IJRSET Volume 2, Issue 2 Pages: 32-40

ISSN 2394-739X

International Journal for Research in Science Engineering and Technology

# Traffic Sign Detection and Recognition Using Feature Based and OCR Method

[1] Dilip Singh Solanki, [2] Dr. Gireesh Dixit
[1] Department of Computer Science and Engineering
[1] Madhav Proudyogiki Mahavidyalaya, Bhopal
[2] Department of Computer Science and Engineering
[2] Madhav Proudyogiki Mahavidyalaya, Bhopal

## Abstract:-

Traffic sign detection is a technology by which a vehicle is able to recognize the different traffic signs located on the road and used to regulate the traffic.

Traffic signs are detected by analyzes color information contained on the images having capability of detection and recognition of traffic signs even with bad visual artifacts those originate from different conditions.

The feature based method is intended for traffic sign detection, in this method two sets of features are to be detected in the reference and sensed images, identifying key points in the images and match among those points to find the similarity, the SURF descriptor is used for key points and point matching.

After detecting the shape of the traffic sign the optical character recognition (OCR) method is used to recognize the character present in the detected shape.

A technique, based on Maximally Stable Extremal Regions (MSER) region and canny edge detector is supervised for character recognition in traffic sign detection.

**Keywords:** - Color analysis, Feature Based method (SURF), OCR method, MESR region, canny edge detector

## 1. INTRODUCTION

Detection and recognition of traffic signs is an essential task of regulating the traffic, guiding and warning driver's pedestrians. In traffic sign detection research are categorized into three groups. The first group of researchers believes that traffic sign detection colors are important information by which traffic signs can be detected and classified. The second group believes that detection of traffic signs can be achieved by traffic sign shape only, and the third believes that color together with shape make the backbone for any road sign detection. Although traffic signs are apparent and have several obvious characteristics, some conditions may prevent driver perceiving them. For instance, at night or in bad lighting conditions drivers are less likely to notice the traffic signs. Some distracting events on road may result in a skip of signs.

Moreover, sometimes only the driver himself is not able to notice the signs due to lack of concentration. Driving needs continuous processing of visual information from the road. To avoid accident, driver needs to monitor a lot of traffic signs.

For this purpose traffic signs play an important role to provide information about traffic and road conditions which is necessary for a driver to accomplish a collision free driving environment.

Feature based method is used for traffic sign detection it possesses two sets of feature to get detected in your reference and sensed images. The feature based detection methods are usually deals with the point features. The point feature set consists of method functioning having line road crossing, centers of regions, intersections, corners, end points, etc. A large amount effort has been used up in developing accurate, robust, and fast method for corner detection image mosaic, and automatic change detection would be the matching accuracy which impacts these purposes. Usually these purposes require the corresponding accuracy to sub-pixel, the matching unit with the feature based approaches is “one pixel”. The pair-wise corresponding features can be utilized as suggestions for sub-pixel corresponding with other approaches. Speeded up robust features (SURF) [1] is the best feature-based algorithms have been widely used in computer vision purposes. It extracts Hessian matrix-based interest points and generates a distribution-based descriptor, and is a scale and rotation invariant algorithms. These features help it become perfect for object matching and detection. The SURF (Speeded-Up Robust Features) algorithm is based on the matching principles, but it utilizes a different scheme and it should provide better and faster results. The image descriptors and feature matching are both quite noisy processes, to make our algorithm more robust, we will use the RANSAC algorithm. The concept of Maximally Stable Extremal Regions (MSER) was proposed by Matas et al [2]. MSER is really a computer vision, maximally stable extremal regions used to be a method intended for blob detection in images. This technique of extracting an extensive number associated with corresponding image elements contributes toward the wide baseline matching, and it's led to better stereo corresponding and object recognition algorithms. The MSER algorithm extracts from an image a number of covariant regions, called MSERs. An MSER is a steady associated component of some gray-level sets of the image. MSER is based on the idea of taking regions which stay nearly the same through a wide range of thresholds. The word extremal refers to the property that all pixels inside the MSER have either higher or lower intensity than all the pixels on its outer boundary. This operation can be performed by first sorting all pixels by gray value and then incrementally adding pixels to each connected component as the threshold is changed, regions such that their variation to the threshold is minimal and defined maximally stable. Canny intend to discover the optimal edge detection algorithm [3]. In this situation an optimal edge detector means that it marks as many real edges as possible, that the marked ones are as close as possible to the real edges and noise effects are avoided. Canny algorithm use the calculus of variations which is a technique that finds the function which optimizes a given functional. The optimal function in canny edge detector is described through the sum of four exponential terms, but it can be approximated by means of the first derivative of a Gaussian. To reduce noise it uses a filter based on the first derivative of a Gaussian giving result as slightly blurred version of the original image, not affected by a feasible single noisy pixel. Since an edge may point out in a variety of directions, the canny algorithm use four type of filter to detect vertical, horizontal and diagonal edges in the blurred image. The edge detection operator returns a value for the first derivative in the horizontal and vertical directions; from this the edge gradient and direction can be determined. The edge direction angle is rounded to one of the four angles representing vertical, horizontal and two diagonals, later on from the stage referred to non maximum suppression set of edge points in the form of binary image is obtained and sometimes referred to as thin edges, after that thresholding is complete. We will have a binary image where each pixel is marked as either an edge pixel or a non edge pixel.

## 2. LITERATURE REVIEW

There are several methods for traffic signs detection and recognition. Vision-based sign detection systems mostly suffer from adverse weather and lighting conditions. A sign detection system can be decomposed into two separate parts: detection and classification. Researchers have proposed various techniques for detection and classification. HSI model is regarded as most suitable for traffic sign detection simply Fang et al.

[4] presents human color perception where colors of traffic indication are originally chosen to help attract human attention. The algorithms developed mainly depending on hue component to localize traffic signs. Moreover, they set minimum threshold of saturation component because to ignore unsaturated pixels where hue value may not correspond to true color value. There are some more color connection algorithms to further improve performance of color feature extraction. Piccioli et al.[5] suggested subdividing graphic 16x16 pixels region in addition to classifying each region as '1' if many labeled pixels for selected color exceeded certain threshold otherwise '0'. Piccioli et al. [6] represent edge image by Canny's algorithm applied to color segmented image. Different approaches are applied to detect different geometrical patterns. After obtaining edge image, to identify triangle shapes, a polygonal approximation in the edge chains inside search region color segmented region via previous step to eliminate the main chains strongly departing from a straight segment by Piccioli et al. [7]. Angles among line segments are extracted to choose triangles finally, detecting circles that has a similar method explained above is more challenging because results are unstable therefore there occur very poor rate of good results. Piccioli et al. use a different method to identify circle, which is especially based on radial and angular distribution in the edges. They compare their solutions to detect circles with that of Etemadi et al. [8], Hough et al. [9] and Masciangelo et al. [10] regarding algorithms are usually used to identify elliptic curves within image processing. SURF's detection scheme is based on the concept of automatic scale selection, proposed by Lindeberg et al. [11] in 1998. In this work, Lindeberg experimented with using the determinant of the Hessian matrix for a 2-D Gaussian, as well as the Laplacian (i.e. the Hessian's trace), to detect blob-like structures in images. Mainly motivated by Lindeberg findings, the authors of SURF chose the determinant of Hessian as their target feature. Other well-known feature schemes include the famed Harris et al. [12] corner detector which relies on Eigen values of the second moment, the entropy-based salient region detector proposed by Kadir and Brady et al. [13], and the edge based work of Jurie and Schmid et al. [14]. Furthermore, SURF's detector extends on Lowe's idea of using the Difference of Gaussian as an approximation of the Laplacian of Gaussian filter. Speeded up Robust Features (SURF) et al. [15] is one of the best feature-based algorithms and has been widely used in computer vision applications. It extracts Hessian matrix-based interest points and generates a distribution-based descriptor, and is a scale- and rotation-invariant algorithm. These features make it perfect for object matching and detection. The key advantage of SURF is to use integral image for feature detection and description, which greatly boosts the process efficiency. Even like other feature based algorithms, SURF is computationally expensive and often results in very low frame rate. In order to employ SURF for real-time applications in portable driver-assistance systems, parallel processing architectures and platforms need to be considered. The task of scene text detection in still images, the large number of published methods can be divided into texture based and region-based approaches. To shortly summarize recent leading research on texture-based methods, Coates et al. [16] and Wang et al. [17] have proposed the use of unsupervised feature learning to generate the features for text versus non-text patch classification. Wang have built an end-to-end scene text recognition system based on a sliding window character classifier using Random Ferns, with features originating from the Histogram of Oriented Gradients (HOG) descriptor. On the other hand, among the many recently published region-based methods we can observe an increasing use of the Maximally Stable Extremal Regions (MSER) algorithm for character candidates detection. For example, Chen et al. [18] obtain state-of-the-art performance with a method that determines the stroke width of edge-enhanced MSERs using the distance transform. The effectiveness of MSER is also exploited by Novikova et al. [19] and Merino et al. [20] among others, while Neumann and Matas propose a region representation derived from MSER where character/non-character classification is done for each possible Extremal Region. Wu et al. [21] found candidate locations, using a combination associated with Shi and

Tomasi characteristics Gaussian mixture, and geometric study. The traffic sign text appeared on the vertical plane with regards to the motion and optical axis from the camera. However, it is likely that text signs will appear from a viewpoint that is not pretty front to parallel. As a result, a perspective transform is important to give OCR a better chance of text recognition, as perform in the method. Candidate regions be matched up through consecutive frame, and be interpreted utilize an OCR system once we were looking at an adequate dimension. The authors reported the detection rate of 88. 9% as well as a false detection rate associated with 9. 2%, based on the data set of twenty-two video sequences, each around 30 s long.

## 3. PROPOSED METHODOLOGY

In earlier studies, simple color and shape based detection methods have been mostly used. Recently, feature-based traffic sign detection methods are proposed to obtain more accurate results mainly when combined with the previous methods. The Speeded up Robust Features (SURF) algorithm is an outstanding feature detector and descriptor with rotation and illumination invariance.

### A. Detecting and Extract Feature descriptors

We should detect key points to match two images. The feature based methods are utilized to detect two sets of features from the reference and sensed image. The features can be an edge, a corner, a point, a line or a curve, etc. Feature matching is to discover the pair-wise corresponding features. The actual matching accuracy will result these applications, normally these types of applications require the matching accuracy to sub-pixel. The matching part of the feature based methods is one pixel. The pair wise corresponding features can be used as an input regarding sub-pixel matching with other methods.


**Figure. 1 Extracted Feature**

The extract Features is used for the feature extraction on each key point. The feature extraction methods are used to detect two sets of features within the reference and sense images. A large amount of attempts has been used up in developing accurate and fast method for corner detection.

### B. Matching Features

Matching features including outliers using their descriptors to locate the object in scene using matched points. In order to locate the object in scene estimate geometric transformations i.e. affine transform is calculated. The transformation relating the matched points, allow us to restrict the object in the scene and finally transform the reference image into the match up system of the target image. The transformed image indicates the position of the object in the scene.

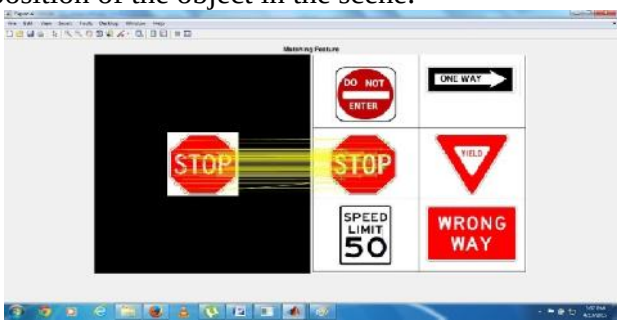


**Figure. 2 Matching Feature**

The transformation between the images needs to be as accurate as possible; however image descriptors and feature matching are both noisy processes. The descriptors are subject to image noise and compression artifacts and not all assumed correspondences are true correspondences due to descriptor error and ambiguities in the matching. Incorrect matches will insert error to our estimation and can adversely artifact the result. To make our algorithm robust, we will use the RANSAC algorithm, it is a method for estimating a parametric model from noisy observations.

### C. Shape Detection

The focus of this technology is to detect the geometric shapes of the traffic signs. Geometric Transform calculates the transformation relating to the matched points. This type of transformation allows us to restrict the object in the scene and finally transform the reference image into the match up system of the target image. The

transformed image indicates the position of the object in the scene and the shape of the traffic sign is detected. After detection of the sign the character has is to be recognized in the traffic sign using character recognition method.


**Figure. 3 Shapes Detected**

## D. Optical Character Recognition Method

The optical character recognition process for the segment the text from a cluttered scene deeply helps among additional tasks. A successful sign translation method relies on two key technologies text detection and optical character recognition. At existing stage of the research we focus our efforts on traffic sign recognition and text detection while taking advantage of existing OCR technologies. Detection of signs from natural scenes is a challenging problem since they usually surrounded in the environment. The task is associated to text detection and recognition from the images so called optical character recognition. Compared to OCR tasks sign detection takes place in additional dynamic environment. Sign detection should also be implemented in real time using limited resources. This features of signs means that lexical mismatch and structural mismatch become more serious problems. Moreover, sign conversion usually requires context or environment information since sign designers assume that a human reader would use such information in understanding signs.

## E. Maximally Stable Extremal Regions (MSER)

The text characters generally have consistent color we initiate finding regions with similar intensities in the image with the MSER region detector. MSER is based on the idea of taking regions which keep on nearly the same through a large range of thresholds. This operation can be performed by first sorting all pixels by gray value and then incrementally adding pixels to each connected component as the threshold is changed. Regions such that their variation to the threshold is minimal and defined maximally stable.


**Figure. 4 MSER Regions**

## F. Canny Edge Detector

After extracting the regions Canny Edge Detector is used for additional segment the text; since written text is usually to be found on clear background, it tends to generate high response toward edge detection. Moreover, an intersection of MSER regions through the edges is obtainable to generate regions that are even more expected to belong to text. The MSER regions still include pixels that are not a part of the text. We can make use of the edge mask together with edge gradients to eliminate those regions.


**Figure. 5 Canny Edge with MSER region**

Filter character candidates using connected component study of the remaining associated components at this instant be removed by means of their region properties. The thresholds used under might be varying for different image sizes, fonts or languages.

## G. Filter Character Using the Stroke Width Image

Filter character candidates using the stroke width image is another useful discriminator for text images, the variation in stroke width within each

text candidate. Characters in most languages have a similar stroke width or thickness throughout. It is therefore useful to remove regions where the stroke width exhibits too much variation

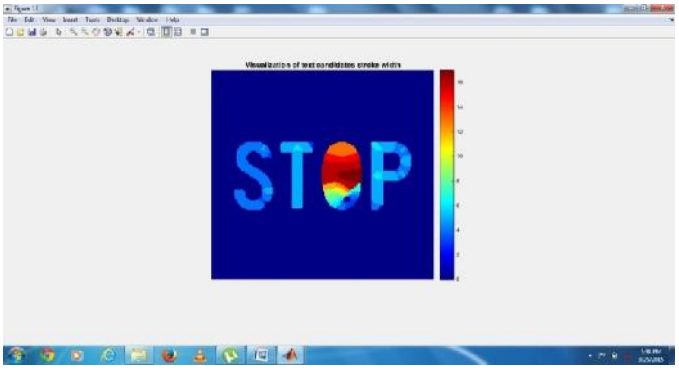


**Figure. 6 Text with Stroke width**

### H. Bounding Boxes Enclosing Text Regions

To compute bounding box enclosing the text region we initially join the individual characters into a single connected component. This can be accomplished by means of morphological closing follow by opening to clean up any outliers.


**Figure. 7 Image joining individual characters**

### I. Optical Character Recognition on Text Region

The segmentation of text from a cluttered scene can significantly improve optical character recognition results. The optical character recognition functions returns the recognized text, and the location of the text is in the original image.


**Figure. 8 Text Character Detected**

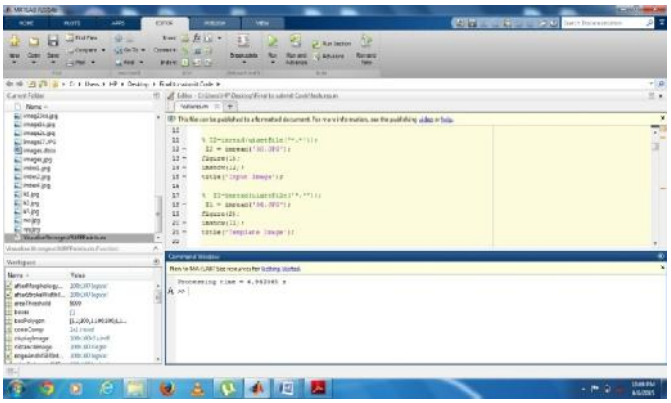

**Figure. 9 Processing Time**

This test is to measure the processing time of the algorithm. Processing times from experiments are computed for algorithm and the processing time is 6.42 Seconds.

## 6. TEST RESULTS

The test result is applied to the group of 80 images which contains 28 circular, 19 rectangular, 21 triangular and 22 octagonal traffic signs. Detection rates decrease with the improper or low illumination conditions.

| Traffic Signs | Signs to be Detected | Detected Traffic Signs | % of Recognition | False Positives |
|---|---|---|---|---|
| **Circular** | 28 | 26 | 92.85 | 2 |
| **Rectangular** | 19 | 18 | 94.73 | 1 |
| **Triangular** | 21 | 18 | 85.71 | 3 |
| **Octagonal** | 22 | 21 | 95.45 | 1 |

**Table1: Detection results for poor lighting condition**

Although improper illumination makes color segmentation harder, it has no such adverse effect on the edge information of the image. Therefore, recognition rate of algorithms using shape properties does not reduce. Only in case of very low illumination these algorithms may fail. The average success rate is 92.18%, the result being independent of the number of traffic signs that the image contains. The success rate for each type of shape is shown in Table 1. The lower rate of success in triangle shaped sign with 85.71%

## 5. CONCLUSION

In this paper, detection of traffic signs was studied. The applied method is tested with static images. In order to provide more accurate and realistic results, images taken in different time intervals and different weather conditions are tested. The proposed system is an adaptive and standalone which means that it works without any kind of previous training. Software implementation of the method was conducted using the development environment MATLAB.

## REFERENCES


[1] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, “Speeded-up robust features (surf),” Computer vision and image understanding, vol. 110, no. 3, pp. 346–359, 2008.

[2] J. Matas, O. Chum, M. Urbana, and T. Pajdla, Robust wide-baseline stereo from maximally stable extremal regions," Image and Vision Computing, vol. 22, no. 10, pp. 761{767, 2004.

[3] Canny operator," Image Processing Learning Resources, School of Informatics, The University of Edinburgh, 2003.

[4] C.Y. Fang, S. W. Chen, “Road-Sign Detection and Tracking” Vehicle Technology, IEEE Transactions on Volume 52, Issue 5, Sept. 2003.

[5] G. Piccioli, E. De Michelli, P. Parodi, M. Campani, “A Robust Method for Road Sign Detection and Recognition,” Image and Vision Computing, vol. 14, pp.209-223, 1996.

[6] G. Piccioli, E. De Michelli, P. Parodi, M. Campani, “A Robust Method for Road Sign Detection and Recognition,” Image and Vision Computing, vol. 14, pp.209-223, 1996.

[7] A. Etemadi, “Robust Segmentation of Edge Data,” IEEE Image Processing Conference, Maastricht, 1992.

[8] P. Hough, “A Method and Means for Recognizing Complex Patterns”. US Patent, 3069654, 1962.

[9] S. Masciangelo, “3-D cues from a Single View: Detection of Elliptical Arcs and Model Based Perspective Back Projection,” Proc. British Machine Vision Conference “90, pp. 223-228, Oxford, 1990.

[10] Lindeberg, Tony. Feature Detection with Automatic Scale Selection. International Journal on Computer Vision. 1998, pp. 79-116.

[11] Harris, Chris and Stephens, Mike. A Combined Corner and Edge Detector. ALVEY Vision Conference. 1988, pp.147-151.

[12] Kadir, Timor and Brady, Michael. Scale, Saliency and Image Description. International Journal on Computer Vision. 2001, pp. 83-105.

[13] Jurie, Frederique and Schmid, Cordelina. Scale-Invariant Shape Features for Recognition of Object Categories Computer Vision and Pattern Recognition. 2004, pp. 90-96.

[14] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, “Speeded-up robust features (surf),” Computer vision and image understanding, vol. 110, no. 3, pp. 346–359, 2008.

[15] A. Coates, B. Carpenter, C. Case, S. Satheesh, B. Suresh, T. Wang, D. Wu, and A. Ng, “Text detection and character recognition in scene images with unsupervised feature learning,” in ICDAR, 2011.

[16] T. Wang, D. J. Wu, A. Coates, and A. Y. Ng, “End-to-end text recognition with convolutional neural networks,” in ICPR, 2012.

[17] K. Wang, B. Babenko, and S. Belongie, “End-to-end scene text recognition,” in ICCV, 2011.

[18] H. Chen, S. Tsai, G. Schroth, D. Chen, R. Grzeszczuk, and B. Girod, “Robust text detection in natural images with edge-enhanced maximally stable extremal regions,” in ICIP, 2011.

[19] T. Novikova, O. Barinova, P. Kohli, and V. Lempitsky, “Large-lexicon attribute-consistent text recognition in natural images,” in ECCV, 2012.

[20] C. Merino-Gracia, K. Lenc, and M. Mirmehdi, “A head-mounted device for recognizing text in natural scenes,” CBDAR, 2011.

[21] W. Wu, X. Chen, and J. Yang, “Detection of text on road signs from video,” IEEE Trans. Intell. Transp. Syst., vol. 6, no. 4, pp. 378–390, Dec. 2005.