

# AUTOMATIC RECOGNITION OF TRAFFIC SIGNS IN NATURAL SCENE IMAGE BASED ON CENTRAL PROJECTION TRANSFORMATION

Ka Zhang<sup>a</sup>, Yehua Sheng<sup>a</sup>, Peifang Wang<sup>b</sup>, Lin Luo<sup>c</sup>, Chun Ye<sup>a</sup>, Zhijun Gong<sup>d</sup>

<sup>a</sup> Key Laboratory of Virtual Geographic Environment, MOE, Nanjing Normal University, No. 1, Wenyuan Road, Nanjing, P. R. China, 210046, - zhangka81@126.com, -(shengyehua, yechun)@njnu.edu.cn

<sup>b</sup> Liuji Secondary School of Wuhe County, Liuji Town, Wuhe County, Anhui Province, P. R. China, 233300, - wangpfbwh@126.com

<sup>c</sup> Wuxi City Communication Engineering Ltd., No. 888, Taihu East Avenue, Wuxi, Jiangsu Province, P. R. China, 214026, -luolin8398@163.com

<sup>d</sup> Nanjing Lear Xindi Automotive Interiors System Co, Ltd., No. 376, Hebanacun, Maigaoqiao, Nanjing, P. R. China, 210028, - gzj4573@126.com

**KEY WORDS:** Vision sciences, Photogrammetry, Recognition, Image, Transformation, Neural, Networks

## ABSTRACT:

Considering the problem of automatic traffic signs recognition in natural scene image (mainly including three kinds of traffic signs: yellow warning signs, red prohibition signs and blue mandatory signs), a new method for traffic signs recognition based on central projection transformation is proposed in this paper. In this method, self-adaptive image segmentation is firstly used to extract binary inner images of detected traffic signs after they are detected from natural scene images. Secondly, one-dimensional feature vectors of inner images are computed by central projection transformation. Lastly, these vectors are input to the trained probabilistic neural networks (PNN) for exact classification, the output of PNN is final recognition result. The new method is applied to 221 natural scene images taken by the vehicle-borne mobile photogrammetry system in Nanjing at different time. Experimental results show a recognition rate of over 98%. Especially, the problem of confirming optimal projection number in central projection transformation is solved by the information entropy in this paper. Moreover, the proposed recognition method is compared with other recognition methods based on three kinds of invariant moments. Results of contrastive experiments also show that the method proposed in this paper is effective and reliable.

## 1. INTRODUCTION

With the development of society and economy in recent years, the problem of traffic safety and traffic jam is becoming more and more serious. Thus the development and application of intelligent transportation system (ITS) have been attached great importance by governments and academia. Traffic sign automatic recognition system, which is an important sub-system in intelligent transportation system, has become one of the hot spots in the field of ITS, and it's also a very difficult problem in real scene image recognition. How to detect traffic signs from complex real scene images rapidly and efficiently is the key step for traffic sign automatic recognition. Colour and shape features are usually used to detect traffic signs. The method of edge extraction and shape analysis is a common one for traffic sign detection (Loy, et al, 2004; Alefs, et al, 2007). In addition, the method based on combination of colour and shape features is another common one for traffic sign detection. In this kind of method, colour segmentation is firstly used to eliminate background-objects and obtain binary image. Then shape analysis is used to detect traffic sign regions in the segmented binary image (Gao, et al, 2006; Filipe, et al, 2007; Zhang, et al, 2007).

Once candidate regions of traffic signs are detected from the image by colour segmentation or edge detection, traffic signs can be recognized according to their shape features. Hsu firstly used template images in database to train matching pursuit (MP) filters to find a set of best MP filter bases of each road sign, then used trained MP filter bases to find the best match between

detected road sign and template sign generated by MP filter bases (Hsu, et al, 2001). Perez firstly carried out Fourier transformation to input image and template image, then computed similarity measure between them by nonlinear correlator for road sign recognition (Perez, et al, 2002). Escalera carried out normalization processing and binary processing on detected road signs first, then input binary image to the adaptive resonance theory paradigm neural network for road sign recognition (Escalera, et al, 2003). Fang presented an automatic road sign detection and recognition system based on a computational model of human visual recognition processing (Fang, et al, 2004). Paclik proposed a novel concept of a trainable similarity measure, which can alleviate some shortcomings of traditional cross correlation similarity. He used this new similarity measure to recognize prohibition signs (Paclik, et al, 2006). Maldonado-Bascon used DtB, which is distance vector composed of distances between edges of road signs and borders of sign candidate regions, as the shape feature of road sign, then input the feature to the trained support vector machines for road sign recognition (Maldonado-Bascon, et al, 2007). Apart from the above methods, many methods based on template matching, neural networks and genetic algorithm are used for traffic sign recognition (Miura, et al, 2000; Liu, et al, 2002; Gil-Jimenez, et al, 2007; Kuo, et al, 2007; etc.).

However, the main shortcoming of above methods is that these methods need the same size between detected real sign and reference sign. Detected sign regions must be normalized to the uniform size, which will damage the original information of detected signs and influence the recognition rate. In this paper,

we present a new shape feature computation method for traffic sign recognition based on central projection transformation. Extracted shape feature can reflect the global feature of traffic sign and stay invariant to object scales and rotations. Research object in this paper mainly includes three kinds of traffic signs: yellow warning signs, red prohibition signs and blue mandatory signs. Experimental results show that author's method has a higher recognition rate for traffic sign recognition.

## 2. DESCRIPTION OF THE METHOD

After traffic signs are detected from natural scene image by the method based on combination of colour and shape features (Zhang, et al., 2007), Self-adaptive image segmentation is firstly used to extract binary inner images of detected traffic signs. Then one-dimensional feature vectors of inner images are computed by central projection transformation. Lastly, for each detected traffic sign, its feature vector is input to the trained probabilistic neural networks (PNN) for exact classification, the output of PNN is final recognition results.

### 2.1 Extract binary inner image of traffic sign

Each type of traffic signs consists of special outline and inner image with specific pattern. If the inner image of traffic sign is effectively extracted, it can provide a stable basis for traffic sign recognition. Within detected traffic sign region in natural scene image, the binary inner image of traffic sign *BImg* can be effectively extracted by self-adaptive image segmentation. The formulas of image segmentation for extraction of binary inner image are described as follows.

$$B\_Img(i, j) = \begin{cases} 0, & \text{if } (R(i + Hh, j + Lh) < T) \\ 1, & \text{if } (R(i + Hh, j + Lh) > T) \end{cases} \quad (1)$$

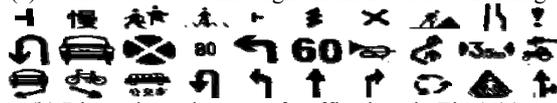
$$T = a \times pjRhd + b \times MinRhd$$

Where, *Hh*, *Lh* respectively represents row and column coordinates of top and left corner of the traffic sign region. *R(i+Hh, j+Lh)* shows pixel's R channel gray value on location (*i+Hh, j+Lh*) of original natural scene image. *pjRhd*, *MinRhd* respectively represents average value and minimal value of all pixels' R channel gray value in traffic sign region. Coefficient *a*, *b* are definite value in the interval [0, 1].

Figure 1 shows some examples of extracted binary inner image of traffic sign from different natural scene images. Figure 1(a) shows some detected traffic signs from natural scene images, Figure 1(b) shows corresponding binary inner images of traffic signs.



(a) Some detected traffic signs from natural scene images



(b) Binary inner images of traffic signs in Fig.1 (a)

Figure 1. Examples of traffic signs and their inner images

### 2.2 Compute feature vector of inner image

In pattern recognition, features are used to distinguish one pattern from the other. Feature extraction is the key step for

pattern recognition. In this paper, we use central projection transformation (Tao, et al., 2001) to compute feature vector for inner image of traffic sign. Through feature computation, we can transform two-dimensional image to one-dimensional vector. For the binary inner image *BImg* of traffic sign,  $m(x_0, y_0)$  is its centroid, and *M* denotes the maximal distance between each pixel of the inner image and its centroid. Then central projection transformation on inner image can be computed as following equations.

$$f(\theta_k) = \sum_{\gamma=0}^M p(\gamma \cos \theta_k, \gamma \sin \theta_k) \quad (2)$$

$$M = \text{Max} \|B\_Img(x, y) - m(x_0, y_0)\|$$

Where,  $\|B\_Img(x,y)-m(x_0,y_0)\|$  represents the Euclidean distance between any point and the centroid of inner image.  $\theta_k = k \times (2\pi/N) \in [0, 2\pi]$ ,  $k=0, 1, 2, \dots, N$ , *N* is the number of projection rays during central projection transformation.  $p(\gamma \cos \theta, \gamma \sin \theta)$  represents the gray value of the pixel at coordinates ( $\gamma \cos \theta, \gamma \sin \theta$ ) in the Cartesian coordinates frame.

The central projection vector ( $f(\theta_1), f(\theta_2), \dots, f(\theta_N)$ ) is the feature vector used for pattern recognition. In practical application, each element in feature vector ( $f(\theta_1), f(\theta_2), \dots, f(\theta_N)$ ) should be normalized by the length of vector. Thus the feature vector used to recognize traffic signs is ( $f^*(\theta_1), f^*(\theta_2), \dots, f^*(\theta_N)$ ) in this paper.

$$f^*(\theta_k) = f(\theta_k) / \sqrt{\sum_{k=0}^N (f(\theta_k) \times f(\theta_k))} \quad (3)$$

From the definition of central projection vector, we can see that the number of *N* will influence the quality of central projection vector. If the number of *N* is too small, a lot of pixels in the binary inner image can't be projected, which will lead to insufficient statistical information for traffic sign recognition. Otherwise, too large *N* will lead to complex computation and a lot of computing time. Thus the optimal projection number can acquire balance between projection quality and computing time. But how to confirm the optimal projection number can't be solved from theoretical aspect in previous research. In this paper, we use the theory of information entropy to solve the problem of confirming optimal projection number in central projection transformation.

Our main thought of confirming optimal projection number is that the information entropy of central projection vector will increase with the increment of projection number *N*. But after projection number *N* increases to a certain large number, the ratio of two neighbouring information entropies should gradually reach a constant. When the ratio of two information entropies approximately reaches a constant, number *N* is considered as the optimal projection number. The steps of confirming optimal projection number are described as follows. (1) Select three standard traffic signs in national standard as experimental data (Shown in Table 1).

Type	Binary inner image	Meaning of sign
Warning sign		Caution pedestrian crossing
Prohibition sign		No honking
Mandatory sign		Turn left and right

Table 1. Experimental data for confirming optimal number *N*

(2) Provide  $N$  with following 103 values: 1, 2, 4, 10, 20, 30, ..., 100, 110, ..., 900, 910, ..., 1000. For three kinds of traffic signs in table 1, compute each central projection vector corresponding to each  $N$  with equation (2) and equation (3). Based on computation of central projection vector, compute its information entropy with following equation.

$$S(N_i) = -\sum_{k=0}^N f'(\theta_k) \log f'(\theta_k), i=1,2,\dots,103 \quad (4)$$

(3) Compute ratio  $\Delta S$  of two neighbouring information entropies with following equation.

$$\Delta S = S(N_{i+1})/S(N_i) \quad (i=1,2,3,\dots,102) \quad (5)$$

(4) For three kinds of traffic signs, the statistical relation between the ratio  $\Delta S$  and central projection number  $N$  is shown in Figure 2. From Figure 2, we can see that ratio  $\Delta S$  gradually reaches constant 1 when  $N$  is 360. Therefore, the central projection number is 360 during central projection transformation in this paper.

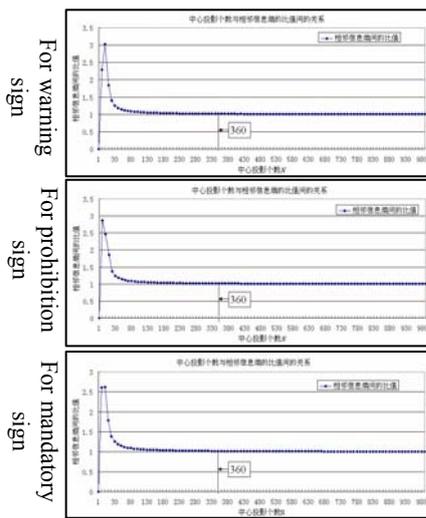
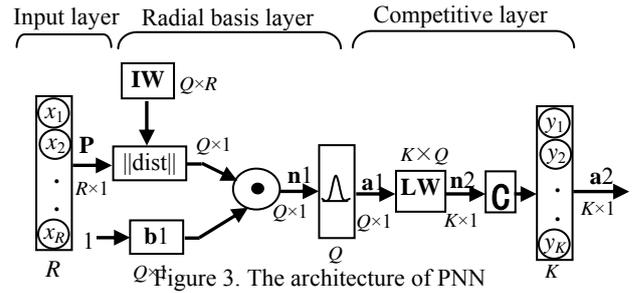


Figure 2. The statistical relation between the ratio  $\Delta S$  and  $N$

### 2.3 Recognize traffic sign through PNN

For each detected traffic sign, after its feature vector is computed by the central projection transformation, we can use the trained probabilistic neural networks (PNN) to recognize its exact meaning. The PNN was firstly proposed by Specht in 1990 (Specht, 1990). It is a kind of forward-propagation neural network developed from the radial basis function neural network. Its characteristics include less training time, fixed architecture and generation output with Bayes posterior probability. Therefore, it has powerful capability of nonlinear recognition. The PNN has three layers: the first layer is input layer, the second layer is radial basis layer and the third layer is output layer. The number of neurons in radial basis layer is equal to that of input samples. The number of neurons in output layer is equal to that of classifications of training samples. Neurons in radial basis layer have thresholds. The architecture of PNN is shown in Figure 3.



In Figure 3,  $R$  shows the number of elements in input vector.  $Q$  shows the number of input/target pairs.  $K$  shows the number of neurons in output layer. The function of module  $C$  is to find the maximal value in elements of its input vector  $\mathbf{n}2$ , produce output value 1 of neuron corresponding to the largest element of  $\mathbf{n}2$ , and produce output value 0 of other neurons. In Chinese standard on traffic signs, the numbers of yellow warning signs, red prohibition signs and blue mandatory signs are respectively 45, 40 and 29. However, actual kinds of limited speed signs, limited height signs and limited width signs are more than one. Therefore, plus actual limited speed signs, limited height signs and limited width signs, the number of prohibition signs is 53 in this paper. Training set is obtained from two sources: inner images of traffic signs in national standard and actual inner images from detected traffic signs in natural scene images. In this paper, for yellow warning signs, red prohibition signs and blue mandatory signs,  $R$  in PNN classifiers are 360.  $K$  in three classifiers are respectively 45, 53 and 29.  $Q$  in training set for three classifiers are 61, 110 and 43 respectively.

The process from input to output of PNN is described as follows.

- (1) Input layer receives input vector  $\mathbf{P}_j = (P_1, P_2, \dots, P_R)^T$ ,  $j=1, 2, \dots, Q$ .
- (2) For the  $i^{\text{th}}$  ( $i=1, 2, \dots, Q$ ) neuron in radial basis layer, the distance  $d_i$  between input vector  $\mathbf{P}_j$  and weight matrix  $\mathbf{I}W_i$  is firstly computed by following equation.

$$d_i = \text{sqrt}((\mathbf{P}_j - \mathbf{I}W_i^T)^T (\mathbf{P}_j - \mathbf{I}W_i^T)) \quad (6)$$

Then according to threshold  $\mathbf{b}1$ , the output of  $i^{\text{th}}$  neuron  $\mathbf{a}_i1$  is computed by following equation.

$$\begin{cases} \mathbf{a}_i1 = e^{-n_i1 \times n_i1} \\ n_i1 = d_i \times \mathbf{b}_i1 \end{cases} \quad (7)$$

- (3) The  $m^{\text{th}}$  element ( $m=1, 2, \dots, K$ ) in input vector  $\mathbf{n}2$  of compete layer is computed as follows.

$$\mathbf{n}_m2 = \mathbf{L}W_m \mathbf{a}1 \quad (8)$$

- (4) The final output vector  $\mathbf{a}2_j$  is computed as follows. Only two kinds of value (0&1) are found in the output vector, and serial number of element 1 is that of classification of input data.

$$\mathbf{a}2_j = \text{compet}(\mathbf{n}2), \mathbf{a}2_j^m = \begin{cases} 1, m = m^* \\ 0, m \neq m^* \end{cases} \quad (9)$$

Where,  $m^*$  is the serial number of maximal value in vector  $\mathbf{a}2_j$ .

### 3. EXPERIMENTS AND CONCLUSIONS

To verify the validity of author’s method, it is realized by Visual C#.Net and Matlab program language. Many experiments are also completed with natural scene images in Nanjing. These images, with size of 1392×1040 pixels, are taken by the vehicle-borne mobile photogrammetry system at different time and in different lighting conditions. The natural scene image taken by its CCD camera is shown in Figure 4.



Figure 4. The experimental data-natural scene image

Traffic signs detected from image of Figure 4 are shown in Figure 5(a), corresponding binary inner images of these traffic signs are shown in Figure 5(b), and recognition results of these traffic signs are shown in Figure 5(c).

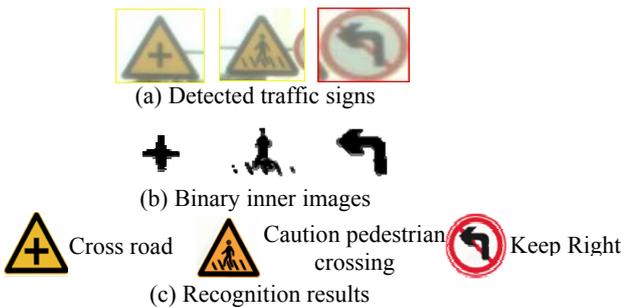


Figure 5. Traffic sign recognition results in image of Figure 4

Experimental result shows that author’s traffic sign recognition method obtains good effect. The run time of author’s method is about 0.4 second under the condition of serial compiling. If the special image processing unit and technique of parallel compiling are used, the recognition speed will be faster.

Besides the above experimental results, the paper totally selects 221 natural scene images taken by the vehicle-borne system at different time and different locations to test the proposed method. There are totally 500 different kinds of traffic signs in these images. The number of detected signs is 480. To compare with other recognition methods based on invariant moments, the paper selects such three kinds of invariant moment as Hu

moment (Hu, 1962), Tchebichef moment (Li, et al., 2006), and Zernike moment (Fleyeh, et al., 2007). Compared statistical results are shown in table 2. In this table, H.M, T.M and Z.M respectively represents Hu moment vector, Tchebichef moment vector and Zernike moment vector.

feature vector		Author’s vector	H.M	T.M	Z.M
yellow warning signs(105)	recognized rate %	105 100	25 23.8	54 51.4	31 29.5
red prohibition signs(221)	recognized rate %	217 98.2	65 29.4	98 44.3	68 30.8
blue mandatory signs(154)	recognized rate %	154 100	69 44.8	103 66.9	86 55.8

Table 2. Compared statistical results between author’s central projection vector and invariant moments

Some compared experimental results between author’s recognition method based on central projection vector and other methods based on invariant moments are shown in Figure 6. From left to right, Figure 6(a) shows natural scene images, detected traffic signs and their binary inner images. Figure 6(b) shows recognition results based on Hu moment. Figure 6(c) shows recognition results based on Tchebichef moment. Figure 6(d) shows recognition results based on Zernike moment. Figure 6(e) shows recognition results based on author’s central projection vector.

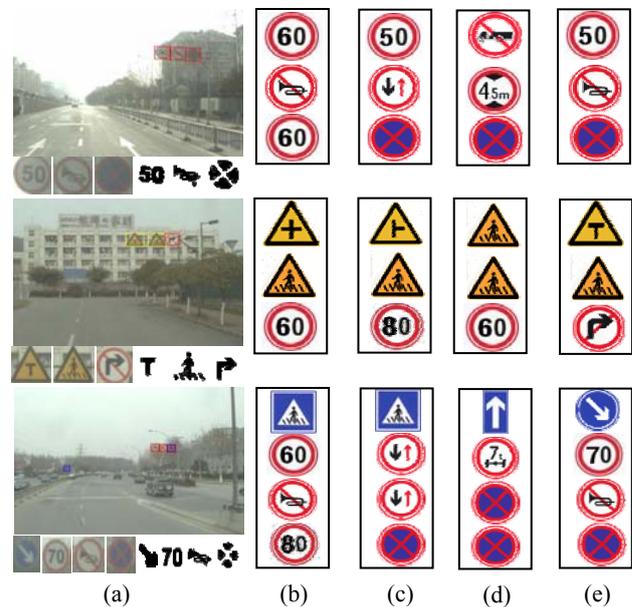


Figure 6. Compared experimental results between author’s recognition method based on central projection vector and other methods based on invariant moments

From experimental results, we can see that the recognition rate of proposed traffic sign recognition method is over 98%. This recognition rate is higher than that of other methods based on invariant moments, which shows that central projection transformation can obtain better effect on feature representation of traffic signs than invariant moments. The shape feature

extracted by central projection transformation is a global feature and invariant to object scales and rotations. Moreover, before the central projection transformation, detected traffic signs are not required to be normalized to the uniform size. Especially, the problem of confirming optimal projection number in central projection transformation is solved by the information entropy in this paper. Therefore, the proposed traffic sign recognition method in this paper is simple, reliable and high-speed for traffic sign recognition in natural scene images.

## REFERENCES

1. Alefs, B., Eschemann, G., Ramoser, H., et al, 2007. Road sign detection from edge orientation histogramms. In: *Proceedings of the 2007 IEEE Intelligent Vehicles Symposium*, Istanbul, Turkey, pp. 993-998.
2. Escalera, A. de la., Armingol, J. M., Mata, M., 2003. Traffic sign recognition and analysis for intelligent vehicles. *Image and Vision Computing*, 21(3), pp. 247-258.
3. Fang, C. Y., Fuh, C. S., Yen, P. S., et al, 2004. An automatic road sign recognition system based on a computational model of human recognition processing. *Computer Vision and Image Understanding*, 96(2), pp. 237-268.
4. Filipe, C., Correia, P. L., 2007. Automatic detection and classification of traffic signs. In: *Proceedings of IEEE Eighth International Workshop on Image Analysis for Multimedia Interactive Services*, Santorini, Greece, pp. 8-11.
5. Fleyeh, H., Dougherty, M., Aenugula, D., et al, 2007. Invariant road sign recognition with fuzzy ARTMAP and Zernike moments. In: *Proceedings of the 2007 IEEE Intelligent Vehicles Symposium*, Istanbul, Turkey, pp. 31-36.
6. Gao, X. W., Podladchikova, L., Shaposhnikov, D., et al, 2006. Recognition of traffic signs based on their colour and shape features extracted using human vision models. *Journal of Visual Communication & Image Representation*, 17(4), pp. 675-685.
7. Gil-Jimenez, P., Gomez-Moreno, H., Siegmann, P., et al, 2007. Traffic sign shape classification based on support vector machines and the FFT of the signature of blobs. In: *Proceedings of the 2007 IEEE Intelligent Vehicles Symposium*, Istanbul, Turkey, pp. 375-380.
8. Hsu, S. H., Huangl, C. L., 2001. Road sign detection and recognition using matching pursuit method. *Image and Vision Computing*, 19(3), pp. 119-129.
9. Hu, M. K., 1962. Visual pattern recognition by moment invariant. *IEEE Trans. On Information Theory*, 28(8), pp. 179-187.
10. Kuo, W. J., Lin, C. C., 2007. Two-stage road sign detection and recognition. In: *Proceedings of 2007 IEEE International Conference on Multimedia and Expo*, Beijing, China, pp. 1427-1430.
11. Li, Q. H., Zhang, H., 2006. A modified PNN-based method for traffic signs classification. *Systems Engineering*, 24(4), pp. 97-101.
12. Liu, H., Liu, D., Xin, J., 2002. Real-time recognition of road traffic sign in motion image based on genetic algorithm. In: *Proceedings of the 1st International Conference on Machine Learning and Cybernetics*, Beijing, China, pp. 83-86.
13. Loy, G., Barnes, N., 2004. Fast shape-based road sign detection for a driver assistance system. In: *Proceedings of International Conference on Intelligent Robots and Systems*, Sendai, Japan, pp. 70-75.
14. Maldonado-Bascon, S., Lafuente-Arroyo, S., Gil-Jimenez, P., et al, 2007. Road-sign detection and recognition based on support vector machines. *IEEE Transactions on Intelligent Transportation System*, 8(2), pp. 264-278.
15. Miura, J., Kanda, T., Shirai, Y., 2000. An active vision system for real-time traffic sign recognition. In: *Proceedings of IEEE Intelligent Transportation Systems*, Dearborn, USA, pp. 52-57.
16. Paclik, P., Novovicova, J., Duin, R. P. W., 2006. Building road-sign classifiers using a trainable similarity measure. *IEEE Transactions on Intelligent Transportation System*, 7(3), pp. 309-321.
17. Perez, E., Javidi, B., 2002. Nonlinear distortion-tolerant filters for detection of road signs in background noise. *IEEE Transactions on Vehicular Technology*, 51(3), pp. 567-576.
18. Specht, D. F., 1990. Probabilistic neural networks and polynomial adaline as complementary techniques for classification. *IEEE Transactions on Neural Network*, 1(1), pp. 111-121.
19. Tao, Y., Lam, E. C. M., Tang, Y. Y., 2001. Feature extraction using wavelet and fractal. *Pattern Recognition Letters*, 22(3-4), pp. 271-287.
20. Zhang, K., Sheng, Y. H., Gong, Zh. J., et al, 2007. A self-adaptive algorithm for traffic sign detection in motion image based on colour and shape features. In: *Proceedings of SPIE, Geoinformatics 2007: Remotely Sensed Data and Information*, Nanjing, China, 6752, pp. 6752G\_1-6752G\_13.

## ACKNOWLEDGES

This research is supported by National Foundation of Sciences under the grant number 40671147, and Momentous Science Foundation of Jiangsu province under the grant number 07KJA42005.

