

# A Survey on Text Detection from Natural and Scanned Images

**Govardhan Gayake**, ME Student, Savitribai Phule University of Pune, Department of Computer Engineering, JSPM COE, Hadapsar, Pune

## ABSTRACT

*In today's fast moving world images lays a crucial role in day to day life for various part of daily life. Text detection of captured image is one of the emerging technique where, from images we just identify section from where we can extract text using Optical Character Recognition method. There number of application of this research where we can apply and find out best solution for given problem. Number detection, Text detection from images has various application in day to day world. OCR has three different phases as pre-processing, segmentation, character recognition. In segmentation, each character is identified. Then this will be given to recognition of OCR which will compare it with training data-set and will identify character. In this survey paper, various techniques for OCR are discussed.*

## Keywords

*Optical Character Recognition (OCR), Pattern Recognition, Automatic Number Plate Recognition (ANPR)*

## 1. INTRODUCTION

Now a days use of digital documents are increased as they are easy to handle, share and store than hard copy of documents. So the use of camera is increased drastically as we are getting these images from camera only. Different books, documents are converted into digital form using scanner as well as camera. Due to this people rather prefer digital document over hard copy of documents. Text detection is very important process as it required to extract text from images. Extracting this text data from images are used in order to store it digitally. To perform these operations, Optical Character Recognition technique can be used. There are different pre-processing techniques for captured images are available from which text can be extracted. Pre-processing phase is used for increasing accuracy and efficiency of algorithms. For scanned and Natural images, different techniques are required. An image with simple background and text is easy to recognize. But for natural images with complex background like images which were taken by camera such

as road signs, it may have different contrast, and may contain non textual part. So to detect text from these images is difficult to recognize.

This survey paper presents the picture of different Text detection algorithm for images which have complex background. Natural images contains a lot of non-text objects. To remove this non-text objects and detect accurate text from these images, text detection algorithm with Maximally Stable Extremal Regions can be used. This algorithm uses Maximally Stable Extremal Regions as basic letter candidates. Candidate is then further processed with geometric constraints and stroke width transformation to remove non text objects. Automatic Number Plate Recognition(ANPR) system is used to detect the numbers from Number Plate. It localizes the number plate within the captured image and then apply OCR on that region only.

OCR needs training data set for comparison with its calculated sets and those data is font dependent. It requires large amount of storage and computation time thus increases the complexity of the system. Feature extraction technique is used for text to get recognize from image.

This paper is divided into different sections. The following section represents Text Detection algorithm and various types. Finally the conclusion is presented in last section.

## 2. TEXT DETECTION ALGORITHM

### 2.1 Maximally Stable Extremal Regions

In Maximally Stable Extremal Regions (MSER), the most structure and well defined block is detected from captured image. Captured image could be traffic sign board, Number plate of vehicle. Stable regions detected from MSERs are referred as individual characters and then grouped into lines for recognition through OCR. MSER algorithm works in two steps i.e. detection and recognition. In detection step a particular structure or block is set in the image and the text which is lie inside the block/region is detected. Once that block is detected then recognition step is applied. Recognition steps performs extraction from that located region and text is retrieved.

Figure 1 shows the Detection and Recognition step from MSER [1].

➤ Detection Phase

Detection Phase has three subparts and they are:

- Determination of search region (where we are expecting our text to be located)
- Detection of candidates from the region

- Reduction of candidate using contextual constraints [1].

The dimension of these search bocks are given explicitly to the system which were depends upon the position of the texts from the captured image. Next will be the reduction of candidates using temporal and contextual information is perform. Then each candidate or character is compare with the previously defined frame. If match is found in the previously extracted text then use previously extracted text as output else treat the text as a newly identified text.

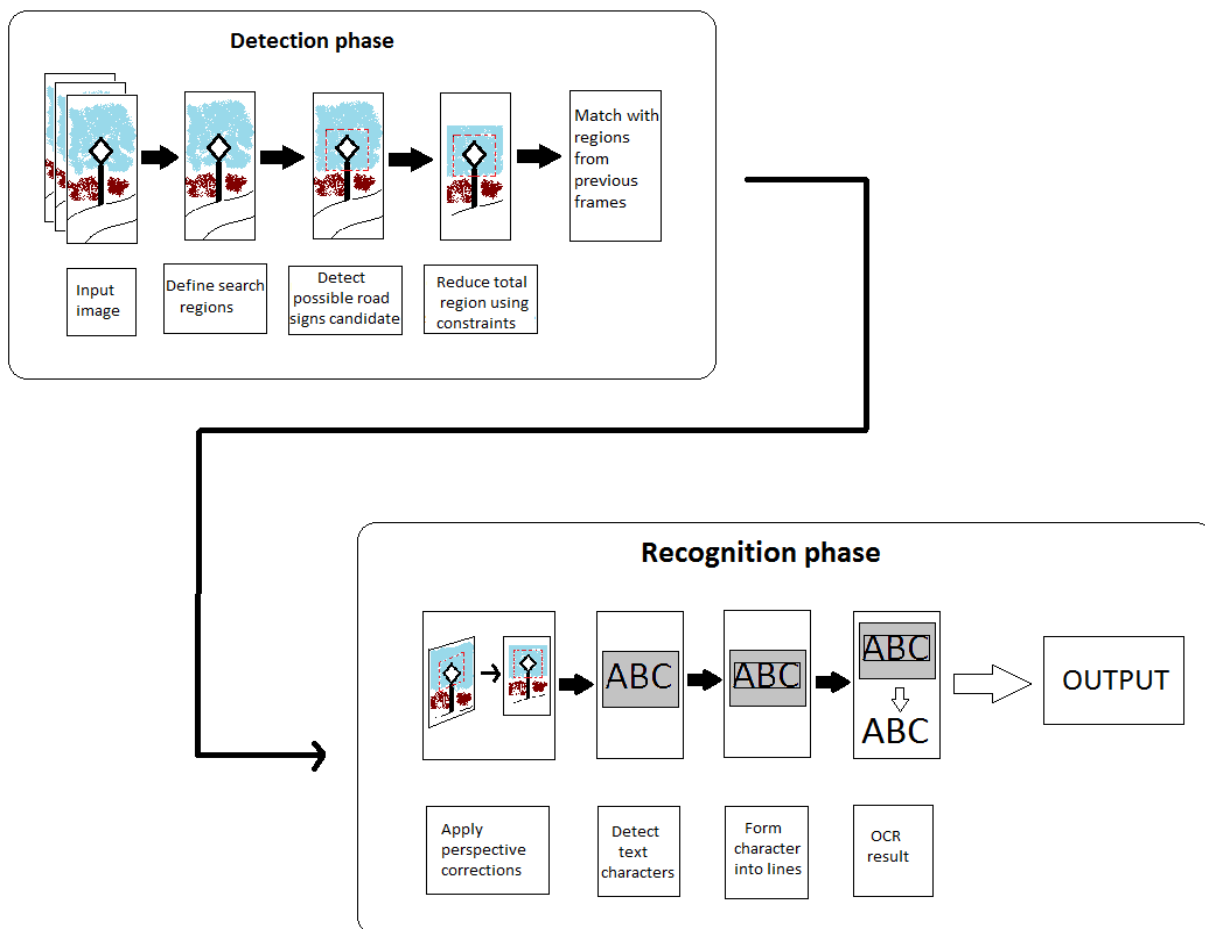


Figure 1: MSER phases [1]

➤ Recognition Phase

This is the step where text is recognizes the content from the detected region of Detection Phase. Different transformations are applied for alignment of the region to be vertical and then individual character are segmented, formed into words. Characters which was segmented then sent to OCR for recognition. This transformation is done by fitting the quadrilateral to detected text region. Then within the candidate blocks, lines of text in the detected

text area were located. Lines which are aligning vertically and unwanted text characters are removed from line and then they were feed to OCR. Detected lines obtained are in the form of gray scale image and then those images are given to open source Tesseract tool which retrieve the text.

2.2 Texture and Connected Component Techniques

Visual based search systems, image features are extracted and then those features were being compared with large

dataset to recognize text from images. The text which was located should be robust because in visual search system there are number of various distortions, different font styles and different lightening conditions. Texture-based and Connected Component (CC) based are the two techniques which can be used as Text Detection techniques.

#### ➤ Texture-based Technique

Texture-based approaches view text as a special texture that is differentiable from the background of images. Typically, the features are extracted over a certain region and a classifier (trained using machine learning techniques or by heuristics) is used to identify the existence of text[2]. In [3] assume that text has certain horizontal and vertical frequencies and extract features to perform text detection in the discrete cosine transform domain.

#### ➤ CC-based Technique

As alternative to texture-based methods, the CC-based approach extracts regions from the image and uses geometric constraints to rule out non-text candidates. Text lines are then formed by linking the CCs based on geometric properties [2].

In this algorithm, first we need to enhance contrast of input images. Next step will be getting located MSER regions from an image and based on that regions, edges are obtained from gray scale image. Then on the resulting output, Connected Components are filtered out using geometric constraints on properties like aspect ratio and number of holes. The stroke width information is computed using different transform e.g. distance transform and high variation objects from the stroke width are rejected. Resulted candidates were grouped pairwise and from them text lines were formed. And the last step will be the words within a text line are extracted separately.

### 2.3 Automatic Number Plate Recognition

#### ➤ Number Plate Localization Module

Basically, this module localizes number plate from the given image. There are different morphological operations are available for feature extraction. These operations, does feature extraction and removal of noise from the NP region. Computational intensity of system is decrease by Morphological operations and also performance increases with respect to other algorithms such as edge detection algorithm. With this operations, system will recognize text more accurately as compared to other algorithms.

#### ➤ Character Segmentation Module

In character segmentation module, all the pre-processing techniques are applied on localized region of image. Pixel projection operation is one of the pre-processing technique which was carried out in different phases. Resulted image from the NPL module is binarized and if there is a titled image is present then it is rotated before going to fed to CS module. In the first step of CS module, the unnecessary parts of Number Plate are removed by a height optimization. By using vertical and horizontal projection, the vertical positions of the characters and the horizontal positions of the characters is recognized.

#### ➤ Optical Character Recognition Module

The OCR module includes character segmentation and recognition. It uses Neural Network to translate the scanned character into text. For neural network, Binary image matrix is given in the form vector as an input. First of all we need to trained neural networks with training characters and then it is going to test against the input. Throughput and accuracy of OCR system is better than other systems.

Most of the researchers uses DSP (Digital Signal Processors) or FPGA (Field Programmable Gate Arrays) for performing Number Detection in ANPR systems to get real time and improved performance [4].

### 3. CONCLUSION

For text detection from images there are various algorithms are available and most of them uses MSER as basic letter candidate. MSERs are unable to retrieve character with small font size and they are sensitive to blur. Images which are going to use text retrieval are captured using low resolution cameras thus edge-enhance MSER can be used to overcome sensitivity to image blur. So, Text detection algorithm with enhanced edge MSER can be used efficiently on blur images. Artificial neural networks and vector based training data sets can be used to recognize the text from images accurately. Traditional algorithms are font dependent and they need a lot of redundant processing. Thus we can use feature extraction on complete image for enhancing the algorithm which is much faster and memory efficient as feature extraction does not require any training data or mechanism. And proposed algorithm is independent of fonts.

## ACKNOWLEDGMENT

I express my gratitude towards the department of Computer Engineering, JSPMCOE for giving me the opportunity to present the paper. A special thanks to my guide Prof. Arati Dandavate, whose help, stimulating suggestions and encouragement helped me to coordinate this paper.

## 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] Huizhong Chen, Sam S. Tsai, Georg Schroth, David M. Chen, Radek Grzeszczuk and Bernd Girod, Robust Text Detection in Natural Images with edge-enhanced Maximally Stable Extremal Regions, ICIIP 2011.
- [3] Y. Zhong, H. Zhang, and A. K. Jain, "Automatic caption localization in compressed video," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 4, pp. 385–392, 2000
- [4] Xiaojun Zhai, Faycal Bensaali and Klaus McDonald-Maier, Automatic Number Plate Recognition on FPGA, 978-1-4799-2452-3/13 IEEE, 2013.
- [5] Sushruth Shastry, Gunasheela G, Thejus Dutt, Vinay D S and Sudhir Rao Rupanagudi, i - A novel algorithm for Optical Character Recognition (OCR), 978-1-4673-5090-7/13 IEEE, 2013.
- [6] Julinda Gllavata, Ralph Ewerth and Bernd Freisleben, A Robust Algorithm for Text Detection in Images, University of Marburg, D-35032 Marburg, Germany.
- [7] Jayshree Ghorpade, Raviraj Palvankar, Ajinkya Patankar and Snehal Rathi, EXTRACTING TEXT FROM VIDEO, Signal & Image Processing : An International Journal (SIPIJ) Vol.2, No.2, June 2011
- [8] Mrunmayee Pati, Ramesh Kagalkar, An Automatic Approach for Translating Simple Images into Text Descriptions and Speech for Visually Impaired People, International Journal of Computer Applications (0975 8887) Volume 118 No. 3, May 2015.
- [9] Binh Quang Long Mai, Tue Huu Huynh, Anh Dong Doan, A Study about the Reconstruction of Remote, Low Resolution Mobile Captured Text Images for OCR, 978-1-4799-6956-2/14 IEEE, 2014.
- [10] Sushruth Shastry, Gunasheela G, Thejus Dutt, Vinay D S and Sudhir Rao Rupanagudi, i - A novel algorithm for Optical Character Recognition (OCR), 978-1-4673-5090-7/13 IEEE, 2013.
- [11] Faisal Mohammad, Jyoti Anarase, Milan Shingote, Pratik Ghanwat, Optical Character Recognition Implementation Using Pattern Matching, International Journal of Computer Science and Information Technologies, Vol. 5 (2), 2014.
- [12] Adam Coates, Blake Carpenter, Carl Case, Sanjeev Satheesh, Bipin Suresh, Tao Wang, David J. Wu, Andrew Y. Ng, Text Detection and Character Recognition in Scene Images with Unsupervised Feature Learning, ICDAR.
- [13] Jisha Gopinath, Aravind S, Pooja Chandran, Saranya S S, Text to Speech Conversion System using OCR, IJETAE, Volume 5, Issue 1, January 2015.
- [14] Yao Li and Huchuan Lu, Scene Text Detection via Stroke Width, 21st International Conference on Pattern Recognition (ICPR 2012), November 11-15, 2012, Tsukuba, Japan.
- [15] Er. Kavneet Kaur, Vijay Kumar Banga, Number Plate Recognition using OCR Technique, International Journal of Research in Engineering and Technology eISSN: 2319-1163 — pISSN: 2321-7308.