Traffic Sign Detection Using Template Matching Technique

Pranjali Pandey
Electronics and Telecommunication
V.E.S. Institute of Technology
Mumbai, India
pranjali.pandey24@gmail.com

Dr. Ramesh Kulkarni
Electronics and Telecommunication
V.E.S. Institute of Technology
Mumbai, India
ramesh.kulkarni@ves.ac.in

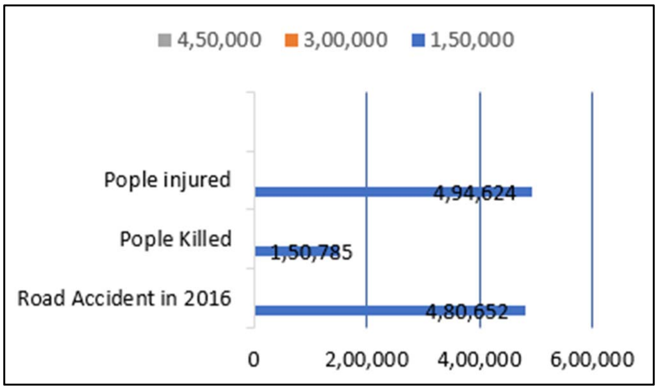
***Abstract*— Advancement in safety features to prevent ignorance of traffic sign boards is a challenge faced by drivers in daily routine. By overlooking traffic road signs, one can lead to road accidents which can be of high-risk for the driver as well as the fellow passengers. Therefore, using Driver Assistance System to assist the driver with mandatory and cautionary road signs can be of important help in prevention of road accidents. The paper explains Template Matching technique of image processing for detecting Traffic signs mounted alongside the road. Here, the object of interest, i.e., traffic sign board, is selected and learned by the system which is then detected where the Region of Interest is in form of green-yellow bounding box. The system is alert about the reappearance of the object, that is, after the departure of the current frame. The output is given both in audio as well as in display message format. This combination of output enables the driver to stay alert while driving the vehicle and prevents any kind of mishap.**

Keywords*— *Driver Safety; Computer Vision; Template Matching ; Traffic Sign Detection

I. Introduction

Cars became a primary means of transport in the 20th century and their number has been dynamically growing since the time they were invented. At the present the motorways and the urban roads in many developed and developing countries are full of vehicles, which has exposed the drivers to various risks. Development of such traffic sign detection systems comes with lots of problems like poor visibility, occlusion, speed of performance, real time execution, variation of traffic signs in different countries [1]. As road signs are an important part of the traffic infrastructure which plays a key role in regulating flow of the vehicles, it was soon found necessary to include in the Driver Support Systems (DSS), as they were often called, the visual Traffic Sign Recognition (TSR) functionality. Survey conducted by Ministry of Road Transport & Highways, India states that there happens 17 deaths and 55 road accidents every hour in 2016, which is one of the highest number of accident in the world [2]. The graph in Table 1 shows the statistics. An estimation carried out by World Bank states that the road accidents occurring in India cost about 3% of its gross domestic product every year.

TABLE I. NUMBER OF ROAD ACCIDENTS IN INDIA IN YEAR 2016 [2]



There are three potential scenarios of a traffic sign recognition system, ordered by the level of interference of TSR in the human activity:

- *The Present-Day Scenario:* Traffic Sign Recognition is used to detect and recognize the traffic signs passed by and to present the appropriate information to the driver. No further actions are taken. This information can have a form of an arbitrary sound, an audio message, or a visual presentation of the detected and interpreted sign's prototype on the dashboard, as well as their combinations.
- *The Near-Future Scenario:* Traffic Sign Recognition system not only detects and recognizes road signs, but is also allowed to trigger certain, limited mechanical actions preventing the vehicle from achieving various dangerous states, and hence ensuring safety of a human.
- *The Future Scenario:* TSR is used to detect and recognize traffic signs and is fully integrated with the other DSS components, which jointly take full responsibility for a vehicle navigation. In this scenario the vehicle would be entirely autonomous and the role of a human would be reduced to merely providing the journey endpoints and possibly some additional requirements to the system, just in a way the GPS is nowadays used.

978-1-5386-5257-2/18/$31.00©2018IEEE

Traffic signs are designed to help people streamline traffic flow, maintain road order and guide with parking requirements to enhance security. Drivers should make appropriate response to different traffic signs such as various mandatory signs, information signs and other signs to secure safety. Traffic signs are designed with such shapes and striking color so as to get easily observed. The information extracted from geometric shape and its color are very useful [3]. These algorithms have to cope with natural and complex dynamic environments, high accuracy demands, and real time constraints. The most challenging issue faced by machine vison-based detection is recognition of objects in natural scenes. This challenge faced by the system is common in every detection and recognition system, especially with the driver support systems, since quick decisions will have an impact on road and driver safety. It is important that these algorithms have a balance between complexity and accuracy keeping the limitations faced due to this conciliation as low as possible. The technique proposed in this paper uses Template Matching which has several methods, these methods are chosen as per trails for the desired application [4]. They are given as follows,

SQDIFF: This method matches the squared difference. Here, a perfect match will be 0 and bad matches will be greater than 0. It is given as,

$$R(x,y) = \sum_{x',y'} (T(x',y') - I(x+x',y+y'))^2 \tag{1}$$

SQDIFF_NORMED: As with Square Difference, a perfect match for Square Difference Normalized will return a 0. It is given as,

$$R(x,y) = \frac{\sum_{x',y'}(T(x',y') - I(x+x',y+y'))^2}{\sqrt{\sum_{x',y'} T(x',y')^2 . \sum_{x',y'} I(x+x',\ y+y')^2}} \tag{2}$$

CCORR: Cross Correlation method multiplicatively match the template against the image, here a good match is large and bad matches will be small or 0. It is given as,

$$R(x,y) = \sum_{x',y'} (T(x',y') \ .\ I(x+x', \quad y+y')) \tag{3}$$

CCORR_NORMED: As with CCORR, an extreme mismatch for Cross Correlation Normalized will return a score near 0. It is given as,

$$R(x,y) = \frac{\sum_{x',y'}(T(x',y') \ .\ I(x+x',y+y'))}{\sqrt{\sum_{x',y'} T(x',y')^2 \sum_{x',y'} I(x+x',y+y')^2}} \tag{4}$$

CCOEFF: Correlation Coefficient matches a template against the image which is relative to its mean, so a perfect match is 1 and a perfect mismatch is –1; and 0 means that there is no correlation. It is given as,

$$R(x,y) = \sum_{x',y'} (T'(x',y') \ .\ I(x+x',y+y')) \tag{5}$$

Where,

$$T'(x',y') = T(x',y') - \frac{1}{(w.h)} . \sum_{x'',y''} T(x'',y'') \tag{5.1}$$

$$I'(x+x',y+y') = I(x+x',y+y') - \frac{1}{(w.h)} . \sum_{x'',y''} I(x+x'',y+y'') \tag{5.2}$$

CCOEFF_NORMED: As with CCOEFF, a relative match for Correlation Coefficient Normalized will return a positive score and a relative mismatch will return a negative score. It is given as,

$$R(x,y) = \frac{\sum_{x',y'}(T'(x',y') . I'(x+x',\ y+y'))}{\sqrt{\sum_{x',y} T'(x',y')^2 \sum_{x',y'} I'(x+x',y+y')^2}} \tag{6}$$

Where; $I\ (x,y)$ = Image pixel value at location (x, y)

$T\ (x,y)$ = Template value of image pixel in location (x, y)

$R\ (x,y)$ = Resultant value in location (x, y)

Where; *Source Image (I):* W*H pixels.

Template Image(T): w*h pixels.

Size of resulting images (R): W- w+1 * H- h+1 pixels.

II. LITERATURE REVIEW

In India, roads are regularly checked for missing or damaged sign boards to avoid any safety threats. This task to investigate the road sign condition is carried out by driving a car on the road of interest and noting problems on the way manually, making it difficult, lengthy, and prone to operator malfunction. The main issue of the traffic sign recognition system does not arise with the detection and recognition of traffic sign in still image, but it is rather, how to obtain a high precision result in a big real-time video data [5]. Overall, the Traffic sign detection methods can be classified into three categories: Color-Based Methods, Shape-Based Methods and Learning-Based Methods. We can apply the best method depending upon the nature of the problem and system requirements; for example, methods based on color information can used with high-resolution dataset, however, not with grayscale images.

A. Color-Based Detection Method

The dominant color-based segmentation is applied to detect regions of interest. Red, blue and yellow are few specific colors in which the traffic sings are there. These characteristics, however, are sensitive to factors, such as, how long the sing have been up and the changing light conditions, making segmentation difficult. In this method simple thresholding or more advanced image segmentation methods are used to find the region of interest. Once the result is obtained, it is then stored as traffic signs. The drawback of this method is that coloration of the image obtained changes depending on the time of day, climate conditions, shadows, and other factors. The various color-based methods proposed by different researchers is discussed. Benallal and Meunier [6] carried out an experiment where the color of a red STOP sign was observed throughout 24 hours and showed that it has more influence between 6:30am to 9pm. Broggi et al. put forward a technique to beat color dominion of light source. Broggi et al. [7] found a piece of road assuming it is gray in color and hence estimating the light source color, which was followed by chromatic equalization, have gamma function linearization. Escalera et al. [8] proposed a method for identifying red in Hue, Saturation, Intensity (HSI) color space. Where the input image is transformed from Red, Green, Blue (RGB) to Hue, Saturation, Intensity (HSI), where for every pixel, hue and saturation value are again calculated to emphasize the domain of saturated hue. These hue and saturation values are then multiplied and upper bounded by 255. The color spaces that are generally used are:

- **RGB Color Model:** The RGB color space model is based on how our eye perceives light and decodes to understand the color. This model works with the addition and subtraction of the colors. The CMYK color model is a subset of the RGB model and it is an acronym for cyan, magenta, and yellow along with black which is noted as K.

- **YUV Color Model:** The YUV color space is derived from the RGB space. It comprises the luminance denoted by Y and the color difference components denoted by U and V. Luminance component is calculated by the sum of red, green, and blue components, whereas, the color difference components, U and V are calculated by subtracting Y from blue and red respectively. The main advantage of this color model is the separation of luminance and color information components.

- **HSV Color Model:** This color model was developed to be more innate in deploying with color to interpret the colors approximately the way we humans perceive them. Here, Hue is the colorfulness, saturation gives the colorfulness of a color relative to its own brightness and value is the visual perception in which the sours appears to be radiating.

B. Shape-Based Detection Method

Shape-based methods are generally robust than colorimetric methods. They can treat grayscale images; for example, in some countries such as Japan where there are few road signs which, when converted to grayscale, appear exactly the same. To be able to distinguish them, an amount of color information is absolutely needed [9].

However, some authors adopt the color feature for identifying the region-of-interest and then they complete the procedure with some shape-based detection methods [5]. Loy and Barnes [10] gave a regular polygon detector for detecting traffic signs. It is based on fast-radial symmetry transform, where the approach has close similarity with Hough transform. Paulo and Correia [9] identify triangular and rectangular signs by the Harris corner detector method, and then look for corners in six pre-defined control areas of the region. Once the approach is carried out, shape of the object is then judged based on the configuration of the control regions where corners are found. Gavrila [11] used distance transform-based template matching method. Here, edges in the original image are found where distance transform (DT) image is built. Basically, it is an image with every pixel representing distance to the closest edge. The idea is to match a template against the Distance Transform image hence obtaining shape of image.

C. Leaning-Based Methods

Viola and Jones [12] introduced a significant milestone in computer vision. They developed an algorithm for detecting objects reliably and in real time where, by using positive and negative sets, training of detector is executed. This algorithm is divided into four stages:

- Haar Feature Selection
- Creating an Integral Image
- Adaboost Training
- Cascading Classifiers

Increasing complexity detectors cascade is used by Viola and Jones [13], where, each detector is a set of classifiers which is based on Haar wavelet, and these classifiers use a learning algorithm called AdaBoost. The detector is a cascade of boosted Haar-like classifiers where it combines AdaBoost and Haar-like classifiers. AdaBoost assigns weights depending on classifier quality, and the resulting classifier is a linear combination of appropriate weights. Viola and Jones group multiple strong classifiers which results into faster processing. In [14] Viola–Jones detector was used to detect triangular traffic sign where 1000 images of relatively poor quality were used for training the detector. The detector achieved high positive rate from 90 to 96% depending on the training set and the configuration. There are two drawbacks of the Viola–Jones detector, first one is that it requires a huge number images for training purpose and the second one is that the rate of false positive is high [14], [16]. Nevertheless, their research also indicates that the Viola-Jones detector is robust to noise and low-quality training data [17], [5]. Histogram of Oriented Gradients (HOG) transformer is also used by Wang et al. [18], and they use Liner Discriminat Analysis (LDA) and Support Vector Machine (SVM) as a classifier.

III. Proposed Method

The system proposed here uses OpenCV library and Emgu files, which is a cross platform .Net wrapper to the OpenCV image processing library, allowing OpenCV functions to be called from .NET compatible languages. The method of detection used is template matching, which is a technique for finding matching areas. The proposed technique is divided into two phases, first training phase and second is testing phase. For the training purpose, traffic signs are imported from Indian Regulatory Road Signs which are freely available. These signs are categorized into Mandatory Signs, Cautionary Signs, Information Signs and Additional Signs. The signs are learned by the system and are saved in a database. Along with these signs, real time traffic scenarios where also captured for accurate detection. There are six methods available for Template Matching, which are Square Difference (SQDIFF), Square Difference Normalized (SQDIFF_NORMED), Cross Correlation (CCORR), Cross Correlation Normalized (CCORR_NORMED), Correlation Coefficient (CCOEFF) and, Correlation Coefficient Normalized (CCOEFF_NORMED). The table below shows the percentage of detection by each of the methods.

TABLE II. COMPARISON OF TEMPLATE MATCHING METHODS

Methods	Matching percentage in good light condition	Matching Percentage in dull light condition
SQDIFF	$\geq$ 39.065	False Detection
SQDIFF_NORMED	False Detection	False Detection
CCORR	Improper Detection	False Detection
CCORR_NORMED	**$\geq$ 99.238**	**$\geq$ 99.001**
CCOEFF	$\geq$ 49.832	$\geq$ 51.678
CCOEFF_NORMED	$\leq$ 99.013	$\leq$ 94.277

From the above table, we can say CV_TM_CCORR_NORMED gives better result in good as well as in dull lighting condition. It is given as, stated in equation (4). The training phase flow diagram is shown in fig below. The template is learned by manually dragging the mouse from left to right and selecting Region of Interest stead of using complete image. Once the template is selected, it is named as per the traffic sign and saved into the database. The working of template matching is explained in the flow chart below. Histogram equalization is used for color correction of the captured frames when the lighting condition is bad.

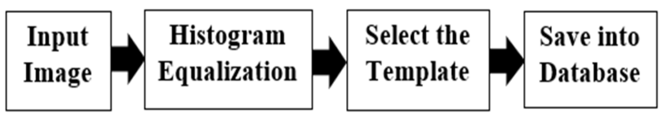


Fig 1: Training Phase

The template matching mechanism carried out in training as well as testing phase is shown in the flow chart below. Once the camera is turned on, the captured frames are taken which are called source images, denoted by I. Next step is to slide this source image over already learned template image, denoted by T. By sliding means, moving patch pixel by pixel in left to right, top to bottom fashion. A metric is calculated at each location to see how alike the patch is to that particular area form source image. For each location of template image (T) over source image (I), the matric is stored in result R matric. Each location (x, y) in R matrix contains the match metric. The brightest locations indicate the highest matches. Once the final highest intensity matric is obtained, the output message is give via voice trigger and display.

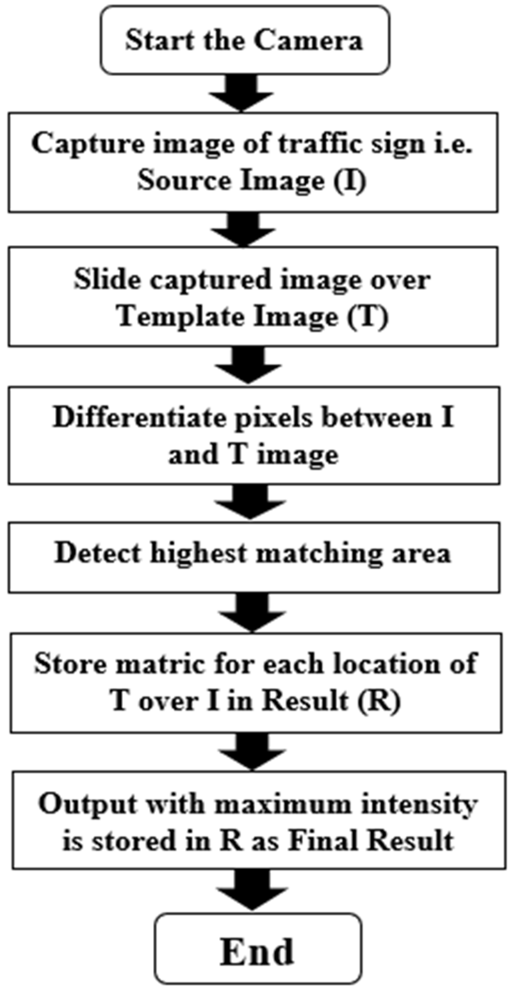


Fig 2: Template Matching Flow Graph

Once the system is trained, the next step is testing phase. The flow diagram is as shown in fig below. Here cross correlation normalized techinique is used as it yeilds better result as shown in table 2.

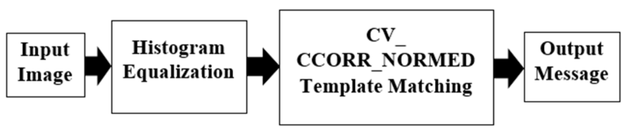


Fig 3: Testing Phase

A. Result

The figure below shows the final user interface that has been developed. The object to be tracked can be seen within the green yellow bounding box. The matching value is displayed on the user interface. A display message and audio message are given as output of the system. Histogram equalization is used for enhancing images in case of bad lighting condition. The tests image used are real time captured images, whereas, the Indian regulatory signs are also used for training purpose of the system.

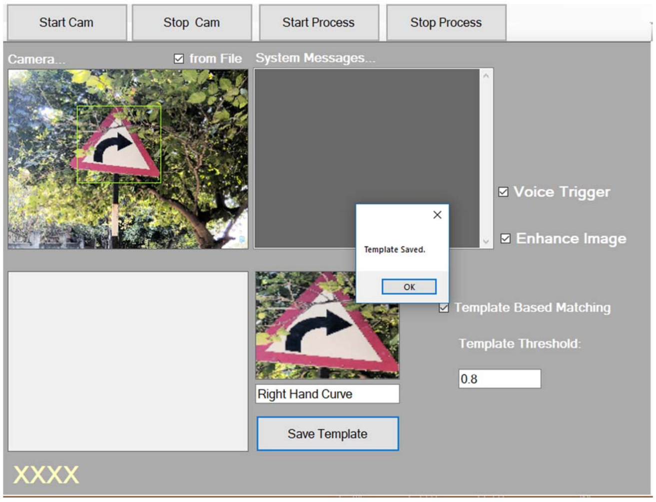


Fig 4: Output of Template Learning

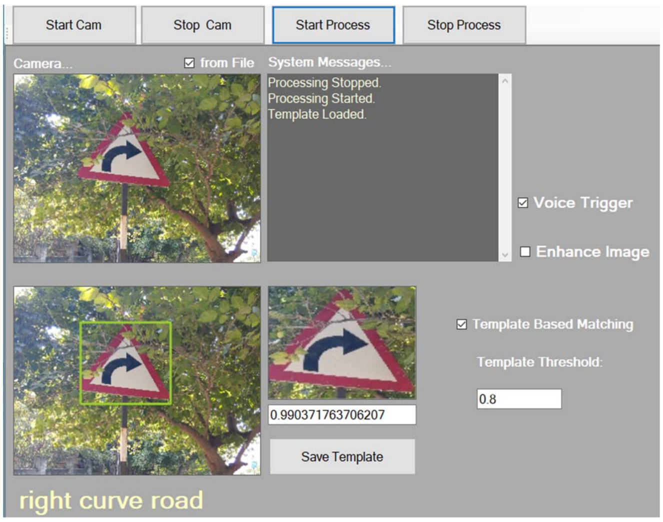


Fig 5: Output of Template Matching

The Table below shows a sample of 12 traffic signs detected in day and night light condition. Each sign is fed 10 times into the system for checking the correctness of detection every time. The camera input used are a mix of freely available Indian traffic signs, and real time captured photos of traffic signs in both day time and night time scenarios.

TABLE III. OUTPUT OF THE SYSTEM WITH MATCHING VALUES

Camera Input	Output	Matching Value
	Left Hand Curve	0.9999
	No Parking	0.9998
	Left Curve Road	Improper Detection
	No Overtaking	0.8736
	No Horn	0.9955
	School	0.8943
	One Way	0.8591

IV. CONCLUSION

From the literature survey in section II, it can be said that real time tracking, learning and detecting of traffic road signs will only improve by time. The proposed system attempted to detect correctly eliminating false positive detection and loss of detection through the use of Template Matching algorithm. The proposed system can detect road signs in difficult light conditions. Further improvement will be made towards better detection accuracy when partial occlusion and low-quality images are taken into account.

REFERENCES

[1] Nilesh J Uke, Ravindra C Thool, Shailaja Uke, "A Vision based Driver Support System for Road Sign Detection", Int. J. on Recent Trends in Engineering and Technology, Vol. 10, No. 1, Jan 2014.

[2] indianexpress.com/article/india/road-accidents-in-india-2016-17-deaths-on-roads-every-hour-chennai-and-delhi-most-dangerous-4837832

[3] Rongqiang Qian, Bailing Zhang, Yong Yue, Frans Coenen, "Traffic Sign Detection by Template Matching based on Multi-Level Chain Code Histogram", in 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), 2015

[4] Vivek Nair, Nadir Charniya, "Driver Drowsiness Detection Using Non-Intrusive Technique", In IOSR Journal of Engineering (IOSR-JECE), vol.12, issue 5, pp.41-19

[5] Yassmina Saadna, Ali Behloull"An overview of traffic sign detection and classification methods", In International Journal of Multimedia Information Retrieval, September 2017, vol.6, issue 3, pp.193-210

[6] M. Benallal and J. Meunier, "Real-time color segmentation of road signs", IEEE CCECE 2003, Canadian Conference on, vol. 3, pp. 1823–1826 vol.3, May 2003.

[7] A. Broggi, P. Cerri, P. Medici, P. Porta, and G. Ghisio, "Real time road signs recognition", Intelligent Vehicles Symposium, 2007 IEEE, pp. 981– 986, June 2007

[8] A. D. L. Escalera, J. M. A. Armingol, and M. Mata, "Traffic sign recognition and analysis for intelligent vehicles", Image and Vision Computing, vol. 21, pp. 247–258, 2003.

[9] C. Paulo and P. Correia, "Automatic detection and classification of traffic signs", in Image Analysis for Multimedia Interactive Services, 2007. WIAMIS '07. Eighth International Workshop on, June 2007.

[10] G. Loy, "Fast shape-based road sign detection for a driver assistance system", in IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2004, pp. 70–75.

[11] D. Gavrila, "Traffic sign recognition revisited", in DAGM Symposium, 1999, pp. 86–93.

[12] P. Viola and M. Jones, "Robust real-time object detection", in International Journal of Computer Vision, 2001.

[13] Viola P, JonesM (2001) "Rapid object detection using a boosted cascade of simple features", Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition, CVPR 2001, vol 1, IEEE, p I

[14] BrkicK, PinzA, Segvic S "Traffic sign detection as a component of an automated traffic infrastructure inventory system", Stainz, Austria 2009

[15] Brki K, Segvi S, Kalafati Z, Sikiri I, Pinz A "Generative modeling of spatio-temporal traffic sign trajectories". In: IEEE computer society conference on computer vision and pattern recognition workshops (CVPRW), 2010, IEEE, pp 25–31

[16] Brkic K (2010) "An overview of traffic sign detection methods. Department of Electronics", Microelectronics, Computer and Intelligent Systems Faculty of Electrical Engineering and Computing Unska, 3, 10000

[17] Wang G, Ren G,Wu Z, Zhao Y, Jiang L "A robust, coarse-to-fine traffic sign detection method". In: The 2013 international joint conference on neural networks (IJCNN), 2013, IEEE, pp 1–5