

Real Time Traffic Signboard Detection and Recognition from Street Level Imagery for Smart Vehicle

Aparna A. Dalve
Department of Computer Engineering, MMCOE.
Pune, Maharashtra, India.

Sankirti S. Shiravale
Department of Computer Engineering, MMCOE.
Pune, Maharashtra, India.

ABSTRACT

The text or symbol detection and recognition from traffic panels is a challenging problem. Number of important application areas is dependent on text detection and recognition, including advanced driver assistance systems, road surveying, and autonomous vehicles. In this research project a novel system for the automatic detection and recognition of text and symbol in traffic signs is proposed. Search regions within the image must be defined. In this particular region locate a large number of candidates, which are then reduced by applying constraints based on temporal and structural information. This problem can be divided in two stages; First stage will be detection of region and second will be character recognition. The detection stage exploits knowledge of the structure of the scene, the size and location of the road in the frame. Once a potential traffic panels has been located, the next stage attempts to recognize text and symbols within the region. For the purpose of text detection MSER is used and for recognition purpose optical Character Recognition method is used. Automatic testing using XML files provide better accuracy.

General Terms

Image Processing, Text Detection, Text recognition, Automatic Testing using XML files

Keywords

Text detection, Text recognition, maximally stable extremal regions (MSERs), Optical Character recognition (OCR)

1. INTRODUCTION

Each government imposes some sets of rules and regulations to ensure a safe traffic system. Each person specially the vehicle driver must obey these rules and regulations for a secure travel. Some of those laws are represented as visual language such as different signs and texts that are known as traffic signs. There are various categories of traffic signs that we can see beside the roads. An efficient driver must notice each of the road signs in front of him and need to act accordingly. Otherwise disastrous things can happen. A driver may not notice each of the road signs in front of his car due to lack of care or human perception errors. As shown in Fig. 1 there is a huge Percentage of road accidents in India. Therefore, it is desirable of having a automatic road sign detection and recognition system to assist the driver to ensure a safe travel.

Estimated Deaths Due To Road Accidents In India By Age Group

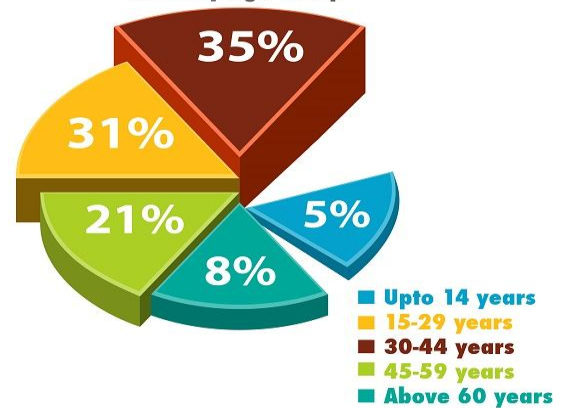


Fig. 1 Accidents in India

Automatic detection of road sign is a challenging but demanding job. Text detection and recognition system have a number of important application areas, including advanced driver assistance systems, road surveying, and autonomous vehicles. Automatic detection and recognition of traffic signs is a challenging problem. While much research exists on both the automatic detection and recognition of symbol-based traffic signs, and the recognition of text in real scenes there is far less research focused specifically on the recognition of text on traffic information signs. This could be partly due to the difficulty of the task caused by problems, such as illumination and shadows, blurring, occlusion, and sign deterioration. Main task in automatic text detection and recognition system is how to identify traffic signs from no traffic signs. Without the use of additional temporal or contextual information, there is few information to determine traffic signs from no traffic signs on the y, while driving, other than basic features, such as shape or color.

As shown in the Fig. 2 there are many traffic signboards present in India which are text based. It is impossible to notice all of them manually in a moving vehicle. Chances of missing very important signboards will be high. Manual capturing of signboards also diverts driver's attention. This may cause many accidents on Indian roads. To prevent accidents advanced driver assistance system plays an important role.

Traffic signboard detection and recognition have been an important issue for research recently. Traffic signs have a dual role: First, they regulate the traffic and, second, indicate the state of the road, guiding and warning drivers and pedestrians.

System can be comprises of two main stages:

1. Detection
2. Recognition

Detail explanation of these two stages is given in next sections.



Fig. 2 Traffic Signboards

2. RELATED WORK

Numerous research works have been conducted for automatic detection and recognition of road signs in order to assist the driver. More research works are being conducted on the issue, because of its enormous potential in practical traffic control system applications. Early text detection and recognition research was a natural extension of document analysis and recognition research, moving from scanned page images to camera captured imagery, focusing on basic preprocessing, detection and OCR technology. Recently, the application of sophisticated computer vision and learning methods has resulted from the realization that the problems do not lend themselves to a sequential series of independent solutions. The trend is to integrate the detection and recognition tasks into an end-to-end text recognition system. Most of the research exists on the detection and recognition of text in natural scenes. To solve these problems main two approaches are present. Approaches to this problem can be broadly divided into two Parts: 1) Region-based methods and 2) Connected component (CC) - based methods. Region-based text detection methods use local features, such as texture, to locate text regions, whereas CC-based methods attempt to segment text characters individually by using information such as intensity, color distribution, and edges. They usually consist of three phases: a first stage to detect CCs within the image, a second stage to eliminate unlikely CCs based on their features, and a final stage that attempts to group the remaining CCs into words or lines. Main purpose of this study is to get text from traffic signboards. More relevant to the context of this paper, the amount of research focused specifically on the detection of text within traffic signs is limited, due to the difficulty of the task.

Many methods for text detection and recognition from Traffic signboards have been proposed over the past years; we will brief review some of these methods and the comparative study of these methods in next section.

2.1 Text Detection Methods

1. Connected Components based method

The method consists of two steps. The first step is to draw CC from images using a specific method and the second step is to estimate whether the CC is text CC or not based on CC feature and CC relative feature.

2. Sliding window based method

Sliding window based methods, also known as region-based methods, use a sliding window to search for possible texts in the image and then use machine learning techniques to identify text. These methods are slow as the image has to be processed in multiple scales.

3. Hybrid method

The hybrid method presented by Pan et al exploits a region detector to detect text candidates and extracts connected components as character candidates by local binarization; non-characters are eliminated with a Conditional Random Fields model, and characters can finally be grouped into text.

4. Edge based method

This method is based on the factor like edge of character; edge is reliable feature of the text regardless of color/intensity, layout, orientations, etc. As the text region has high contrast to its background, the edges of character can be easily detected. There are two steps used in this method: first, an edge extraction algorithm (such as canny edge detector) is used to get the edges and second, smoothing algorithm or morphology is used for edges connections and obtaining a full character boundary. The main disadvantage of this method is that small image regions and stroke may be misidentified. Therefore this method needs to be verified using other methods.

5. Color based method

In this method, color clustering is done by categorizing the pixels with same or similar colors and forming a candidate region. Then the candidate regions are analyzed and the CC is estimated. The main challenge of this method is the degree of clustering. If the data is over clustered, the background and text region may be mixed together. And if the data is under clustered, the number of clustering will be increased and the system performance will be degraded.

6. Combination of edges and colors

Some methods combine Method 1 and Method 2, which detects both edges and colors of the text. This method has achieved better results by combining both features together than using these features separately.

7. Texture based method

This method deals with text regions as a special texture. The region is identified as text region or not according to the extracted relevant texture of the candidate regions. To overcome the disadvantages mentioned above, hybrid approach is presented, which takes the advantages of both texture-based and CC-based methods, to robustly detect and localize texts in natural scene images. In this method, a text region detector is designed which is based on the texture. This can be used to estimate the probabilities of the position and the scale of the text and then it is analyzed to be text region or not.

8. Corner based method

This approach is inspired by the observation that the characters in the text, usually contains multiple corner points. The method is to describe the text regions formed by the corner points using several discriminative features. The research on the method based on corners is still in the early stage. Compared with texture based method, this method is faster but the performance is less satisfied.

9. Stroke based method

As a basic element of text strings, strokes provide robust features for text detection in natural scene images. Text can be modeled as a combination of stroke components with a

variety of orientations, and features of text can be extracted from combinations and distributions of the stroke components. One feature that separates text from other elements of a scene is its nearly constant stroke feature like stroke width. This can be utilized to recover regions that are likely to contain text. For stroke-based methods, text stroke candidates are extracted by segmentation, verified by feature extraction and classification, and grouped together by clustering. These methods are easy to implement on specific applications because of the intuition and simplicity. However, complex backgrounds make text strokes hard to segment and verify.

10. Semiautomatic Ground Truth Generation method

The semiautomatic ground truth generation system for text detection and recognition includes text with different orientation and language. In this method, the system allows user to manually correct the ground truth if the automatic method produces incorrect results. This method uses eleven attributes at the word level, namely: line index, word index, coordinate values of bounding box, area, content, script type, orientation information, type of text (caption/scene), and condition of text(distortion/distortion free), start frame, and end frame to evaluate the performance of the method.

2.2 Text Recognition

It is a process of electronically extracting text from images. OCR (optical character recognition) is the recognition of printed or written text characters by a computer. This involves photo scanning of the text character-by-character, analysis of the scanned-in image, and then translation of the character image into character codes, such as ASCII, commonly used in data processing. Optical Character Recognition, or OCR, is a technology that enables you to convert different types of documents, such as scanned paper documents, PDF files or images captured by a digital camera into editable and searchable data.

2.3 Comparative Study

Comparison of MSER with other region detectors: Six region detectors are studied (Harris-affine, Hessian-affine, MSER, edge based regions, intensity extremal, and salient regions).

- Region density: MSER detects about 2600 regions for a textured blur scene and 230 for a light changed.
- Region size: MSER tended to detect many small regions, versus large regions which not cover a planar part of the scene.
- Viewpoint change: MSER outperforms the 5 other region detectors in both the original images and those with repeated texture motifs.
- Scale change: Following Hessian-affine detector, MSER comes in second under a scale change and in-plane rotation.
- Blur: MSER proved to be the most sensitive to this type of change in image, which is the only area that this type of detection is lacking in.
- Light change: MSER showed the highest repeatability score for this type of scene, with all the other having good robustness as well.

Table 1 Comparison of Text Detection and Recognition Methods

No.	Authors	Method	Comment
1.	Jack Greenhalgh and Majid Mirmehdi	MSER hue, saturation, and value color thresholding OCR	Search area for traffic signs was reduced
2.	Qixiang Ye David Doermann	Survey	OCR typically achieves recognition rates higher than 99% on scanned documents
3.	Jack Greenhalgh and Majid Mirmehdi, Senior	Histogram of oriented gradient features, maximally stable extremal regions ,support vector machines	Significantly insensitive to variations in illumination and lighting conditions
4.	Luis M. Bergasa, and J. Javier Yebe	Bag of visual words (BOVW)	Images are represented as a “bag of visual words” and classified using Naïve Bayes or support vector machines.
5.	Jack Greenhalgh and Majid Mirmehdi	MSER, HOG	Robust to variations in lighting and illumination
6.	Andreas Møgelmoose, Mohan Manubhai Trivedi, and Thomas B. Moeslund	Survey of all methods	Provided an overview of the state of sign detection
7.	Saturnino Maldonado-Bascón,, Sergio Lafuente-Arroyo, Pedro Gil-Jiménez, Hilario Gómez-Moreno,, and Francisco López-Ferreras	Hue saturation intensity (HSI), support vector machines (SVMs)	Traffic-sign detection by shape classification using linear SVMs. Content recognition based on Gaussian-kernel SVMs.

3. PROPOSED METHOD

3.1 Problem Statement

To design a system which detect and recognize text and symbols from traffic panels from street level imagery. A lot of text-based traffic signboards are mounted on the road side. Manual capturing of all the signboards is difficult. Current Automatic Signboard Recognition Systems (ASRs) are mostly symbol based. Requirement is to research more on text-based ASR.

3.2 Proposed System Architecture

System for traffic signboard detection and recognition consists of main two parts: detection and recognition. Sub parts of system are sign detection and text detection. System is capable of detecting both sign as well as text at a time. There are main two modules to design this system: First one is sign detection and recognition module. And second is Text detection and recognition module. Main focus is given for text detection and recognition. More research is for getting more appropriate and ancient method to detect and recognize text from traffic signboards. This job is challenging because of many distortions arises during the image capture procedure. This system provides a real time application for automatic traffic signboard detection and recognition system. As shown in Fig.3 Input image is taken for processing by system. There are two parts both will work simultaneously. One is Symbol detection and recognition and second one is text detection and recognition. The first stage of the proposed system detects candidates for text-based traffic signs. This consists of three phases: determination of search regions (regions of interest where the text sign is expected to be found), detection of all possible candidates within these regions, and reduction of candidates using contextual constraints. In this way we detect region of symbol and text. On next step we have to recognize text and symbols. For Text recognition Optical Character Recognition System is used. For the purpose of Symbol recognition MSER is useful method. Automatic testing is used to find out the accuracy of output. Finally output with accuracy percentage will be provided by this system. In Fig.4 shows a fragment of XML used to markup the data captured. The root element is tagset and consists of a sequence of image elements – one for each image in the dataset. The imageName element gives the relative path to the image file, And the resolution element gives the width and height of the image. The taggedRectangle element contains a taggedRectangle element for each word in the image. The x, y, width and height. Attributes specify the location (top left corner) and size of the word, while the offset and rotation Specify the slant (e.g. for italicized text) and rotation of the word. The text of the word is given as the body of the tag element. The segmentation element specifies a sequence of offsets for each character segmentation point in the word.

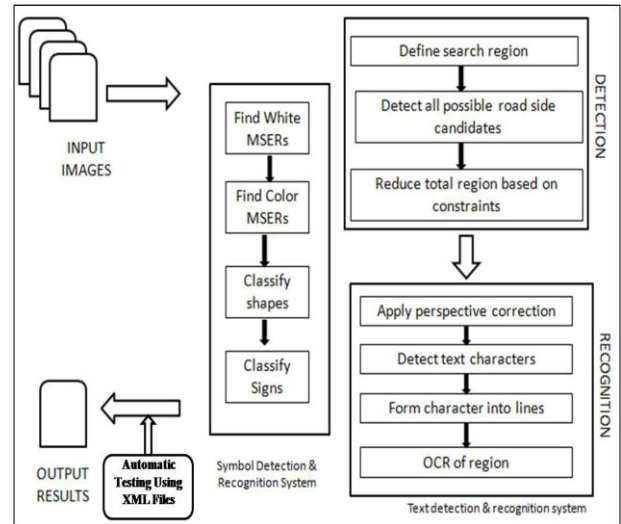


Fig 3: Architecture for Text and symbol Detection Recognition System



Fig 4: A sample of XML format for marking up the words in images.

3.3 Proposed Algorithm

The algorithm will have two phases- The Detection and the Recognition

3.3.1 Detection Phase

For the purpose of detection of text from traffic signboard following steps is followed:

- Preprocess input image
- Detect all possible road sign candidates
- Reduce total number of candidates by applying contextual and temporal information of traffic panels.

3.3.2 Recognition Phase

A) Recognition of Text from Traffic signboards:

- Perspective correction
- Detect text characters
- Form characters into lines
- OCR of region

B) Symbol Recognition from Traffic Signboards:

- Find White MSERs
- Find color MSERs
- Classify shapes
- Classify signs

C) Automatic Testing Of Output Using XML Files

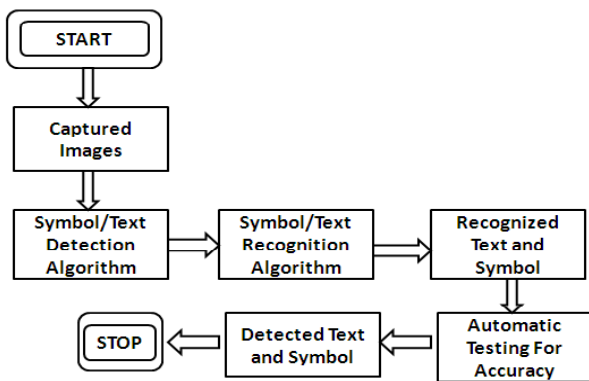


Fig 3: Algorithm for Traffic signboard detection and recognition system

4. INPUT AND EXPECTED OUTPUT

The proposed work is in the phase of pre- processing; hence no finite results can be shown. In this paper we present the idea of input and the expected output. In the training phase of the algorithm, the input images will consist of the training dataset images captured. The input images will have both text as well as symbols. All traffic panels from road sides are captured. These images will be formed and given as an input to the algorithm. The algorithm will be trained to detect and recognize text and symbols from traffic signboards.

For the testing phase, XML files are used. All the input images are stored in the form of XML files. Automatic texting of results will be done using XML files. The input test image will be divided into no of parts; these parts are stored in the XML files.

The output of this project work will be the detection of text as well as symbols from real time captured images of traffic signboards. The proposed system will automatically test the output of previous modules and able to show the accuracy percentage for better understanding of traffic signboards for advanced driver assistance system.

5. CONCLUSION

This text and symbol detection system with MSER enables to detect most characters from the traffic signboard images. The MSER method is proven better results as compare to various techniques, because it easily finds the extremal regions from the input image. By applying constraints based on temporal and structural information candidates are reduced. This gives better accuracy. Perspective recitation and temporal fusion of candidate regions of text were used to improve OCR results. OCR provides better results for text recognition.

6. REFERENCES

[1] Jack Greenhalgh and Majid Mirmehdi “ Recognizing Text-Based Traffic Signs” IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 16, NO. 3, JUNE 2015

[2] Qixiang Ye, David Doermann ” Text Detection and Recognition in Imagery: A Survey” IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 37, NO. 7, JULY 2015

[3] A. González, L. Bergasa, and J. Yebes, “Text detection and recognition on traffic panels from street-level

imagery using visual appearance,” IEEE Trans. Intell. Transp. Syst., vol. 15, no. 1, pp. 228–238, Feb 2014.

[4] F. Zaklouta and B. Stanculescu, “Real-time traffic-sign recognition using tree classifiers,” IEEE Trans. Intell. Transp. Syst., vol. 13, no. 4, pp. 1507– 1514, Dec. 2012.

[5] J. Greenhalgh and M. Mirmehdi, “Traffic sign recognition using MSER and random forests,” in Proc. EUSIPCO, Aug. 2012, pp. 1935–1939.

[6] J. Greenhalgh and M. Mirmehdi, “Real-time detection and recognition of road traffic signs,” IEEE Trans. Intell. Transp. Syst., vol. 13, no. 4, pp. 1498–1506, Dec. 2012.

[7] A. Møgelmoose, M. M. Trivedi, and T. B. Moeslund, “Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey,” IEEE Trans. Intell. Transp. Syst., vol. 13, no. 4, pp. 1484–1497, Dec. 2012.

[8] M. A. García-Garrido et al., “Complete vision-based traffic sign recognition supported by an I2V communication system,” Sensors, vol. 12, no. 2, pp. 1148– 1169, Jan. 2012.

[9] Ahmed Hechri, Abdellatif Mtibaa "Automatic Detection and Recognition of Road Sign for Driver Assistance System" IEEE 2012

[10] Fatin Zaklouta , Bogdan Stanculescu "Real-time tra_c sign recognition in three stages" Elsevier 2012

[11] Jesmin F. Khan, Sharif M. A. Bhuiyan, and Reza R. Adhami "Image Segmentation and Shape Analysis for Road-Sign Detection" IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 12, NO. 1, MARCH 2011

[12] Ian Sebanja, D. B. Megherbi "Automatic Detection and Recognition of Tra_c Road Signs for Intelligent Autonomous Unmanned Vehicles for Urban Surveillance and Rescue" IEEE 2010

[13] Saturnino Maldonado-Bascn, Member, IEEE, Sergio Lafuente-Arroyo, Pedro Gil-Jimnez, Hi-lario Gmez-Moreno, Member, IEEE, and Francisco Lpez-Ferreras "Road-Sign Detection and Recognition Based on Support Vector Machines" IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 8, NO. 2, JUNE 2007

[14] A. VZQUEZ REINA, R. J. LPEZ SASTRE, S. LAFUENTE ARROYO, P. GIL JIMNEZ. "Adaptive tra_c road sign panels text extraction" Proceedings of the 5th WSEAS Int. Conf. on Signal Processing, Robotics and Automation, Madrid, Spain, February 15-17, 2006 (pp295-300)

[15] Wen Wu, Xilin Chen and Jie Yang "Detection of Text on Road Signs From Video" IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 6, NO. 4, DECEMBER 2005

[16] S. M. Lucas, A. Panaretos, L. Sosa, A. Tang, S. Wong and R. Young "ICDAR 2003 Robust Reading Competitions" Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003)