AN IMPROVED DtBs METHOD FOR AUTOMATIC TRAFFIC SIGN RECOGNITION

Tsung-Hsien Juan, Jen-Yu Han
Department of Civil Engineering, National Taiwan University
E-mail: r02521113@ntu.edu.tw, jyhan@ntu.edu.tw

KEY WORDS: Traffic sign, object recognition, rotation correction, Support Vector Machine, Distances to Boundary (DtBs) method.

ABSTRACT: In this study, we developed an improved image-based traffic sign recognition approach based on the Distances to Boundary (DtBs) method. The proposed algorithm is capable of detecting and adjusting the in-plane and out-of-plane rotations of each candidate object. Consequently, this approach is translation-, scale-, and rotation-invariant. Based on the preliminary results from a real field experiment, it has been demonstrated that a significant improvement on the traffic sign detection result can be achieved after implementing the proposed approach.

1. INTRODUCTION

Road traffic signs are important facilities for road safety. Information carried by traffic signs can indicate road conditions, regulate traffic behaviors, provide navigation and warnings, and ensure road functionality (Shoba & Suruliandi, 2013). Automatic road sign detection and recognition technique can be used in many fields such as automatic driver assistant, intelligent navigation, and smart road system (Waite & Oruklu, 2013). Road sign detection and recognition techniques is also useful for road maintenance, which can provide field information in short time, making maintenance tasks more efficient. For the above reasons, efforts are continuously made in developing an automatic traffic sign recognition system using vehicle-borne images or videos (Bui-Minh et al., 2012).

There are a number of challenges for the automatic recognition of traffic signs from vehicle-borne images (see Table 1). First, the lighting conditions in an outdoor environment may vary drastically, caused by the time of day or night, and by the changes of the weather. Another problem is that as the distance between camera and traffic sign changes while collecting data, the imaging scale also varies. Therefore, the traffic sign may have varied sizes in the recorded images. The next problem is the possible distortion of traffic signs in the image caused by rotation and translation. Although the position of traffic signs is supposed to be perpendicular to the driving direction of vehicles, in many cases the sign is not properly positioned. Besides, the traffic sign itself might somehow be rotated. In this study, rotations bringing about to signs that are not perpendicular to the vehicle trajectory is defined as out-of-plane rotations, and the rotation within the sign plane is defined as in-plane rotations, both of them are needed to be eliminated. Finally, occlusions caused by adjacent objects such as trees, cars, and other signs can reduce the visibility of target and lead to the failure of detection.

Table 1. Challenges of sign recognition using image-based approaches

Different sizes and lighting conditions	Occlusion	In-plane rotation	Out-of-plane rotation
 			

In this study, we developed an image-based traffic sign recognition approach. This approach is mainly composed of two phases: traffic sign detection phase and traffic sign recognition phase. In the sign detection phase, candidate objects are extracted from input images by a color-based method. In sign recognition phase, the shape of each candidate is determined using an improved algorithm based on the Distances to Boundary (DtBs) method. It is capable of detecting and adjusting the in-plane and out-of-plane rotations of each candidate object. Consequently, this approach is translation-, scale-, and rotation-invariant, and is robust with variable lighting conditions.

2. TRAFFIC SIGN DETECTION

In order to draw people's attention and stand out from a cluttered background, traffic signs are designed in showy and intense color. With this characteristic, color-based segmentation by thresholding a certain component in color space to extract traffic sign candidates is frequently used in many studies. However, a direct thresholding over the red-green-blue (RGB) space is seldom used because the components of RGB space are very sensitive to lighting condition. Instead, the relative relationship between RGB components is more frequently employed. In (Soheilian et al., 2013), the ratio between the intensities of a specific color component and of other components is calculated to detect the color of interest. A different method is to generate a vector containing the RGB components of each pixel in the whole image. It is then used as a feature vector and a neural network is performed for color classification and segmentation (Zhang et al., 2013). In addition to the RGB space, another color space that is also frequently employed is the hue-saturation-intensity (HSI) space. Because in the HSI space the reflect intensity is recorded in the intensity component, and the color information is encoded in the hue and saturation components, it has a higher degree of invariance to illumination condition. For this reason, many studies have used the HSI space on sign detection and obtained good results (Yok & Kouzani, 2006; de la Escalera et al., 2004).

The segmentation in RGB space is robust enough under a well-lighted environment, and the segmentation in HSI space can provide assistance while the illumination condition is not stable enough. For this reason we use a hybrid method to extract the color of interest. Take the red color as an example, the criterion can be written as:

$$\begin{cases} (R(i,j) - [G(i,j) + B(i,j)] > 0 \\ H(i,j) < 0.05 \text{ or } H(i,j) > 0.95 \end{cases} \tag{1}$$

where *R(i, j)* represents the intensity value in red band. *G(i, j)* and *B(i, j)* are the intensity values in green and blue bands, respectively. *H(i,j)* denotes the corresponding hue value. After the segmentation step, a binary image containing blobs identified as red color is first obtained. Next, a cluster analysis is employed and the blobs whose sizes are within a predefined range are identified. The blob with an unreasonable size is treated as noise and removed from the binary image. Finally, an adaptive bounding box is applied to each object in the image, and proper candidate objects can be finally extracted (see Figure 1).

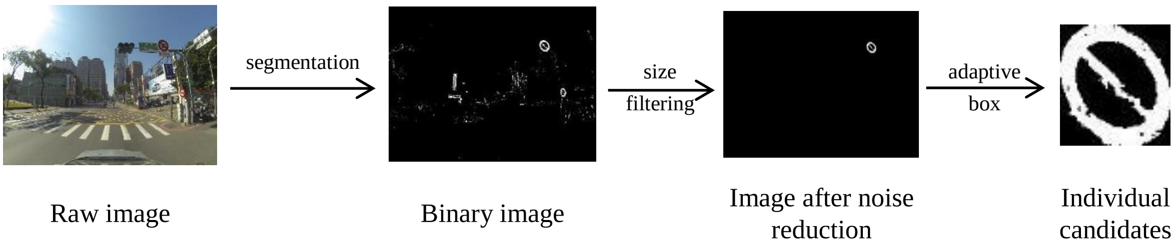
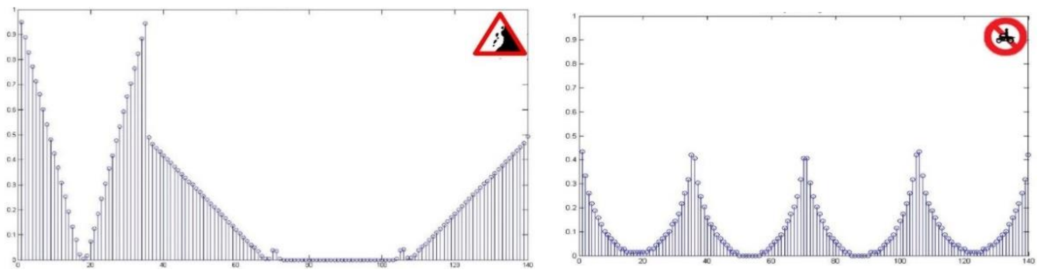


Figure 1. Data flow in the traffic sign detection phase

3. IMPROVED DtBs ON SHAPE RECOGNITION

The second phase of the proposed approach is on the geometric shape classification for the candidates extracted from the segmentation phase. As mentioned above, the translation, rotation, and varied scales may cause distortions in the recorded images; these distortions need to be corrected in order to give better recognition results. In this study, an improved classification approach derived from the DtBs method is proposed. The original DtBs method is first presented in (Lafuente-Arroyo et al., 2004). It calculates the distances between blobs to the four adaptive box boarders. The DtB vectors are first normalized making their values ranging between 0~1. Because of the normalization, the effects of varied scales and the distortion caused by out-of-plane rotations can be eliminated (Maldonado-Bascon et al., 2007). Figure 2 illustrates the DtB vectors of triangular and circular shape. It is shown that each object shape has a distinct DtB vector. Consequently, the DtB vectors can be used as feature vectors for the shape classification.



(a) (b)

Figure 2. DtBs of (a) triangular shape and (b) circular shape

The workflow of the proposed shape recognition is illustrated in Figure 3. The DtB vector of each candidate object is first obtained and fed into the linear support vector machines (SVMs) to determine the possible shape. Support vector machines were introduced by Cortes and Vapnik (Cortes & Vapnik, 1995), whose main idea is to map data into a higher-dimensional feature space, and a hyperplane that best separates the two classes can then be determined. For the training dataset $\{x_i, y_i\}$, where $i = 1,2,...,l$, $y_i \in \{-1,1\}$, $x_i \in \{\mathbf{R}^\mathbf{d}\}$, the decision hyperplane can be written as:

$$\mathbf{w}^T \cdot \mathbf{x}_i + b = 0 \tag{2}$$

where $\mathbf{w}$ is the normal vector of hyperplane, $b/\|\mathbf{w}\|$ is the offset from the origin along the vector $\mathbf{w}$. In the binary classification problem, the following constraint holds:

$$y_i(\mathbf{w}^T \cdot \mathbf{x}_i + b) \geq 1 \quad \forall i \tag{3}$$

Vectors for which the equality in Eq. (3) is held lie on two support hyperplanes: $H_1 : \mathbf{w}^T \cdot \mathbf{x}_i + b = 1$ and $H_2 : \mathbf{w}^T \cdot \mathbf{x}_i + b = -1$. Therefore, the margin between these two hyperplanes is $2/\|\mathbf{w}\|$. Since the optimal decision hyperplane needs to have the maximum margin to the two support hyperplanes, the SVM solves the following optimization problem:

$$\begin{cases} \min \frac{1}{2}(\mathbf{w}^T \mathbf{w}) \\ s.t.\ \ y_i(\mathbf{w}^T \cdot \mathbf{x}_i + b) \geq 1 \end{cases} \tag{4}$$

In order to find the optimal solution, the Lagrange multiplier α is introduced. Once the optimization is completed, the decision hyperplane is also determined as:

$$y_i = f(x) = \mathbf{w}^T \mathbf{x} + b = \sum_i \alpha_i y_i \mathbf{x}_i{}^T \mathbf{x} + b \tag{5}$$

Based on Eq. (5), one can simply determine the class on which side of the hyperplane a given test vector **x** lies. In our case, that is to classify it to one class/shape (y=1) or to the other (y=−1). For the reason that SVM can only perform a binary classification, we designed numbers of one-against-all SVM classifiers to classify the candidate object that belongs a certain shape. The distance between candidate vector and the decision hyperplane is calculated as the classification score, and the class with a highest score is assigned to the corresponding candidate.

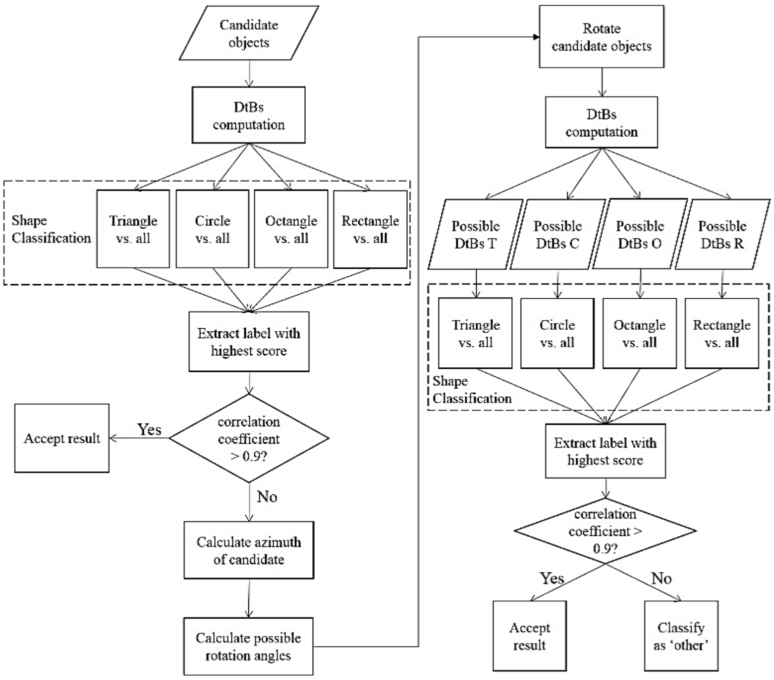
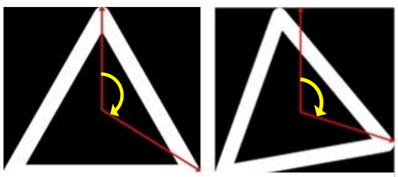


Figure 3. Workflow of the proposed approach on shape recognition

To avoid misclassification and false negatives, the correlation coefficient between candidate DtBs and template DtBs is computed after the above classification analysis. If the value does not pass through a predefined criterion, the workflow switches into the second stage of recognition. In the second stage, the azimuth angle of the longest vector in the candidate image



(a) (b)

Figure 4. Azimuth of (a) a correct image and (b) a rotated image

is computed (see Figure 4). For a regular polygon without a rotation, the azimuths from its centroid to each vertex are constant values. Furthermore, the longest vector in a polygon should always coincide with one of the vertices, so the difference of azimuth between the same vector in the candidate image and the template image can be determined as the in-plane rotation angle. With this method, the in-plane rotation angle of a candidate with respect to each matching shape is obtained. Therefore, we can rotate the images back to their nominal orientation and compute the corresponding DtBs. With this treatment, the shape recognition is performed again and the shape with the highest score is classified to the corresponding candidate.

Following the proposed workflow, sign objects with very little or no rotations can be recognized quickly in the first stage, and the in-plane rotation angle of candidate object can be determined in the second stage of the process. Also, this method is not affected by out-of-plane rotation and different scales due to the normalization of the DtBs vectors.

4. EXPERIMENTAL RESULT

A field test was performed to evaluate the proposed approach. A digital camera with a resolution of 1296 × 964 pixels was installed on a vehicle and used to acquire the images at a rate of 20 fps (frames per second). The averaged driving speed of the vehicle was around 20 km/hr while acquiring the images. Totally 10738 images were recorded along a 2.4 km route in the test field. Two cases were tested in the experiment. Case 1 classified the candidate objects using the original DtBs without employing a rotation correction, while case 2 processed images with the proposed method with a rotation correction. Table 2 summarizes the test results for the two cases. It can be seen that there were totally 9 traffic signs in the test filed. In case 1 when the original DtBs was employed, only 5 traffic signs have been successfully detected. On the other hand, when the improved approach with a rotation correction was applied, all the 9 traffic signs in the test field have been successfully detected. Figure 5 illustrates those traffic signs with rotations or under a bad lighting condition. They were not correctly detected by the original DtBs. However, they can all be identified by the proposed approach. As a consequence, a significantly improved detection result has been achieved after implementing the proposed approach.

Table 2. Summary of experimental result

	Case 1: without rotation correction	Case 2: with rotation correction
Number of recorded images	10,738	10,738
Detection of traffic signs	5	9
Total of traffic signs	9	9
Detection rate	55.6%	100%

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

Figure 5. Examples of successful detections by the proposed approach for the traffic signs with in-plane rotations. Note that (d) is also under a bad lighting condition.

5. CONCLUSION

In this study, an improved approach based on the Distances to Boundary (DtBs) method is proposed. Based on the field experimental result, there are three conclusions to our approach. First, the segmentation by combining RGB and HSI spectrum spaces is invariant to outdoor lighting conditions. Traffic signs in a shadowed area can be extracted successfully with this hybrid approach. Secondly, the implementation of the in-plane rotation correction results in a significant improvement to the shape recognition result. It is capable of detecting and correcting the rotations correctly without incurring an iterative computation. Finally, the normalization to DtBs makes the algorithm invariant to out-of-plane rotations and variable scales.

In the future, we will focus on the texture recognition of traffic signs and their 3-D reconstruction in the object space. Our goal is to develop a complete automatic traffic sign recognition and 3-D reconstruction system for transportation applications.

6. REFERENCE

Bui-Minh, T., Ghita, O., Whelan, P. F., & Hoang, T., 2012. A robust algorithm for detection and classification of traffic signs in video data. Paper presented at the Control, Automation and Information Sciences (ICCAIS), International Conference on.

Cortes, C. and V. Vapnik, 1995. Support-vector networks. Machine learning, 20(3), pp: 273-297.

de la Escalera, A; Armingol, J.M.; Pastor, J.M.; Rodriguez, F.J., 2004. Visual sign information extraction and identification by deformable models for intelligent vehicles. Intelligent Transportation Systems, IEEE Transactions on, vol.5, no.2, pp.57-68

Lafuente-Arroyo, S., Garcıa-Dıaz, P., Acevedo-Rodrıguez, F., Gil-Jiménez, P., & Maldonado-Bascón, S., 2004. Traffic sign classification invariant to rotations using support vector machines. Proceedings of Advanced Concepts for Intelligent Vision Systems, Brussels, Belgium.

Maldonado-Bascon, S., Lafuente-Arroyo, S., Gil-Jimenez, P., Gomez-Moreno, H., and Lopez-Ferreras, F., 2007. Road-Sign Detection and Recognition Based on Support Vector Machines. Intelligent Transportation Systems, IEEE Transactions on, 8(2), pp. 264-278.

Shoba, E., & Suruliandi, A., 2013. Performance analysis on road sign detection, extraction and recognition techniques. Circuits, Power and Computing Technologies (ICCPCT), 2013 International Conference on

Soheilian, B., Paparoditis, N., & Vallet, B., 2013. Detection and 3-D reconstruction of traffic signs from multiple view color images. ISPRS Journal of Photogrammetry and Remote Sensing, 77(0), pp. 1-20.

Waite, S., & Oruklu, E., 2013. FPGA-Based Traffic Sign Recognition for Advanced Driver Assistance Systems. Journal of Transportation Technologies, 3, 1.

Yok, Y. N., Kouzani, AZ., 2006. Automatic Road Sign Recognition Using Neural Networks. Neural Networks,. IJCNN '06. International Joint Conference on, vol., no., pp.3955-3962

Zhang, K., Sheng, Y., & Li, J., 2012. Automatic detection of road traffic signs from natural scene images based on pixel vector and central projected shape feature. Intelligent Transport Systems, IET, 6(3), pp. 282-291.