ISSN: 2277-9655
(I2OR), Publication Impact Factor: 3.785

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

ENHANCED TECHNIQUES TO DETECT, RECOGNIZE AND EXTRACT TEXT ON ENGLISH TRAFFIC PANELS ALONG ROAD SIDE

Ms. Pallavi R. Vyas*, Prof. Vinay S. Kapse
M. Tech. (2nd Year) CSE,Vidarbha Institute of Technology, Nagpur, India
Asst. Professor, Dept. Of CSE, Vidarbha Institute of Technology, Nagpur, India

ABSTRACT

Traffic Panel Detection and Text Extraction is an important issue in road maintenance system. This is also important because it provides effective and secure medium to drivers to drive safely. Traffic Panels differ in appearance because of different size and style of text written on Traffic Panels. Language is also a problem for drivers. As language differs as we travel country to country and throughout the world. Detection of Traffic Panel also becomes difficult due to different variety of text appearance because of different fonts, thicknesses, colors, sizes, textures, as well as the presence of geometrical distortions and partial occlusions in the images, different lighting conditions and image resolutions, etc. This paper mainly deals with English Traffic Panels Detection and Extraction of text written on them. Firstly it detects a Proper English Traffic Panel from all the panels arriving along a road side and in second part it then extracts text from that panel. Finally, Text Detection and Recognition method is applied on those images where a traffic panel has been detected. Text detection and recognition in images is a system which attempts to develop a computer system with the ability to automatically read from images the text content visually embedded in complex backgrounds. This extracted text can then be used to help drivers to drive considering the safety issues.

KEYWORDS: Panel Detection, Text Recognition, Text Extraction, Different Text Pattern, Complex Background

INTRODUCTION

Traffic Panels are markers placed along roads to inform drivers about road conditions and restrictions or which direction to go. They communicate a wealth of information but are designed to do so efficiently and at a glance. This also means that they are often designed to stand out from their surroundings, making the detection task fairly well defined. The designs of traffic signs are standardized through laws but differ across the world.

Nowadays, improving road safety is a key matter for road network management. Any company, organization, or institution responsible for the management and operation of any road network should be able to diagnose the problems related to the safety of drivers, set up a joint action plan, coordinate efforts among all the involved organizations, assign funds and resources, supervise the implementation of the action plan, and evaluate the effectiveness of the taken measurements. When the system is considered a secured system where the driver is an integral part, it allows for the driver to contribute what he is good at (e.g., seeing speed limit, identifying traffic signs etc).

However, text from traffic panels is difficult to be detected and recognized due to their various sizes, grayscale values and complex backgrounds. This technique investigates methods for building an efficient application system for detecting a proper traffic panel and recognizing text of any grayscale values embedded in images. This technique is mainly designed for English Traffic Panels. Once traffic panel is identified then both empirical image processing methods and statistical machine learning and modeling approaches are studied in two sub-problems: text detection and text extraction.

ARCHITECTURE OF PROPOSED MODEL

To achieve the proposed technique the following architecture is designed. Whenever any vehicle travel, all the traffic panels can be captured using camera, or mobile phone or any input image capturing device attached to the system in the vehicle. Scanned image then goes through the

following process and text on proper traffic panel is extracted.

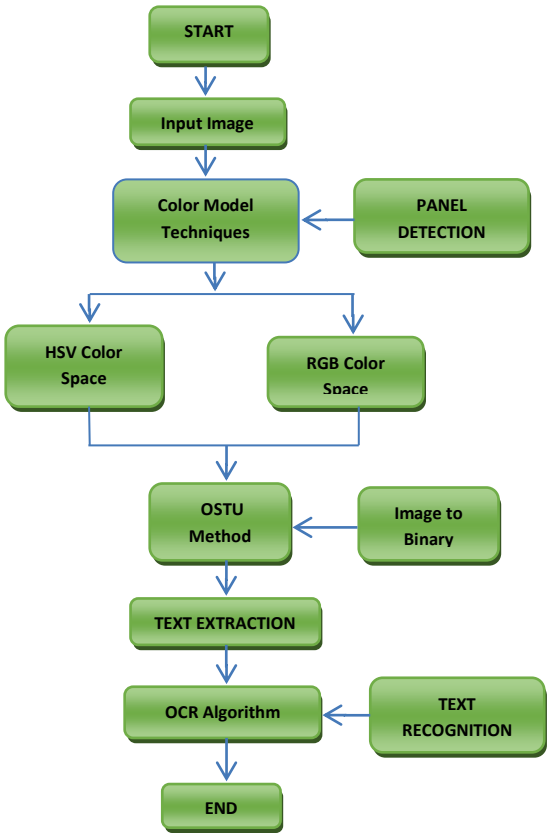

Figure 1. Architecture of Traffic Panel and Detection and Text Extraction.

The proposed model basically use the techniques of image Extraction and text recognition and text extraction.Proper Traffic Panel Detection considers all features of detection and extraction.

METHODOLOGY USED

COLOR SEGMENTATION

One major task of pattern recognition, image processing, and related areas: is to segment image into homogenous regions. Image segmentation is the first step towards image understanding and image analysis. Image segmentation has been acknowledged to be one of the most difficult tasks in computer vision and image processing. Image segmentation can be performed with the help of color segmentation.

Unlike other vision tasks such as parametric model estimation, fundamental matrix estimation, optical flow calculation etc., there is no widely accepted model or analytical solution for color image segmentation. There probably is no "one true"segmentation acceptable to all different people and under different psychophysical, conditions.

A lot of image segmentation methods have been proposed: These methods can be classified into: (1) Histogram thresholding (2) Clustering (3) Region growing (4) Edge-based (5) Physical model- based (6) Fuzzy approaches and (7) Neural network methods

Clustering techniques identify homogeneous clusters of points in the feature space(such as RGB color space, HSV color space, etc.) and then label each cluster as a different region. HSV color Space and RGB color space

Color image segmentation is useful in many applications.From the segmentation results, it is possible to identify regions of interest and objects in the scene, which is very beneficial to the subsequent image analysis or annotation. Recent work includes a variety of techniques: for example, stochastic model based approaches, morphological watershed based region growing, energy diffusion, and graph partitioning. Quantitative evaluation methods have also been suggested. However, due to the difficult nature of the problem, there are few automatic algorithms that can work well on a large variety of data.

The problem of segmentation is difficult because of image texture. If an image contains only homogeneous color regions, clustering methods in color space such as are sufficient to handle the problem. In reality, natural scenes are rich in color and texture. It is difficult to identify image regions containing color-texture patterns.

The approach taken in this work assumes the following:

• Each region in the image contains a uniformly distributedcolor-texture pattern.

• The color information in each image region can be represented by a few quantized colors, which is true for most color images of natural scenes.

• The colors between two neighboring regions are distinguishable - a basic assumption of any color image segmentation algorithm.

1. OCR

Optical Character Recognition is a process that can convert text, present in digital image, to editable text. It allows a machine to recognize characters through optical mechanisms. The output of the OCR should ideally be same as input in formatting. The process involves some pre-processing of the image file and then acquisition of important knowledge about written text. That knowledge or data can be used to recognize characters.

ISSN: 2277-9655
(I2OR), Publication Impact Factor: 3.785

Figure 2 shows stages in OCR Design. Input given to OCR system is a Scanned Image. In this proposed the input image is a traffic panel.

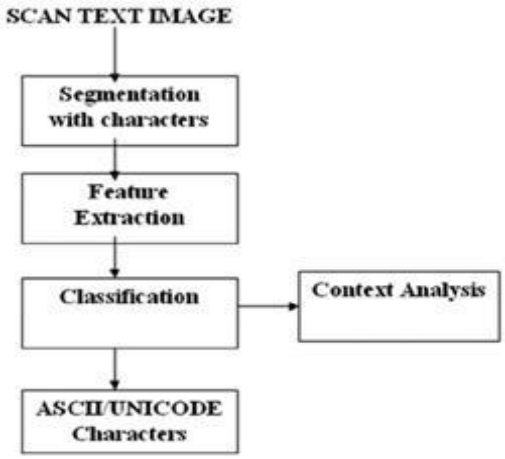

Figure 2. Design of OCR

Various stages of OCR system design are:-

1. Segmentation with characters
2. Feature Extraction
3. Classification
4. Context analysis
5. ASCII/ UNICODE Characters

Segmentation: This is important operation of OCR as rate of recognition is directly proportional to segmentation. In this process, every individual character is separated. This isolates the different sub-parts of an image. It is used to separate pixels of an image as per the contents in data like words, paragraph etc.

Feature Extraction: For the accuracy of OCR system, the appropriate Feature Extraction method should be selected. While processing over the image some features should be separated. The typical features are Edges, Corners, Ridges, etc. This method of separation is called as Feature Extraction. The accuracy of an OCR technique depends on selection of proper feature extraction method.

Classification &Context Analysis: According to extracted features. Text is classified in proper text category. And the extracted text is identified.Extracted text is then arranged and can be recorrected with context analysis.

ASCII/UNICODE Characters: Extracted Text are represented according to their ASCII/UNICODE values.

IMPLEMENTATION

1. PANEL DETECTION

Panel Detection is the first step in Text Extraction on Traffic Panel. There are several methods used for Panel Detection. Panel Detection basically helps to identify accurate traffic panel from all panels coming across street. The proposed paper work for Green Color and Yellow Color Traffic Panel. Green Traffic Panel has white text depicted on it. And Yellow color panel has black text on it.

Color Segmentation is being used in the proposed technique. One major task of pattern recognition, image processing, and related areas: is to segment image into homogenous regions. Image segmentation is the first step towards image understanding and image analysis. Image segmentation has been acknowledged to be one of the most difficult tasks in computer vision and image processing. Image segmentation can be performed with the help of color segmentation.

Unlike other vision tasks such as parametric model estimation, fundamental matrix estimation, optical flow calculation etc., there is no widely accepted model or analytical solution for color image segmentation. There probably is no "one true" segmentation acceptable to all different people and under different psychophysical, conditions.

Some Snapshots of Traffic Panel Detection in MATLAB 2015a are shown below. Figure 3 shows Image captured by camera and is processed to Detect Proper Traffic Panel and Figure 4 shows the detected area of green traffic panel.

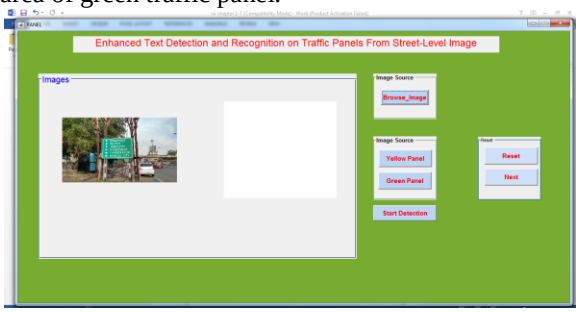

Figure 3. Original Input Image

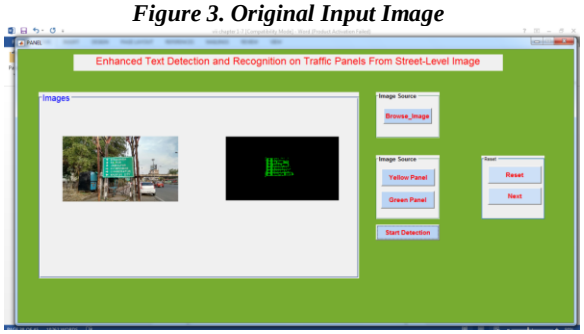

Figure 4. Detected Traffic Panel

2. TEXT RECOGNITION AND TEXT EXTRACTION

Text Recognition is the next phase before Text Extraction. Text Recognition allows text extraction easier. Proposed approach uses OCR Technique for

text extraction. Optical character recognition is the electronic conversion of binary images of printed as well as handwritten text into machine encoded text. It is mostly used as a type of data entry from the printed text documents, whether invoices, passport documents, computerized receipts, bank statements, printouts of static-data, mail, business cards. The output of the OCR should ideally be same as input in formatting.

The process involves some pre-processing of the image file and then acquisition of important knowledge about written text. That knowledge or data can be used to recognize characters. OCR is becoming an important part of modern research based computer applications. Especially with the advent of Unicode and support of complex scripts on personal computers, the importance of this application has increased.

Optical Character Recognition is classified into two types: Offline recognition and Online recognition. In offline recognition the source is either an image or a scanned form of the document whereas in Online recognition the successive points are represented as a function of time and the order of strokes are also available.

To detect text, object segmentation is applied and object extraction for the separation of the entire object which in kind of text and image. Then convert the colored (rgb) image into grayscale using rgb2gray function. Then the gray threshold for gray intensity is calculated and coverted to black and white using im2bw function. The the image is labeled using bwlabel with 8-way connectivity i.e. to find all objects in image. Then plot the bounding box around those objects which comes under above condition only assume that it would be the text.

Figure 5 shows Text Recognition on selected and detected Traffic Panel.

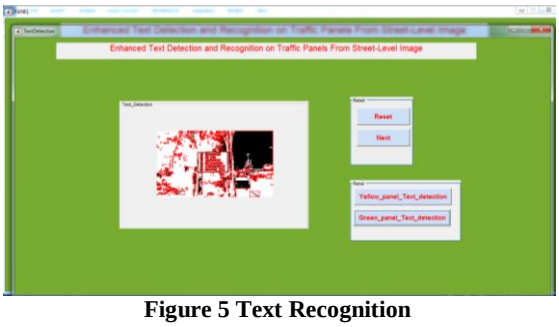

Figure 5 Text Recognition

Finally OCR Technique displays the extracted text on the final window. Figure 6 shows Extracted Text on Detected Green Traffic Panel.

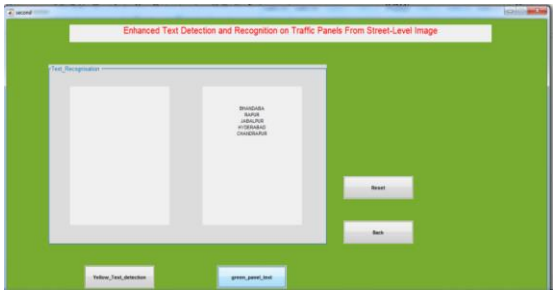

Figure 6 Text Extraction

RESULT ANALYSIS & DISCUSSION

In below analysis Table 1 and graph, the results are shown of accuracy shown for different input images in proposed approach. Number of Total Words and Number of Output Words detection ratio is recorded. And total accuracy level achieved in proposed system is 84.16 %.

IMAGES	TOTAL WORD	OUTPUT WORD
image1	3	3
image2	5	5
image3	14	13
image4	7	4
image5	2	1
image6	11	5

Table 1 Performance of Proposed Approach

Figure 7 shows performance and accuracy level of proposed technique of Enhanced Techniques To Detect, Recognize and extract text on english traffic panels along road side

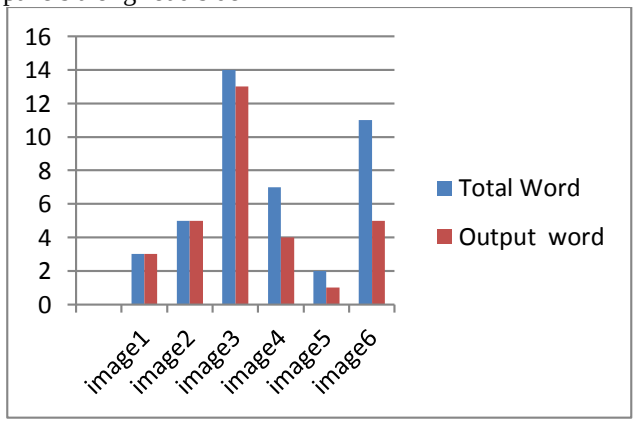

Figure 7 Graph Representing Accuracy level of proposed approac

CONCLUSION AND FUTURE SCOPE

Traffic Panel Recognition and Text Detection is a technique to detect English Traffic Panels from all panels arriving the road side and extract text written on that traffic panels. The proposed approach is a real life application which provides an automatic inventory to drivers and helps them to drive safely. Road maintenance is also possible with proposed approach to improve road and traffic systems. This project uses color segmentation to identify panels of Green and Yellow color. And then use OCR method to extract text. Green Panel contains White color text whereas Yellow Panel Contain black color text.Traffic Panel Recognition and detection is useful for identifying images in all atmospheric conditions.Traffic Panel Detection System could be implemented and used in vehicles having a camera to capture all panels along road side. Various processing techniques could be used to improve quality of scanned image if panel is not a good condition traffic panel. Long distance panels are a challenge for identification of panel and text depicted on them. And also different styles of writing text could be a problem to identify proper text. So a standard text font should be used worldwide. So that an inventory could be maintained to store information about traffic systems.

REFERENCES

1. Álvaro González, Luis M. Bergasa, Member, IEEE, and J. Javier Yebes "Text Detection and Recognition on Traffic Panels From Street-Level Imagery Using Visual Appearance" IEEE Transactions On Intelligent Transportation Systems, Vol. 15, No. 1, February 2014
2. Álvaro González, Miguel ÁngelGarcía-Garrido, David FernándezLlorca, Miguel Gavilán, J. Pablo Fernández, Pablo F. Alcantarilla, Ignacio Parra, Fernando Herranz, Luis M. Bergasa, Member, IEEE, Miguel ÁngelSotelo, Member, IEEE, and Pedro Revenga de Toro "Automatic Traffic Signs and Panels Inspection System Using Computer Vision" IEEE Transactions On Intelligent Transportation Systems, Vol. 12, No. 2, June 2011
3. AyatullahFarukMollah,NabamitaMajumder, Subhadip BasuandMitaNasipuri"Design of an Optical Character Recognition System for Camera-based Handheld Devices" IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 4, No 1, July 2011
4. Hanzi Wang and David SuterProc.Color"Image Segmentation Using Global Information and Local Homogeneity" VIIth Digital Image Computing: Techniques and Applications, Sun C., Talbot H., Ourselin S. and Adriaansen T. (Eds.), 10-12 Dec. 2003, Sydne
5. Shalin A. Chopra, Amit A. Ghadge, Onkar A. Padwal, Karan S. Punjabi, Prof. Gandhali S Gurjar "Optical Character Recognition" International Journal of Advanced Research in Computer and Communication Engineering Vol. 3, Issue 1, January 2014
6. Sukhpreet Singh "Optical Character Recognition Techniques: A Survey" Journal of Emerging Trends in Computing and Information SciencesVol. 4, No. 6 June 2013
7. K. E. A. van de Sande, T. Gevers, and C. G. M. Snoek, "Evaluating color descriptors for object and scene recognition" IEEE Trans. Pattern Anal.Mach. Intell., vol. 32, no. 9, pp. 1582–1596, Sep. 2010
8. A. González, L. M. Bergasa, J. Yebes, and M. Sotelo, "Automatic information recognition of traffic panels using SIFT descriptors and HMMS", in Proc. ITSC, 2010, pp. 1289–1294
9. N. Kulkarni, "Color thresholding method for image segmentation of natural images",Int. J. Image, Graph. Signal Process. vol. 4, no. 1, pp. 28–34, Feb. 2012.
10. G. Csurka, C. R. Dance, L. Fan, J. Willamowski, and C. Bray "Visual categorization with bags of keypoints ",in Proc. Workshop Stat. Learn.Comput. Vis. ECCV, 2004, pp. 1–22.
11. Gregory K. Myers · Ramesh Nallapati · Julien van Hout · Stephanie Pancoast ·RamakantNevatia · Chen Sun · AmirhosseinHabibian · Dennis C. Koelma· Koen E. A. van de Sande · Arnold W. M. Smeulders · Cees G. M. Snoek "Evaluating multimedia features and fusion for example-basedevent detection" Published online: 23 July 2013
12. A. Vázquez Reina, R. J. LópezSastre, S. Lafuente Arroyo, P.GilJiménez."Adaptive traffic 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)
13. Hanzi Wang and David SuterProc.Color"Image Segmentation Using Global Information and Local

ISSN: 2277-9655
(I2OR), Publication Impact Factor: 3.785

Homogeneity" VIIth Digital Image Computing: Techniques and Applications, Sun C., Talbot H., Ourselin S. and Adriaansen T. (Eds.), 10-12 Dec. 2003, Sydney