TWO-STAGE ROAD SIGN DETECTION AND RECOGNITION

Wen-Jia Kuo and Chien-Chung Lin

Department of Information Management
Yuan Ze University, Chung-Li, Taiwan 32003, R.O.C.
wjkuo@saturn.yzu.edu.tw

ABSTRACT

We propose a road sign detection and recognition method using two-stage classification strategy. In the detection phase, geometric characters of road traffic signs, Hough transformation, corner detection, and projection are used to detect the exact position of the road sign in the image under noisy and complicated environment. In the recognition phase, convolution, radial basis function (RBF) neural network and K-d tree are used to recognize the road signs in two stages. Experimental results show that most road signs can be correctly detected and recognized by our proposed method with the accuracy of 95.5%. Moreover, the method is robust against the major difficulties of road sign detection and recognition. The proposed approach would be helpful for the development of intelligent Driver Support System and provide effective driving assistance message.

1. INTRODUCTION

With the rapid development of technologies, automatic detection and recognition of road signs have already become an important research topic in the field of intelligent vehicle. In the past few years, many techniques have been developed to detect road signs already [1]-[5]. Most studies dealt with single images under the assumption with simple backgrounds. For example, the SIFT transform has been employed to extract a set of invariant features for detection and recognition [1]. Genetic algorithm and neural network had also been used for road signs classification [3]. Hsu used templates to detect road signs and the matching pursuit method to identify road signs [4]. In this paper, we demonstrate a visual system for the detection and recognition of road signs in images of cluttered urban streets.

In this paper, we propose a road sign detection and recognition method using two-stage classification method. The RBF neural network is first used to group the possible candidates to decrease the error rate of false recognition. The K-d trees are then used to classify results of RBF neural network in the previous step. The K-d tree can also be used to modify the results of classifying the road sign into the wrong group. The organization of the proposed paper is as follows. Section 2 introduces the road signs detection method. The recognition method is presented in section 3. Section 4 demonstrates the experimental results with real images. Conclusions are discussed in Section 5.

2. ROAD SIGNS DETECTION

In general, there are three types of road signs that are shown in the traffic code including warning, prohibition, and informative. They can be roughly recognized depending on the form and the color.

2.1. Color thresholding

For the reason that the lighting conditions are changeable and not controllable in most road sign images, the HSI color space with more immune to lighting changes is used in our proposed method. Each image element is classified according to its hue, saturation, and intensity. The Anisotropic diffusion method [6] is then used to reduce noises. Table 1 shows the threshold values used in this stage.

Table 1 Threshold used for road sign detection.

	Red	Blue
Hue	$H \geq 0$ and $H<0.111\pi$ $H \geq 1.8\pi$ and $H<2\pi$	$1.066\pi \leq H \leq 1.555\pi$
Saturation	$0.1<S \leq 1$	$0.28<S \leq 1$
Intensity	$0.12<I<0.8$	$0.22<I<0.5$

2.2. Geometric characters

This section addresses the problem of detecting road signs within the region extracted in Section 2.1. We aim at the purpose of retrieving circular and triangular contours which represent possible candidates to be road signs. For both the circular and triangular shape detection algorithm, the edge map extracted by Soble operators is used for further detection.

2.2.1. Triangular road signs

To detect the specific position of triangular road signs in the images, we first perform the Hough transform to find possible lines of triangular road sign candidates. Those parameters do not satisfy the criterion for triangular are removed in this stage. Then, the lines are reconstructed using the remained parameters obtained from the Hough transform. The candidate region is then estimated by the intersection of reconstructed lines and edge map of the image. The median filter and dilation processes are used to remove noise and extend the candidate regions. The above steps are performed repeatedly until the number of reconstructed lines is stable to achieve the candidate regions.

2.2.2. Circular road signs

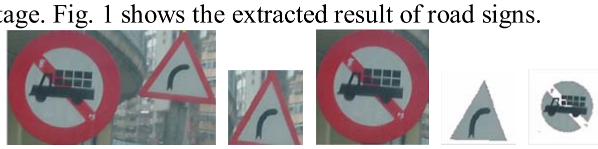
Circle Hough transformation is first performed to estimate the possible center of circle for the circular road signs. Median filter is used to remove the redundant noises. Then, the detected centers are used to extend the candidate region instead of reconstructing the circle by the detected parameters directly. The above steps are performed repeatedly until the region of centers is stable to achieve the candidate regions.

2.2.3. Connected road signs

In general, the road signs are usually appear in the connected form with more than one road signs. Therefore, it is necessary to divide them into individual road sign for recognition. In this paper, the ratio of width to height evaluated from the detected candidate region is considered as the criterion to differentiate the type of connected road signs. If the ratio of width to height is larger than 0.8, the region is considered as the form with horizontal connected road signs. If the ratio of height to width is larger than 0.8, the region is considered as the form with vertical connected road signs. The others are considered as the form with single road sign.

2.3. Division of connected road signs

In this paper, reduction of redundant background and division of road signs are accomplished by the projection method. For triangular road signs, the corner detection method [7] is cooperated with projection to divide the connected road signs. For circular road signs, projection of the centers of circles is used for division. The segmented candidate regions are then normalized to 100×100 using bicubic interpolation [8]. Once the candidate road signs have been segmented, area of the region is then used to recognize the type of shape. Background reduction is then performed to extract the main content of the road sign. The extracted content is then used for recognition at the next stage. Fig. 1 shows the extracted result of road signs.



(a) (b) (c) (d) (e)

Fig. 1. Example of road signs extraction. (a) Extracted candidate region (b), (c) Segmented results (d), (e) Normalized road sign contents.

3. ROAD SIGNS RECOGNITION

Because there are many different types of road signs, it becomes difficult to recognize them well in one stage. Therefore, we use the RBF neural network to classify the group of road sign to decrease the error rate of false recognition in the first stage, and then recognize it with corresponding K-d tree structure in the second stage.

3.1. Wavelet transform and convolution

It has been shown that the convolutional neural network performed well for some recognition problems [9]-[11]. In this paper, we use the Harr wavelet transform [9] to achieve the images for the next recognition steps. The convolution is defined as follows:

$$C(i,j) = \sum_{k=1}^{5} I_v(k,l), \tag{1}$$

where $I_v(k,l) = I(k,l+1) - I(k,l), k = 1, \ldots, 5; \; l = 1, \ldots, 4$.
The convoluted image is then transformed to binary expression with threshold method to extract the road sign components.

3.2. Road sign features

In order to construct the radial basis function neural network and K-d tree for further recognition, we use some features as the discriminative factors. The features we adopted are introduced as follows:

(1) Gradient factors:

We use the Soble operators to evaluate the gradient phase angle $\alpha(x,y)$ which defined as follows:

$$\alpha(x,y) = \tan^{-1}\left(\frac{G_x}{G_y}\right), \tag{2}$$

where $G_x = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}, G_y = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}$.

Then the histogram is evaluated according to the phase angle of pixels in the image. The accumulation of positive infinite, negative infinite, and the other two maximum bins are used as four gradient factor features.

(2) Angular radial transform (ART) coefficients:

The ART has been used as the shape descriptor in the MPEG-7 standard [12]. This descriptor preserves the shape properties under scaling, shift, and rotation. The ART coefficients are defined as follows:

$$\begin{aligned} F_{nm} &= \langle V_{nm}(\rho,\theta), f(\rho,\theta) \rangle \\ &= \int_0^{2\pi}\int_0^1 V_{nm}^*(\rho,\theta) f(\rho,\theta) \rho \, d\rho d\theta \end{aligned} \tag{3}$$

where $V_{nm}(\rho,\theta) = A_m(\theta) R_n(\rho)$,

$$A_m(\theta) = \frac{1}{2\pi}\exp(jm\theta),$$

$$R_n(\rho) = \begin{cases} 2, & \text{n} = 0 \\ 2\cos(\pi n\rho), & \text{n} \neq 0 \end{cases}$$

In this paper, the five real coefficients of the first two order ART coefficients ((n, m) = (0, 1), (1, 0), (1, 1), (1, 2), (2, 1)) are adopted as the features for further recognition.

(3) Entropy:

The entropy is evaluated according to the histogram of phase angles. The entropy is defined as follows:

$$entropy = -\sum p(a_i) \log p(a_i), \tag{4}$$

where $p(a_i)$ is the probability of histogram bin.

(4) Median point and connected component:

In addition to the above features, the connected component analyses of road signs are to be included in the recognition step to increase the completeness of description for road signs. The median point $M(x,y)$

defined as $M(x,y)=\frac{1}{n}\sum_{i=1}^{n}S(x_i,y_i)$ is used to evaluate the space distribution of components. $S(x_i,y_i)$ is the coordinate of pixel i in the convoluted image. The number of connected components *CMS* is used to distinguish variety of road sign components.

3.3. Radial basis function neural network

The RBF neural network has been widely used to classify multiple targets [13], [14]. Therefore, we use the RBF neural network to classify the road sign group at the first stage. In our proposed method, the form of RBF consists of three layers including input layer, hidden layer, and output layer. For the learning strategies, the weighting coefficients among input layer and hidden layer are adjusted by K-means clustering method. The weighting coefficients among hidden layer and output layer are adjusted by back-propagation method. Gaussian function is used as the basis function of RBF neural network.

To achieve better classification results, we modified the training processes of RBF neural network as follows:

Step 1: All the training patterns are used to construct the structure of RBF neural network. The number of designated output is equal to the number of different road signs. Then, choose the two road signs with the largest difference according to the real output value under this structure. Reconstruct a new RBF neural network with these two selected classes of road signs. The number of clusters recognized by this RBF neural network is two.

Step 2: Evaluate the real output value with the remained patterns which are not in the training set. Choose the one class whose variety of output value is minimal as the new class. The selected class of road sign is then added to reconstruct the new RBF neural network. The number of clusters recognized by this new RBF neural network will increase one. The training patterns are then evaluated by this RBF neural network.

(a) If the correct recognition rate is larger than 80%, repeat step 2.

(b) If the correct recognition rate is less than 80%, merge the two classes with the closest variety of output values. Reconstruct the RBF neural network, and the number of clusters recognized by this new RBF neural network will decrease one. Repeat step 2.

(c) If the correct recognition rate can't be improved up to 80% after the adjustment of step (a) and step (b), terminate the training process and go to step 3.

Step 3: If there is no remaining class, terminate the training process and the RBF neural network is used to classify the group of road signs. Otherwise, the remained classes are merged to the class with the closest variety according to the output value evaluated by the RBF neural network.

3.4. K-d tree

The K-d tree is a tree search method using nearest neighbor search. It can find the nearest neighbors quickly and differentiate data with multi-dimensional properties. In this paper, we use K-d tree structure for each group to recognize the road signs in the second stage. The road sign features are both used in the construction and recognition steps.

To achieve better classification results, we modified the rules for K-d tree with adaptive properties as follows:

(1) Each node contains n features and two pointers to the sub-trees.

(2) For each node P in the K-d tree, choose the representative feature sets K from the n features such that the node satisfies the feature sets K is unique. Meanwhile, the number of nodes for the left sub-tree and the number of nodes for the right sub-tree is equal or less than one.

(3) Traditionally, the representative feature sets K is the same for all nodes at the same level. In this paper, the representative sets for nodes of the K-d tree at the same level may be different for the purpose of increasing the flexibility of K-d tree structure.

(4) For node P which does not satisfy the feature sets K in the K-d tree, the number of feature sets for left sub-tree is less than the number of feature sets for right sub-tree.

The K-d tree is constructed as follows:

Step 1: For each road sign feature, evaluate the average value from the training patterns. The feature with maximal variance is used as the feature used for division.

Step 2: Find the median of the selected feature among the road signs to be the representative class of the node.

Step 3: Use the training patterns to modify range of threshold value for the selected feature of the node. If the number of different road signs can't be divided into equal size, take the other features into consideration. Repeat this process until the division result can be achieved.

Step 4: If there exists incomplete sub-tree of the group, repeat step 1 to step 3 until all road signs are inserted into the K-d tree.

In our proposed two-stage recognition method, RBF neural network is used to identify the group of the road sign at the first stage. Then, the corresponding K-d tree is used to recognize the road sign. If the difference between the road sign and the recognized node is larger than a threshold, it means that the road sign does not belong to this group. The other K-d tree will be used for further recognition. Therefore, the K-d tree can also be used to modify the results of classifying the road sign into the wrong group.

4. EXPERIMENTAL RESULTS

In this section, we will present the experimental results of road sign detection and recognition. The standard road signs including 39 triangular warning road signs, 26 circular prohibition road signs, and 15 obligation and informative road signs are used to construct the RBF neural network and K-d tree for recognition. To demonstrate the capability of the proposed system, the experiments are performed for 53 triangular warning road signs, 45 circular prohibition road signs, 12 obligation and informative road signs, respectively. The number of groups classified by RBF neural network and number of road signs for each K-d tree are shown in Table 2. Table 3 shows the overall performance of recognition results. The overall accuracy of proposed road sign recognition system is 95.5%. Fig. 2 shows examples of correct detection and recognition results. Table 4 shows the results compared with other recognition methods. We can find that the proposed method is better than the other methods which take different type of road signs as multiple targets directly.

Table 2 Overall of RBF groups and K-d trees.

	Groups and recognized road signs by K-d tree			
Triangular road signs	Group 1	Group 2	Group 3	Group 4
	4	5	13	17
Circular road signs	Group 1	Group 2	Group 3	Group 4
	3	9	7	7
Informative road signs	Group 1	Group 2	Group 3	
	1	6	8	

Table 3 Overall of system performance.

	Correct recognition	False recognition	Correct recognition rate (%)
Triangular road signs	49	4	92.45%
Circular road signs	44	1	97.78%
Informative road signs	12	0	100%
Overall correction rate	105/110=95.5%		

(a) (b) (c)

Fig. 2. Examples of correct detection and recognition results. (a) Multiple road signs (b) Multiple circular road signs (c) Informative road sign.

5. CONCLUSION

In this paper, we propose an automatic road sign detection and recognition method using two-stage classification method. In the detection phase, geometric characters of road traffic signs, Hough transformation, corner detection, and projection are used to detect the exact position of road signs in the image under noisy and complex environment. In the recognition phase, convolution, RBF neural network and K-d tree are used to recognize the road signs in two stages. Experimental results show that most road signs can be correctly detected and recognized by our proposed method with the accuracy of 95.5%. Moreover, the method is robust against the major difficulties of road sign detection and recognition such as image scaling and rotation, illumination change, deformation, and so on. The proposed approach would be of great help for the development of intelligent Driver Support System and Intelligent Autonomous Vehicles to provide effective driving assistance message.

Table 4 Results compared to other recognition methods.

	Correct recognition	False recognition	Correct recognition rate (%)
Zernike moments with projection	68	42	61.8%
Single back-propagation neural network	10	100	9.1%
Single RBF neural network	14	96	12.7%
Single K-d tree	6	104	5.5%
Proposed method	105	5	95.5%

6. REFERENCES

[1] A. Farag, A. E. Abdel-Hakim, “Detection, Categorization and Recognition of Road Signs for Autonomous Navigation,” in *Proc. ACIVS 2004*, Aug. 2004, Belgium, pp. 125-130.

[2] Y. Y. Nguwi, A. Z. Kouzani, “A Study on Automatic Recognition of Road Signs,” in *Proc. IEEE Conf. on Cyber. And Intell. Sys*., June 2006, pp. 1-6.

[3] A. de la Escalera, J. M. Armingol, M. Mata, “Traffic Sign Recognition and Analysis for Intelligent Vehicles,” *Image and Vision Comp.*, vol. 21, pp. 247-258, 2003.

[4] S. H. Hsu, C. L. Huang, “Raod Sign Detection and Recognition Using Matching Pursuit Method,” *Image and Vision Comp.*, vol. 19, pp. 119-129, 2001.

[5] H. Janssen, W. Niehsen, “Vehicle Surround Sensing Based on Information Fusion of Monocular Video and Digital Map,” in *Proc. 2004 IEEE Intelligent Vehicles Symposium*, June 2004, pp. 244-249.

[6] M. M. Oliveira, B. Bowen, R. McKenna, and Y. S. Chang, “Fast Digital Image Inpainting,” in *Proc. Int. Conf. on Visualization, Imaging and Image Processing, VIIP 2001*, 2001, Marbella, Spain., pp. 261-2665.

[7] A. de la Escalera, L. E. Moreno, M. A. Salichs, and José María Armingol, “Road Traffic Sign Detection and Classification,” *IEEE Trans. Industrial Electronics*, vol. 44, issue 6, pp.848-859, Dec. 1997.

[8] Paul Bourke, Bicubic Interpolation for Image Scaling, http://astronomy.swin.edu.au/~pbourke/colour/bicubic.

[9] Y. N. Chen, C. C. Han, and K. C. Fan, “The Application of Convolutional Neural Network on Face and License Plate Detection,” in *Proc. 18th Conf. on Computer Vision Graphics and Image Processing*, *CVGIP 2005*, Taipei, 2005, pp. 31-37.

[10] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, **“**Face recognition: a convolutional neural-network approach,**”** *IEEE Trans. Neural Networks*, vol. 8, issue 1, pp. 98-113, Jan. 1997.

[11] C. Nebauer, **“**Evaluation of convolutional neural networks for visual recognition,**”** *IEEE Trans. Neural Networks*, vol. 9, issue 4, pp. 685-696, July 1998.

[12] ISO/IEC JTC1/SC29/WG11/N4063, “MPEG-7 Visual part of Experimentation Model Version 10.0,” Singapore, March 2001.

[13] Chien-Cheng Lee, Pau-Choo Chung, Jea-Rong Tsai, and Chein-I Chang, “Robust Radial Basis Function Neural Networks,” *IEEE Trans. System, Man, and Cybernetics—PART B: CYBERNETICS*, vol. 29, no. 6, pp. 674-685, December 1999.

[14] Y. H. Cheng and C. S. Lin, “A learning algorithm for radial basis function network: With the capacity of adding and pruning neurons,” in Proc. *ICNN’*94, vol. 2, pp. 797–801, 1994.