A Novel Approach to Detect and Recognize Text in Signboard Images

Monika Pathak[1], Anuj Kumar Gupta[2], Sohni Singh[3]

[1]*Assistant Professor, Department of Computer Science, Multani Mal Modi College, Patiala*
[2]*Professor, Department of Computer Science & Engg., Chandigarh Group of Colleges, Mohali*
[3]*Teaching Assistant, Department of Electronics Comm. & Engg., Chandigarh Group of Colleges, Mohali*

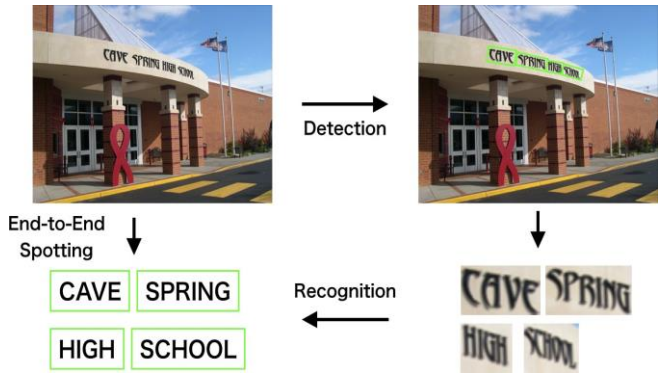
Abstract

This paper describes that how a text can be identified and recognized from any image that has text in it. Sign board has been detected and recognized in this paper. This is performed on the images which are amorphous in nature or containing scenarios which are random or that cannot be determined. Detecting texts from the unstructured images aids in many of the additional applications like Optical Character Recognition (OCR). The algorithm of automatic detection of text also helps in many practical applications. This approach can be used to alert a driver about any road sign even from a captured video. The algorithm uses Maximally Stable Extremal Regions (MSER) feature detector. The algorithm contains several steps which are briefly described in the paper. The approach does not need any external hardware and also the computational time is very less.

Keywords—*Text detection, Text recognition, MSER, natural images, OCR.*

I. INTRODUCTION

Certainly, text is very important and crucial part of the mankind. Text in the written form of the languages used by human makes it very efficient, reliable and feasible to gain and acquire the knowledge and information in time and space. On the one hand, text is a very important device for association and communication, also plays a very vital role in modern society; on the other hand, the exact and specific high level semantics which are exemplified in text are very much advantageous for a better understanding of the world in our surroundings. There are many real-time practical applications around us in which the text detection and text recognition is employed , that include image search [1], [2], instant translation of the text [3], navigation of the robots [4], [5], and in the industrial computerization [6], [7]. Therefore, automatically detecting the text and recognizing it from the natural images [8] or Photo OCR



[9], as shown in Figure 1 has developed a significant and popular topic of research.

Figure 1. Schematic diagram of scene text detection and recognition

The printing technologies were more prominent before the rush of multimedia content. We used books and newspapers to share our knowledge and thoughts. We were unable to reach the mass people around the country at that point of time due to many reasons (i) we utilized products and tools which are not feasible in large scale production (ii) ability to find out the large size content (iii) language obstructions. Nevertheless, with the increasing utilization of computers in the 90s, all the above obstacles were challenged as the data and content are digitized which made very to share things on the internet and also eliminating the language barrier. Automatic detection and recognition of text signs from natural images is a thought-provoking problem as they are typically fixed in the environment. It is very similar to the detection and recognition of text from video images or we can also say Video OCR. The detection of signs takes place in a more vibrant environment as compared to video OCR.

We can implement the text detection and recognition using very limited resources. Sign translation is very much different from a traditional language translation task due to its extraordinary features.

II. TEXT DETECTION AND RECOGNITION

1. ***Text Detection*** – the exact area of the text is located in the natural image which is its main objective. Or we can say, it identifies the region of the minimum area that is of interest to us from a natural scene of regular size. The area to be identified contains two forms, (i) rectangular boxes set which locate the text or (ii) segmentation of the pixels of the text in the form of binary data [10]. A typical image consists of an input and an output. The natural image to be considered is taken as input and the output is the non-overlapping rectangular boxes in the binary form. We can evaluate such systems on the basis of accuracy and precision with the application of intersection upon union statistics. Both the detection methods are contrasting.

- In detection through bounding boxes, word level bounding boxes are anticipated where as in segmentation techniques there is no grouping of character information.
- Segmentation methods helps to identify the text with even more precision by segregating the background and foreground data. While in detection through bounding boxes, the background exists in the text.

2. ***Text Recognition*** – after locating the text region, it is cropped and further processed for the recognition of text. We separate the background and foreground areas to recognize the text. Recognition can be done in many different ways. They are assessed grounded on the definite accuracy of the characters which are predicted with respect to the ground truth word. The text which are present in the scene can have words which are out of dictionary or having huge quantities of dissimilarity which may lead to the misclassification of characters.

III. SYSTEM DESIGN

Following are the steps to be followed in the algorithm or the data flow of text detection and recognition. Each of the point is accurately given in the algorithm flow chart and is also further elaborated thereafter.

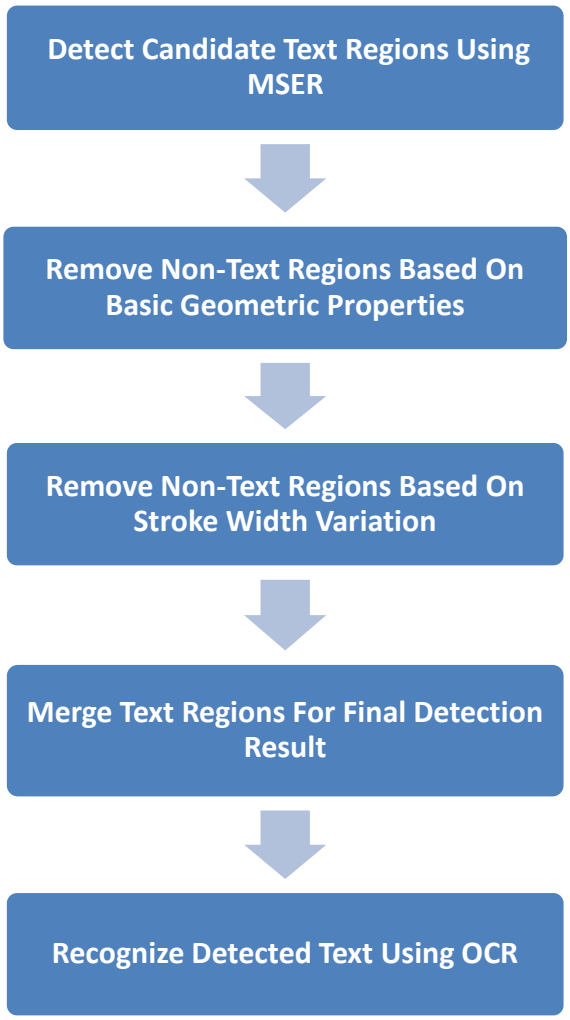


Figure 2. Flow of system design

The MSER feature detector has a very good application in detecting the text from the images. [11]. It works great for text as due to high contrast and color efficiency leading to stable intensity profiles. Other than detecting text, MSER also detects stable images from the input natural scene. Many methods are used to eliminate the non-text regions. For instance, we can remove the non-text region by using the text geometrical properties using their thresholds. Otherwise, we can use a machine learning approach to separate the text from the image. Usually, by using both the approaches, we can get better results [12]. We can use numerous geometric properties of the image to separate the text with non-text regions. [13], [14], including: Aspect ratio Eccentricity Euler number Extent Solidity

Stroke width is also used for the same. Stroke width is a measure curve width and lines that make up a character. Text regions have less variations in the stroke width, and on the other side non-text regions have larger variations.

We need to quantify the entire natural image in order to discriminate the text and non-text regions with the threshold value using stroke width. Therefore, we apply a threshold value to eliminate the non-text regions.

At this point, all the detection results are composed of individual text characters. To use these results for recognition tasks, such as OCR, the individual text characters must be merged into words or text lines. This enables recognition of the actual words in an image, which carry

more meaningful information than just the individual characters. After detecting the text regions, use the ocr function to recognize the text within each bounding box.

IV. RESULTS

For detection of text, firstly the sign board image is entered which is shown in fig 3. Then using MSER Maximally Stable Extremal Regions detection text features are detected. Fig 4 and 6. Shows MSER regions and text detection. Then based on basic geometric properties, non-text regions are removed shown in fig 6. Fig 7 shows merging of text regions and detecting sign and text. Using OCR function the text in image is converted into text shown in fig 8. Further text can be used for any processing like alerting driver or converting into local language.

Figure 3. Original sign board image

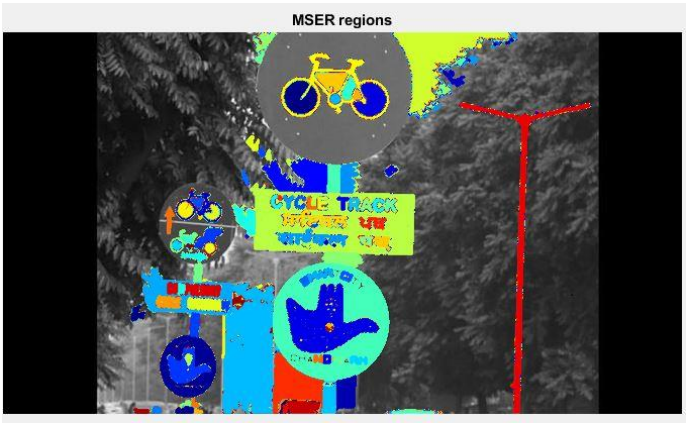


Figure 4. Image after finding MSER regions

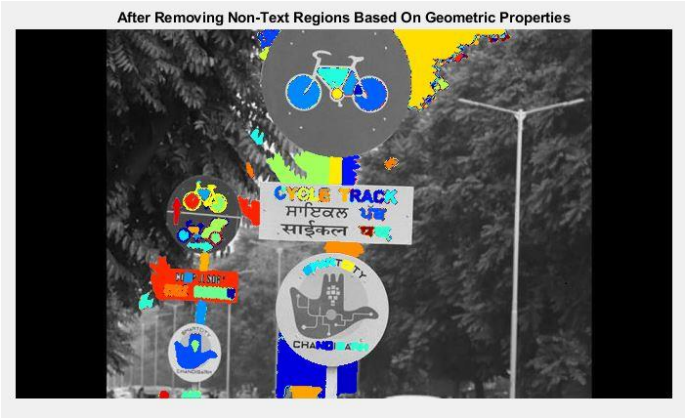


Figure 5. Image after removing non-English text regions

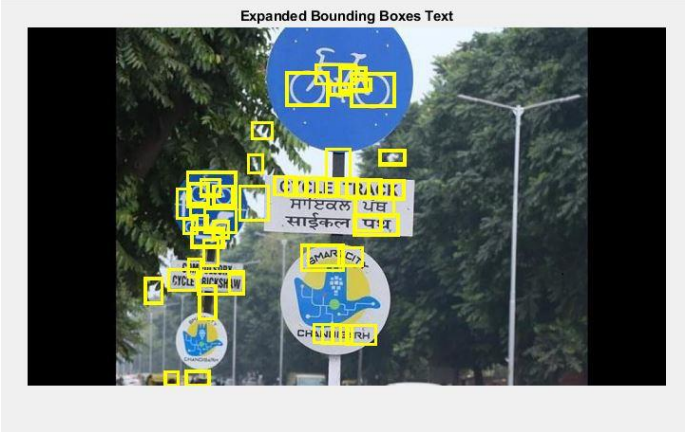


Figure 6. Image detecting the regions

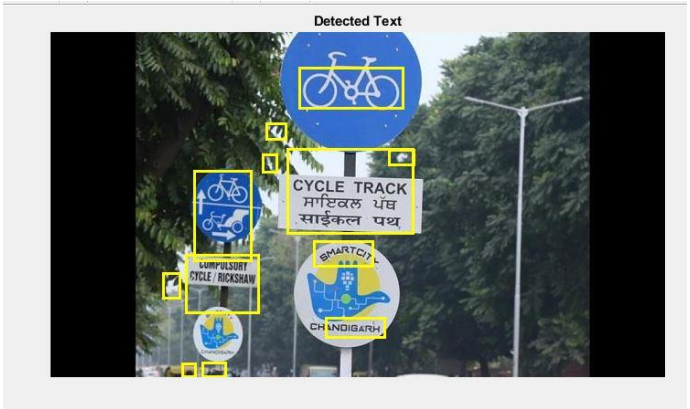


Figure 7. Detected text and symbols

Figure 8. Text converted using OCR from input image

V. CONCLUSION

This paper presents the novel approach for detecting text and recognizing text from any image. The schematic system containing the sign board image as input and the output with bounding rectangular boxes are taken. These boxes act as the method to separate the text from the image. The area containing the text region is separated with the help of a system design contains all the steps in the form of a flowchart and in simulated in MATLAB. The approach used does not need any additional hardware so it can be comparable with other existing approaches. There are many real time applications in the detection and recognition of text from any natural image or any kind of sign board. The detected text can be converted into local language.

References

[1] *Georg Schroth, Sebastian Hilsenbeck, Robert Huitl, Florian Schweiger, and Eckehard Steinbach. Exploiting text-related fea- tures for content-based image retrieval. In 2011 IEEE International Symposium on Multimedia, pages 77–84. IEEE, 2011.*

[2] *Sam S Tsai, Huizhong Chen, David Chen, Georg Schroth, Radek Grzeszczuk, and Bernd Girod. Mobile visual search on printed documents using text and low bit-rate features. In 18th IEEE International Conference on Image Processing (ICIP), pages 2601– 2604. IEEE, 2011.*

[3] *Yuval Dvorin and Uzi Ezra Havosha. Method and device for instant translation, June 4 2009. US Patent App. 11/998,931.*

[4] *Guilherme N DeSouza and Avinash C Kak. Vision for mobile robot navigation: A survey. IEEE transactions on pattern analysis and machine intelligence, 24(2):237–267, 2002.*

[5] *Xiaoqing Liu and Jagath Samarabandu. An edge-based text region extraction algorithm for indoor mobile robot navigation. In Mechatronics and Automation, 2005 IEEE International Conference, volume 2, pages 701–706. IEEE, 2005.*

[6] *MM Aftab Chowdhury and Kaushik Deb. Extracting and seg- menting container name from container images. International Journal of Computer Applications, 74(19), 2013.*

[7] *Young Kug Ham, Min Seok Kang, Hong Kyu Chung, Rae- Hong Park, and Gwi Tae Park. Recognition of raised characters for automatic classification of rubber tires. Optical Engineering, 34(1):102–110, 1995.*

[8] *Yingying Zhu, Cong Yao, and Xiang Bai. Scene text detection and recognition: Recent advances and future trends. Frontiers of Computer Science, 10(1):19–36, 2016.*

[9] *Alessandro Bissacco, Mark Cummins, Yuval Netzer, and Hartmut Neven. Photoocr: Reading text in uncontrolled conditions. In Proceedings of the IEEE International Conference on Computer Vision, pages 785–792, 2013.*

[10] *A. Clavelli, D. Karatzas, and J. Llado´s. A framework for the assessment of text extraction algorithms on complex colour images. In DAS, 2010.*

[11] *Chen, Huizhong, et al. "Robust Text Detection in Natural Images with Edge-Enhanced Maximally Stable Extremal Regions." Image Processing (ICIP), 2011 18th IEEE International Conference on. IEEE, 2011.*

[12] *Neumann, Lukas, and Jiri Matas. "Real-time scene text localization and recognition." Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 2012.*

[13] *Gonzalez, Alvaro, et al. "Text location in complex images." Pattern Recognition (ICPR), 2012 21st International Conference on. IEEE, 2012.*

[14] *Li, Yao, and Huchuan Lu. "Scene text detection via stroke width." Pattern Recognition (ICPR), 2012 21st International Conference on. IEEE, 2012.*