sup>1</sup>

<sup>1</sup> School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing, China

**Abstract.** Traffic sign detection and recognition has been the active research topic due to its potential applications in intelligent transportation. However, detection and recognition of traffic panels containing much information, still remains to be a challenging problem. This paper proposes a method to detect and recognize traffic panels from street-level images in the urban scenes and to analyze the information on them. The traffic panels are detected based on histogram of oriented gradient and linear support vector machines. The text strings and symbols on traffic panels are segmented using connected component analysis method. Finally, the symbols on traffic panels are recognized by means of a model named bag of spatial visual words. Experimental results on images from Baidu Panorama Map prove the effectiveness of the proposed method.

## 1 Introduction

In recent years, the research of intelligent vehicle is attached more and more importance. In the related research of intelligent vehicle, traffic sign detection and recognition has been a major topic of discussion over the past decades and its detection rate and recognition rate have been quite high [1].

Nevertheless, the detection and recognition of traffic panels, which contain characters and symbols, is still a remote area. There are two main reasons. On the one hand, compared with most traffic signs, traffic panels have less priority, to which less attention of researchers is paid. On the other hand, there is many kinds of information, like characters and symbols, appeared on traffic panels and as a consequence, recognition of traffic panels is very difficult, flinching many researchers [2].

For all that, traffic panels are still worthy of studying. Firstly, they can provide a large amount of valuable road information for unmanned driving. The detection and recognition of traffic panels can assist intelligent vehicle with the knowledge of traffic conditions and surroundings [3]. Secondly, even though navigation system mounted in vehicle can also provide road information nowadays, they need periodic renewing, leading to the problem of timeliness. But the detection and layout understanding of traffic panels has better timeliness, to which navigation system mounted in vehicle can't be compared, because the processing objects are images captured by camera mounted in vehicle.

The rest of the paper is organized as follows. Section II shows an overview of our proposed system. Section III describes a traffic panel detection method. Section IV explains an implemented approach for localizing text

strings and symbols on traffic panels and presents a symbol recognition algorithm. Section V provides the experimental results. Section VI concludes with discussion of future work.

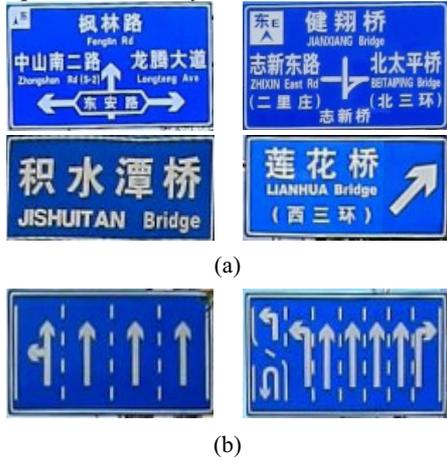
## 2 System Overview

Our system is divided into two parts, as it has been mentioned above.

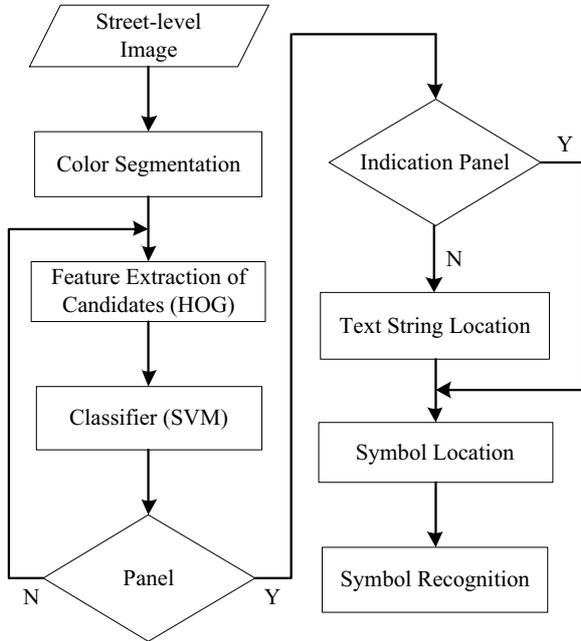
The first one is traffic panel detection. Two kinds of traffic panels with blue background are selected as objects of our study, which are shown in Fig. 1. One is road guide panel, which contains both characters and symbols or only characters and the other is indication panel, which generally contains only symbols. Because traffic panels with blue background are only consider, colour is the most significant feature of traffic panels and candidates can be extracted from input images by colour segmentation. Then the candidate can be decided a road guide panel, an indication panel or neither by using histograms of oriented gradients (HOG) [4] and linear support vector machines (SVM) [5].

The other one is traffic panel recognition, including text string location, symbol location and symbol recognition. Once a traffic panel is detected, image partition of panel region must be done for recognition firstly. In this step, the edge and color information are used for better binarization. Then if it is a road guide panel, text strings can be located by analyzing the alignment of connected components. Symbol location is based on the position of text strings, which are located before and the size of symbol. Finally, symbols detected are recognized by bag of spatial visual words model [6] combined with linear SVM. If it is an indication panel,

we do not need the text string location. Fig. 2 depicts the complete process of our system.



**Figure 1.** Examples of traffic panels. (a) Road guide panels. (b) Indication panels.



**Figure 2.** Complete process of our system.

### 3 Traffic Panel Detection

Traffic panel detection is the first and important step of our proposal, because the following algorithms are applied on traffic panels to increase the efficiency of the system. There are three common methods for traffic panel detection: (1) using edge feature extracted from grayscale image. (2) clustering analysis or extracting intelligent features. (3) combining color and shape information. Our method is developed, based on color segmentation for finding candidates of traffic panel and linear SVM combined with HOG for classification of candidates □ road guide panels, indication panels or neither of both.

An efficient method is applied to segment blue regions as candidates in the images, which has been proposed in [2]. Then, HOG of each candidate is extracted and the classification decision is made by the

linear SVM classifier trained previously. If there is no traffic panel detected by our system in the first time, dilation will be employed on binary images obtained by color segmentation no more than 5 times and classification decision will be made using HOG and linear SVM each time until our system detects at least one traffic panel. If there is still no traffic panel detected, the former step will be repeated by the substitution of erosion for dilation.

### 4 Traffic Panel Recognition

Once there are traffic panels detected in an image, our traffic panel recognition algorithm will be applied only on those areas, including four key steps: image partition to obtain binary image, text string location, symbol location and symbol recognition. These steps will be described in detail as follows.

#### 4.1 Image Partition

To extract Chinese character and symbol from traffic panel, image partition is performed to obtain a binary map of candidate components of character and symbol. Image partition determines the accuracy of Chinese character and symbol detection to some extent. We propose a combination of three independent methods using a logical AND operation in

$$B(x, y) = g_1(x, y) \text{ AND } g_2(x, y) \text{ AND } g_3(x, y) \quad (1)$$

where the first method is a proposal that we are making in this paper but the last two methods have been proposed by other authors .

$g_1(x, y)$  is computed using

$$g_1(x, y) = \text{Otsu\_Block}(T(\text{Gray}(x, y))) \quad (2)$$

$\text{Gray}(x, y)$  is the gray-scale image converted from an color image.  $T$  is an intensity transformation function, which applies piecewise-linear transformation and power-law transformation successively.  $\text{Otsu\_Block}(\theta)$  is an improved Otsu's threshold method [7].

The piecewise-linear transformation used in my paper is given by the expression

$$r_1(x, y) = \begin{cases} 25, & \text{Gray}(x, y) < 25 \\ 135, & \text{Gray}(x, y) > 135 \\ \text{Gray}(x, y), & \text{otherwise} \end{cases} \quad (3)$$

and the power-law transformation is given by the expression

$$r_2(x, y) = 255 \times \left( \frac{r_1(x, y)}{\text{Max} - \text{Min}} \right)^\gamma \quad (4)$$

where Max and Min is the maximum and minimum among  $r_1(x, y)$  computed by (3) respectively.  $\gamma$  is a positive constant and  $\gamma = 0.55$  is the optimum value. To

some extent, the piecewise-linear transformation can reduce the influence of light and the power-law transformation can enhance the contrast of images.

Otsu's threshold method is a global threshold method widely used in image processing, which is sensitive to noise, and  $Otsu\_Block(\theta)$  is proposed, which can be regarded as a combination of global threshold method and local threshold method. The input image, which is  $r_2(x, y)$  in our proposal, is divided into  $n \times n$  regions and each region is binarized by Otsu's threshold method. According to our experiments,  $n=4$  is the optimum value.

$g_2(x, y)$  is computed using

$$g_2(x, y) = Canny(Gray(x, y)) \quad (5)$$

as it is proposed in [8].  $Gray(x, y)$  is the gray-scale image converted from an color image.  $Canny(\theta)$  is a method of gray-scale image edge detection, whose result is 0, if the pixel is regarded as edge, and 255 otherwise. It can solve the problem of adhesion between adjacent characters.

$g_3(x, y)$  is computed using

$$g_3(x, y) = K - Means(255 \times S(x, y)) \quad (6)$$

as it is proposed in [9].  $S(x, y)$  is the saturation component of the image and its value range is  $[0, 1]$ . K-Means clustering algorithm is used to distinguish the foreground, whose color is white and saturation is correspondingly low, and background, whose saturation is high. In my algorithm, two clustering centers are generated after employing K-Means clustering algorithm. The result of  $K - Means(\theta)$  is 255, if the variable is close to the clustering center with low value, and 0 otherwise. It can improve the performance of the first method by color information. Referring to Fig. 3, the advantage of our image partition method can be seen apparently, compared to other methods.



Figure 3. Comparison of image partition algorithm: Otsu, K-means clustering and my proposal.

## 4.2 Text String Location

Once the image partition has been finished, we need to detect potential text strings on the road guide panel detected previously. The flowchart of our text string location algorithm is shown in Fig. 4.

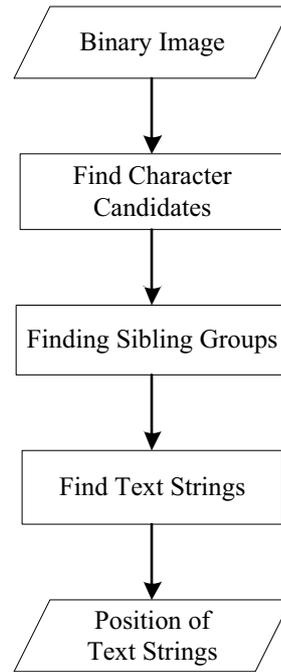


Figure 4. The flowchart of text string location.

The first stage of our algorithm is finding character candidates which are connected components satisfying the constraints (7)  $\square$  (9).

$$5 \leq w \leq 0.5 \times Wid \quad (7)$$

$$0.1 \times Hei \leq h \leq 0.6 \times Hei \quad (8)$$

$$0.5 \leq aspect\ ratio \leq 2 \quad (9)$$

where  $Wid$  and  $Hei$  is the width and height of road guide panel respectively and  $w$  and  $h$  is the width and height of connected component respectively. In (9),  $aspect\ ratio$  is  $w$  divided by  $h$ . The lower limit of  $w$  is independent of  $Wid$ , because it is hard to find their relation, as is shown in Fig. 5.



Figure 5. Comparison of characteres on traffic panels.

In general, characters do not appear isolated but in alignment on road guide panels, namely, each character in a text string must have its siblings at adjacent positions. Thus, the second stage is adjacent character grouping, which is proposed in [10], to extract regions containing text strings. However, it is applied on not only English characters but also Chinese characters in our system and

it is modified based on the arrangement of characters on road guide panels in China. Here, there are four constraints (10) E (13) defined to judge whether two character candidates are siblings.

$$T_1 \leq H/H' \leq 1/T_1 \quad (10)$$

$$D_x \leq T_2 \times \max(W, W') \quad (11)$$

$$D_y \leq T_3 \times \max(H, H') \quad (12)$$

$$T_4 \leq (W \times H)/(W' \times H') \leq 1/T_4 \quad (13)$$

where  $H, W, H', W'$  are height and width of two character candidates  $CC$  and  $CC'$  respectively.  $D_x$  is the difference between x-coordinates of the character candidate centers and  $D_y$  is the difference between y-coordinates of the character candidates centers.  $T_1, T_2, T_3, T_4$  are four thresholds and we set  $T_1 = T_3 = 0.4, T_2 = 3.0$  and  $T_4 = 0.7$ .  $T_2$  is set at 3.0, because it is very common that Chinese characters in a text string are distant, as is shown in Fig. 6.



**Figure 6.** Examples of distant Chinese characters on traffic panels.

These four constraints are applicable to text strings aligned approximately horizontally, while there are many text strings aligned vertically on road guide panels. Therefore, constraints must be modified appropriately in that case.

For two character candidates  $CC$  and  $CC'$ , they can be grouped together if the above four constraints are satisfied and their sibling sets will be updated in accordance with their locations. If  $CC$  is on the left of  $CC'$ ,  $CC'$  will be added into the right-sibling set of  $CC$ , which is added into the left-sibling set of  $CC'$  correspondently. The operation will be reverse when  $CC$  is on the right of  $CC'$ .

For a character candidate  $CC$ , if its left-sibling or right-sibling set is not empty, a sibling group denoted by  $SG$  is obtained, which consists of the character candidate and its sibling sets. Two  $SG$ s are merged together when they intersect and satisfy following constraints.

$$OverlapArea/MaxArea \geq T_5 \quad (14)$$

$$AspectRatio \leq T_6 \quad (15)$$

where  $OverlapArea$  is defined as the area of the overlap between two  $SG$ s,  $MaxArea$  is the maximum of the area of two  $SG$ s and  $AspectRatio$  is the aspect ratio of the new  $SG$  generated after two  $SG$ s are merged together.  $T_5, T_6$  are two thresholds and we set  $T_5 = 0.25, T_6 = 7$ . The merge process will be repeated until there is no new  $SG$  emerged together. In the end, a  $SG$  represents a text string, as is shown in Fig. 7.



**Figure 7.** Correct location of text strings on traffic panels.

### 4.3 Symbol Location

Heretofore, the recognition of traffic panels, on which symbols can also provide important information, only include character detection and recognition except for [2,11]. To improve the performance of our system, symbol location and recognition, which are on both road guide panels and indication panels, are added.

On indication panels, there are only arrow symbols designed to indicate the information of lanes, such as the number of lanes and the type of lanes. Generally, symbols are very eye catching for their size and there are few interference factors. Thus, the size feature is enough to detect symbols on indication panels. A connected component can be determined to be a symbol, if it satisfies following constraints.

$$0.2 \times Hei \leq h \leq 0.75 \times Hei \quad (16)$$

$$5 \leq w \leq 0.2 \times Wid \quad (17)$$

$$aspect\ ratio \leq 1 \quad (18)$$

where  $Wid$  and  $Hei$  is the width and height of indication panel respectively and  $w$  and  $h$  is the width and height of connected component respectively. In (18),  $aspect\ ratio$  is  $w$  divided by  $h$ .

On road guide panels, there are more kinds of symbols and symbol location is harder, compared to that on indication panels. There are two reason why symbol detection is harder on road guide panels. One is the interference of characters which may increase its false-positive rate, and another one is the uncertainty of symbol's position and size, which increase its difficulty. However, text string location done previously is helpful for symbol detection, because there are only symbols except characters on road guide panels, and the position of symbol has relation to the aspect ratio of road guide panel to some extent, therefore, we locate symbols using different method according to different aspect ratio of road guide panel denoted by  $ARP$ , which is the width of road guide panel divided by the height of road guide panel, combining the position of text strings. In our proposal, three instances are taken into account, according to  $ARP$  set as follows:  $ARP \geq 2.3, 0.4 < ARP < 2.3$  and  $ARP \leq 0.4$ .

Case 1:  $ARP \geq 2.3$

Under this condition, there may be only arrow symbols on the left or right of text strings. At first,

candidates must be found through following three constraints.

1) The height of candidate is greater than the average height of text strings.

2) Candidates do not intersect with any text string.

3) The range of the aspect ratio of Candidate is [0.5,2].

Then, if text strings are on the left side of road guide panels, there may be candidates only on the right of text strings. The possibility will be reverse when text strings are on the right side of road guide panels. Some examples of the correct result under this condition are shown in Fig. 8.



Figure 8. Correct symbol detection when  $ARP \geq 2.3$ .

Case 2:  $0.4 < ARP < 2.3$

Symbol detection is the most complex under this condition, because symbols of different sizes can be everywhere. Even though, as is mentioned above, text strings can be helpful, it is inevitable that some characters are left outside text strings located previously. Meanwhile, the constraints of candidates must also be revised as follows.

1) Candidates can intersect with text strings, but the area of the overlap between candidates and any text string should not be greater than  $T_7$  times the area of candidates.

2) To eliminate some interference, the duty-cycle of candidate should not be greater than  $T_8$ .

3) As symbols vary largely, the aspect ratio of candidate falls between  $1/T_9$  and  $T_9$ .

In our system, we set  $T_7 = 0.35$ ,  $T_8 = 0.8$  and  $T_9 = 10$ .

Then, most of candidates which are not symbols can be eliminated by following steps.

1) Only Four biggest candidates are retained if the number of candidates are greater than four. Otherwise, all candidates are retained.

2) The aim of this step is to check whether there are multiple symbols, which have almost the same y-coordinate of center and height, on road guide panels. If there exist candidates whose difference between y-coordinates of centers and difference between height of candidates are less than  $T_{10}$ , other candidates will be eliminated.  $T_{10}$  is set at 4. If such candidates do not exist, the third step will be carried out. Otherwise, symbol location is finished.

3) In most cases, size is the most distinctive feature of symbol. If the area of the biggest candidates is greater than  $T_{11}$  times the area of other smaller candidates, the smaller one will be eliminated.  $T_{11}$  is set at 2.5. If there are at least two candidates left, the fourth step will be carried out. Otherwise, symbol location is finished.

4) Position can also be helpful for the detection of symbol. If x-coordinate of a candidate center falls between  $T_{12}$  and  $T_{13}$  times the width of road guide panel, other candidates will be eliminated.  $T_{12}$  and  $T_{13}$  is set at 0.4 and 0.6 respectively.

Some examples of the correct result under this condition is shown in Fig. 9.



Figure 9. Correct symbol detection when  $0.4 < ARP < 2.3$ .

Case 3:  $ARP \leq 0.4$

Symbol location is similar to that under the first condition, because there may be only arrow symbols above or below text strings. The constraints of candidate selection are the same but the first constraint, which is revised as follows: the height of connected component is larger than the average height of text strings when text strings are aligned horizontally, or the width of connected component is larger than the average width of text strings when text strings are aligned vertically. If text strings are on the upper side of road guide panels, there may be some candidates only below text strings. The possibility will be reverse when text strings are on the lower side of road guide panels. Some examples of the correct result under this condition is shown in Fig. 10.



Figure 10. Correct symbol detection when  $ARP \leq 0.4$ .

#### 4.4 Symbol Recognition

As the performance of optical character recognition (OCR) is good enough at present, symbol recognition is the focus in our proposals. In my opinion, symbol recognition on traffic panels is similar to traffic sign recognition and there should be some methods used for traffic sign recognition for reference. Finally, we choose bag of visual words (BOVW) model [12].

A two-dimensional image can be mapped to a set of visual keywords by BOVW. It not only preserves the local features of an image but also compresses its description. However, the spatial relations between visual words are neglected in traditional BOVW. Because scale invariant feature transform (SIFT) [13] is the most common used feature in BOVW for its strong robustness against rotation, scale variation and light condition, some arrow symbols and intersection symbols shown are hard to distinguish. If BOVW is adopted without any adjustment, the result of symbol recognition will be non-ideal. Thus, BOSVW model is adopted in my proposal.

Actually, bag of spatial visual words model is an extension of traditional bag of visual words model, with the integration of spatial information. As same as traditional one, the most important step is the

establishment of dictionary of spatial visual words, which can be done as follows.

1) SIFT descriptors of input images for training after size normalization are extracted.

2) Images are divided into  $n \times n$  regions and SIFT descriptors in the same region are clustered to obtain visual keywords.

3) Dictionary of spatial visual words can be established on the basis of dictionary of visual words in each region in a given order.

BOVW model owns the spatial information by this means. Each image can be represented by the histogram of visual words in dictionary, which counts the occurrence probability of the visual words in the image.

The categories of input images for training are learned by a classifier, which is linear SVM used in our paper.

In test stage, visual keywords of an input image are extracted and a BOSVW histogram is calculated to represent the image. Finally, the classification decision is made by linear SVM trained previously.

### 5 Experimental Results

The proposed algorithm must be evaluated through experiments to prove its effectiveness. The data sets used in this paper have been obtained from the Panorama Map service developed by Baidu. A total of 1578 images are extracted, which are divided into two independent subsets, one for training composed 1078 images and the other for testing composed of 500 images. In order to prove the diversity of our data set and the robustness of our proposal, all the images are obtained from four metropolitans in China, which are Beijing, Chengdu, Shanghai and Shenzhen.

Test images, with size of  $480 \times 360$ , contains 580 panels, 1708 text strings and 669 symbols. Referring to GB5768E2009 implemented in China and actual situation, 27 common symbols are chosen for recognition. The chosen symbols are shown in Table 1.

Our experiments platform is a 2.66GHz processor personal computer. There are three parts of experiment, including traffic panel detection, text string location and symbol location and recognition.

The results are shown in terms of recall or precision in our experiments. They are defined as in (19) and (20).

$$Recall = \frac{TP}{TP + FN} \tag{19}$$

$$Precision = \frac{TP}{TP + FP} \tag{20}$$

$TP$  is the number of true positives,  $FN$  is the number of false negatives and  $FP$  is the number of false positives.

The results are shown in terms of accuracy in experiments of symbol recognition. It is defined as in (21).

$$Accuracy = \frac{CTP}{TP} \tag{21}$$

$CTP$  is the number of true positives, of which are recognized correctly.

Table 1. Chosen Symbols

Categories	Sequences	Symbols
Arrow	1-13	
Intersection	14-27	

#### 5.1 Traffic Panel Detection

Table 2 shows that our method of traffic panel detection is effective and robust, even though the background in urban is very complex. The main factors related to *Recall* and *Precision* include light condition, occlusion and interference factors that are easily misdiagnosed as traffic panels, such as billboards, trucks.

Table 2. Results of Traffic Panel Detection

Algorithm	Panel Number	Precision	Recall	Time
Wen's	92	92.40%	82.10%	—
Proposed	580	98.96%	98.45%	67ms

#### 5.2 Text String Location

At a long distance, characters are too small and illegible to detect. Thus, only images captured at a short or medium distance are considered as well as symbol location and recognition. Here, short distance means the traffic panel is less than 20 meters far, medium distance means it is in the range 20-40 meters far and long distance means it is more than 40 meters far.

There are both Chinese and English text strings in most traffic panels in China. However, English text strings are only designed for foreigners and have the same meaning with respective Chinese ones. Therefore, only the result of Chinese text strings are computed in our study. It can be seen from Table 3 that our method of

Chinese text string location is satisfactory at a short or medium distance.

Light condition also has a great impact on image partition, which is of significant importance to character detection. In addition, there are some unconventional panels where English characters have almost the same even bigger size compared to respective Chinese characters. To some degree, this factor influences Recall and Precision of Chinese character detection.

**Table 3.** Results of Text String Location

Distance	Text String Number	TP	Recall
Short	1279	1171	91.56%
Medium	412	388	94.17%

### 5.3 Symbol Location and Recognition

As there are few papers having tried symbol detection, it is hard to carry a result contrast. However, it can be concluded that our method of symbol detection is quite effective by experiment from Table 4. Like former two parts, light condition is also an unavoidable challenge in symbol detection. Moreover, some symbols are easily erroneously assumed to be characters and it results in the misdetection of symbol.

As for symbol recognition, Table 5 can prove that our method is available. As the distance varies, the accuracy of symbol recognition remains quite high. In my opinion, the main reason of errors is that some symbols are similar, as is shown in Table 1.

**Table 4.** Results of Symbol Location

Distance	Symbol Number	TP	FP	Precision	Recall
Short	437	409	8	98.08%	93.59%
Medium	155	146	3	97.99%	94.19%

**Table 5.** Results of Symbol Recognition

Distance	TP	CTP	Accuracy
Short	409	372	90.95%
Medium	146	136	93.15%

## 6 Conclusions and Future Works

In this paper, we represent a robust proposal of traffic panel detection and recognition, which few researchers tried before. Even though it is hard, it is valuable and worthwhile to try. It is very useful for driver assistance systems in the future.

Although experiment results are quite good, there exist some problems to be settled. Firstly, our data set has limitation. All images are captured during the day time. However, our method should also work at nighttime. Secondly, our data set is static image. To prove the effectiveness of our proposal more powerfully, our data set should be video. Finally, our system could not meet the requirement of real-time at present, if it is applied to

driver assistance. We should continue to improve our algorithm in our future works.

## Acknowledgements

This work has been supported by the National Natural Science Foundation of China (NSFC) Grants 61271306.

## References

1. M. Gu, Z. Cai and X. Ren, "Detection algorithm of traffic sign with character in nature scenes," *Journal of Central South University*, vol. 44, no. 5, pp. 1861-1868, May. 2013.
2. A. Gonzalez, L. M. Bergasa and J. J. Yebes, "Text detection and recognition on traffic panels from street-level imagery using visual appearance," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 1, pp. 228-238, Feb. 2014.
3. W. Wu, X. Chen and J. Yang, "Detection of Text on Road Signs from Video," *IEEE Transactions on Intelligent Transportation Systems*, vol. 6, no. 4, pp. 378-390, Dec. 2005.
4. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005, pp. 886-893.
5. C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273-297, 1995.
6. Y. Wang, H. Guo and C. He, et.al, "Bag of spatial visual words for scene classification," *Computer Science*, vol. 38, no. 8, pp.265-268, Aug. 2011.
7. N. Otsu, "A threshold selection method from gray-level histogram," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, no. 1, pp. 62-66, 1979.
8. J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679-698, 1986.
9. J. Liu, L. Huang, B. Niu, "Road sign text detection from natural scenes," *International Conference on Information Science, Electronics & Electrical Engineering*, 2014, 3: pp. 1547-1551.
10. C. Yi and Y. Tian, "Text string detection from natural scenes by structure-based partition and grouping," *IEEE Transactions on Image Processing*, vol. 20, no. 9, pp. 2594-2605, Sep. 2011.
11. A. Gonzalez, L. M. Bergasa and M. Gavilan, et al, "Automatic information extraction of traffic panels based on computer vision," *12th International IEEE Conference on Intelligent Transportation Systems*, 2009, pp.184-189.
12. G. Csurka, C. R. Dance, L. Fan, J. Willamowski and C. Bray, "Visual categorization with bags of keypoints," in *Proc. Workshop Stat. Learn. Comput. Vis. ECCV*, 2004, pp.1-22.
13. D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. ICCV*, 1999, pp.1150-1157.