International Journal of Engineering Research in Electronic and Communication Engineering (IJERECE) Vol 3, Issue 5, May 2016

Assistive Scene Text Recognition

[1]Hymavathi K.P, [2] Fouziya v.k
Department of Electronics and Communication Engineering, KMCT College of Engineering, Calicut, India,
[1]hyma.kp@gmail.com, [2]fauzia457@gmail.com

***Abstract*: Abundance in text along with development of digitalized camera system lead to a new era of recognition and analysis. Traffic sign board which meant for road safety is all time a matter of fact for the traffic safety. The traffic board which carries signs along with informative texts are to be recognized for safe driving. Apart from normal sign board recognition, here the informative boards are also considered. This is a fewer step towards driver assistive system. The scene recognition faces a lot of challenges like motion blur, occlusion and lighting factors. The goal is to implement a system (model) for recognizing the text and signs in the traffic board and to create sound alert. The HOG based vector machine (SVM) classifies the traffic sign and the text is recognized using MSER based OCR. Advantages of the proposed system are well condition under illumination changes sound alert and text recognition is an added advantage with other methods.**

***Index Terms* - Histogram oriented gradients (HOG), Maximal Extremal Region (MSER), occlusion, Sign board recognition**

I. INTRODUCTION

Road safety is always a prime fact for the safety of both drivers and pedestrians. Increasing number of vehicles are a common fact today. Traffic sign recognition is a relevant part of road safety and security of drivers and pedestrians. These signs are meant for safety by advising drivers the possible dangers like hump, work a head and etc. Apart from traffic boards information board is also considered this gives valuable information for the drivers including datas regarding places and other spots like hotels, workshops and etc. Along with sign, text is also to be considered. Traffic sign recognition involves following challenges (1) quality of the image (2) size and corrosion effect of the traffic sign board. Assistive scene recognition proposed in this paper have following parts (1) text recognition (2) sign recognition (3) sound alert system .The recognition process consist of detection and classification of signs. Once the area of interested is obtained using MSER algorithm and WaDe, for text recognition the text normalization is done using normalization algorithm this allows to recognize all font styles and orientation of text and further the text is recognized using OCR.In case of sign recognition the sign is classified using SVM classifier. Before classification inorder to enhance sign of interest from background preprocessing step is done. Candidate region is distinguished using context aware filter. Finally a sound alert is generated which gives drivers warning.

II. RELATED WORK

Many researchers have been done in the field of traffic sign and text recognition. Traffic sign recognition algorithm is sectioned as (1) detection (2) classification. The detection is established on the color based segmentation to extract candidate region.RGB channel thresholding is sensitive to change in illumination.RGB color relationship is used. In [1] color enhancement is done for segmenting candidate blobs according to the color. Blobs are created when segmenting using SVM.There is poor performance in non-illumination condition. In diverse lighting condition intensity decoupling color scheme are preferred [1].Other detection algorithm are done based on edge detection, more robust with change in illumination. Candidate edge orientation histogram is used to detect the features [2]. The RSLD algorithm [3] changes triangle detection to simple segment detection. Coding gradient method is used to detect the corner candidates. Many approaches uses orientation based on gradient in sage of classification done by extracting the edge of traffic boards. Many of the pedestrian detection is based on HOG which is modified to [4] traffic sign detection. In case of scene text recognition word and character recognition [5], [6] giving bottom up and top up information. The machine perception problems especially hand writing recognition are very much related to STR. Most of the approaches require character to be isolated in the image before recognition .In some cases detect the character candidates with sliding window. Others using connected components. CC analysis [6] helps in detecting most of the text formats. Combination of both sign and text recognition along with sound alert is fewer step towards driver assistive system. Several researches are done in recent years to improve the efficiency of the system by change in image quality, lighting effects and so on.

The proposed model recognizes both text and sign in a traffic board with an added advantage of sound alert that provides the driver a warning voice regarding the information. This is implemented using HOG based SVM

that classifies sign.Severel researches have done regarding the illumination changes and other lighting effects. Moreover implementing towards a real time application is another major research regarding sign recognition

III. PROPOSED METHOD

The proposed method is simplest and efficient way recognition traffic sign and scene text. This method consist of following steps: detection, normalization and classification.

A. Image preprocessing

Image preprocessing done to enhance the image for recognition of both sign and text. Preprocessing enhances the color in each sign/text of interest and fades the back ground. The preprocessing is done by contrast stretching, RGB normalization and image enhancement .This improves the quality of image by enhancing specific features of the image after normalization.

(a) (b) (c)

Fig. 1: (a) Original image (2) Contrast stretching (c) detected image

B. Regional grouping

Regional grouping is done extract the interest region .This is done by two approaches Maximally Stable Extremal Region (MSER) and Wave Equation (WaDe).In case of MSER algorithm it detects the interested region by thresholding at different levels. This approach extracts the connected components with stable region across a range of threshold, there by MSER+ and MSER- regions are extracted. Whereas highly symmetrical regions are extracted using WaDe. The false positives are fed to the classifier. In case of text grouping often lexicon gap is to be considered. This is because large gap between consecutive letters, create a problem during recognition

.

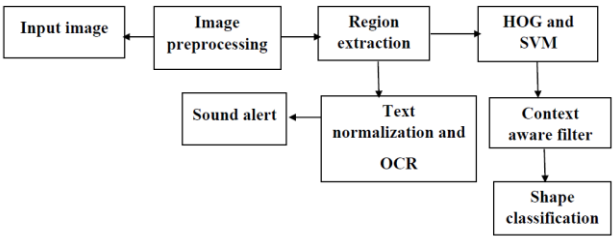

Fig. 2: Block diagram for proposed method

C. Text normalization and OCR

After the extraction of interested region, it is needed to recognize the text candidates. This is done by normalization text by text normalization algorithm. This step allows recognization of all form of texts including different font style and orientation. The region having text is called confidence map. For text reading this is fed to Optical Character Recognization (OCR).The most advanced optical character recognization system are focused over replicating natural recognization. Although OCR faces lot of challenges like blur lexicon gap and so on. This challenges are overcome by the process of regional grouping in which CC analysis is done.CC analysis creates better accuracy for the text recognition.

(a) (b)

Fig. 3 (a) Regional grouping (b) Text with different format

D. HOG and SVM

Object detection is done using feature descriptors like HOG .Histogram of Oriented Gradients capture shape and color as features. Extracted output from HOG is given to SVM, which recognizes and analysis the pattern. Either sign of interest or background is classified by SVM.

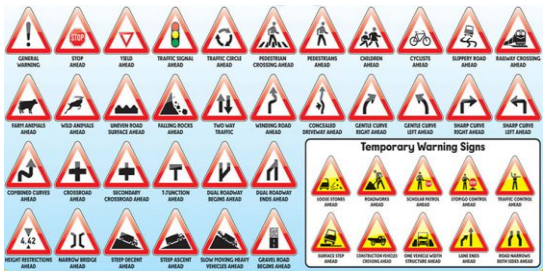

Fig. 4 Indian traffic signs

E. Context Aware Filter

False positives are present in the sign of interest obtained from SVM.This false positives are removed by generative filters. This have built an advantage in traffic sign of urban environment. Size and height are features used to identify sign. The sign size r and height h can be obtained by normalized range [0, 1].The standard deviation for estimated distribution follows parametric law depending upon the independent variable r is expressed as follows:

$$\mu(\gamma) = a\_\mu r + b\mu \quad (1)$$

The wrongly detected false positives are discarded by their scolding. There is no need of negative samples in learning stage of generative approach.

F. Shape Classification

Shape classification is done using Hough transform and verified using peri2area ratio.Peri2area ratio is ratio of square of perimeter to area.

G. Line Detection

Edge, line and curve are detected using a global method Hough transform. This method is not much stable. The general form of line as below.

$$p = x\cos(\theta) + y\sin(\theta) \quad (2)$$

The above equation (2) represent a line passing through (x, y) is perpendicular to the line drawn from the origin (p, θ) in the polar space. Hough transform is also used to represent objects beside line. Generally a circle is represented as

$$r^2 = (x-a)^2 + (y-a)^2 \quad (3)$$

Here (a, b) represents the coordinates of the circle with origin (x, y) and radius r.The number of lines gives the shape of road sign. Three lines represent a triangle, four lines represents either diamond, rectangle or square shape. Eight line represents hexagon and more than eight lines is circle.

H. Peri2area

Verification of the shape is done by peri2area. The total number of pixels in a blob is defined as the area of that blob and the sum of the distances between adjacent pixels is accounted as the perimeter. To determine the shape based on the value of the parameter Peri2Area, if the value is between 9 and 11.75, then the shape is considered an octagon; for values from 11.8 to 14`, it is a circle; from 14.1 to 15.77, it is a pentagon; from 15.78 to 19.14, it is a rectangle; and from 19.15 to 23, it is a triangle. Any value less than 9 or greater than 23 is considered as a non-road sign shape. Mathematical formulas for different geometrical shapes triangle, circle, rectangle and octagonal are used.

Table-I Equation for peri2area calculation

Shape	Formula
Circular	$\frac{(2\ \Pi\ b\ )^2}{\Pi\ b^{\ 2}}$
Triangular	$\frac{(3\ \ b\ \ )^2}{\frac{1}{2}\ bh}$
Rectangular	$\frac{(2\ (b\ +\ h\ ))^2}{bh}$
Octagonal	$\frac{(8b)^2}{\left(2b+\frac{2b}{\sqrt{2}}\right)\frac{b}{\sqrt{2}}+b\left(2b+\frac{2b}{\sqrt{2}}\right)}$

Peri2area ratio gives a clear view of the shape of traffic sign board. Once the sign shape is obtained classifier have limited
datas to analysis.

IV. EXPERIMENTAL RESULTS

The proposed method was tested using a dataset consisting 25 images of each sign in various of lighting conditions and weather conditions. Quality of image is to be considered even if the sign is identified the text is having smaller size which is difficult to recognize. Data set also includes text of different font styles and orientation

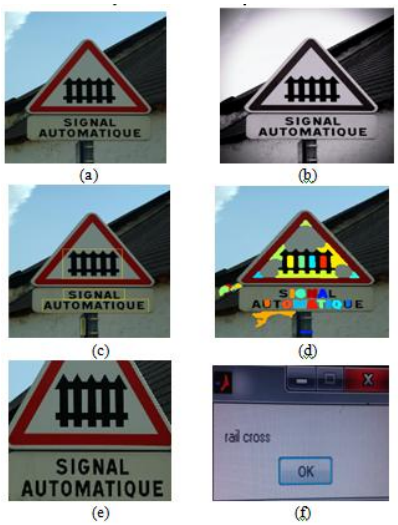

Fig 5: Experimental results (a) Original image (b) Contrast stretching (c) Regional grouping (d) MSER

regions (e) Detected region (f) Recognized sign

Here for this image the detected text is *signal automatique* a warning sound rail cross is generated.

V. CONCLUSION

Assistive scene text recognition which is step towards driver assistive system. This provides the drivers clear idea about the traffic sign. The text recognition also gives the driver a clear information regarding certain locations this will be greater advantage for persons with low visibility.Improoving towards a real time application will be far good. Challenges in text recognition like occlusion and unconstrained lexicons is to cleared so that there will be far better accuracy. Data search, image search are major application of text recognition

ACKNOWLEDGMENT

I would like to thank my project guide Ms. Hymavathi and Head of the Department Ms. Nishida for their guidance and support and also grateful to all the staff members of the Department of Electronics and Communication Engineering of KMCT College of Engineering and Technology, Calicut for their support and for providing all the important facilities like internet access and books, which were essential to carry out the survey. Last but not the least, I would like to thank my family and friends for their whole hearted support and co operation

REFERENCES

[1] A. Møgelmose, M. M. Trivedi. and T. B. Moeslund. Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspective sand survey. IEEE Trans. Intell. Transp. Syst., vol. 13,no. 4, pp. 1484–1497, December2012

[2] F. Zaklouta. and B. Stanciulescu.. Warning traffic sign recognition using a HOG-based K-d tree.in Proc. IEEE IV Symp. pp. 1019–1024.June 2011

[3] J. Greenhalgh, M. Mirmehdi. and S. Member.. Real-time detection and recognition of road traffic signs. IEEE Trans. Intell. Transp. Syst. Vol. 13, no. 4,pp. 1498–1506, December. 2012

[4] K. Elagouni, C. Garcia, F. Mamalet, and P. Se´billot, "Combining Multi-Scale Character Recognition and Linguistic Knowledge for Natural Scene Text OCR," Proc. Int'l Workshop Document Analysis Systems, pp. 120-124, Mar. 2012

[5] M. Boumediene, C. Cudel, M. Basset. and A. Ouamri. Triangular traffic signs detection based on RSLD algorithm. Mach. Vis. Appl. Vol. 24, no. 8, pp. 1721–1732, November. 2013

[6] Yamazoe, M. Etoh, T. Yoshimura, and K. Tsujino, "Towards integerated Scene Text Reading," Proc. Int'l Conf. Document Analysis and Recognition, pp. 359-363, dec.2014