

## Review on Text Detection and Extraction for Catching Identities of Objects in Road Video Scenes

Namrata Choudhary<sup>1\*</sup>, Kirti Jain<sup>2</sup>

<sup>1,2</sup>Dept. of Information Technology, Lakshmi Narain College of Technology, Bhopal, MP. India

\*Corresponding Author: [namratac2014@gmail.com](mailto:namratac2014@gmail.com)

Available online at: [www.ijcseonline.org](http://www.ijcseonline.org)

Accepted: 26/Sept/2018, Published: 31/Oct/2018

**Abstract-** In the current scenario there are different types of research has been performed in the field of text detection and extraction from images. Different types of processes used different types of application of the text extraction from the images. With the help of the Text Extraction process we can easily find important data from images. There are various types of Image processing techniques have been evolved for Text Extraction from the image scene and videos. Every process has different types of factors like Precision, Speed, Complexity and Time required, etc. but each process gives different result in each field. Some process has advantage and disadvantage related to these factors, so we can say single process is inadequate for the complete process of text detection and extraction. For the preferable performance, we present the combine form of different processes. This paper contains combination of two different processes for text detection and extraction from video surveillance system just like processing of images from image scenes.

**Keywords:** -Text detection, Text extraction, Image processing, Precision, Preferable, Video surveillance.

### I. INTRODUCTION

In the present days increasing needs for security applications have inspired the betterment in the field of visual surveillance system. In the busy traffic it is very necessary to control the overall activities happen. This is only possible when we have proper system to see and analyse all these activities. It is not easy to separate movements of a particular vehicle. In this paper we use the text detection and extraction for catching the vehicle images form the input videos of the surveillance system. Now-a-days text detection and extraction from videos is highly demanded. Text data is important because text describes the overall content written on the vehicle. From these videos of surveillance system, we general capture the name, number plate, colour of the vehicle and also the different type of text written on the vehicle.

However, text detection from videos is still difficult task due to the complex background, font size, styles and sporadic lighting condition [1] [2]. A vehicle is thing, or we can say mobile machine which is use for transporting people or cargoes on the land. There are different types of vehicles such as two wheels, three wheels, four wheels, etc. Examples of the vehicles are motorcycles, cars, buses, etc. This paper focus on extraction of text written on the vehicles. We use many pre-processing and extraction techniques which helps us in removing the problems occur in the video text detection.

### II. LITRATURE SURVEY

There are different types of methods proposed in the field of text extraction form the images.

In the paper Wahyono, et.al [12] Multilanguage text is detected by the use of a Fast Stroke Width Transform. This method detect the natural images in place of document images because due to their complex background and different type of variation in colour, text, fonts the natural scenes is very difficult to difficult to detect. Methods used in this paper are Canny Edge Detector, FSWT produces very fast result in the real time scenario but design of this method is difficult.

In the paper C.P. Sumathi et.al [13] extraction of text from videos is done. Video text is divided into two parts Scene text and Caption text. Methods used in this paper are Morphological operation, Neural Network classifier, Wavelet transformation produces efficient result with less error rate and have a capability of noise tolerance. Main drawback of this method is very difficult to design and produce result very slowly.

In the paper Ho Vu, et.al [14] proper design of acknowledgement of character from images is proposed. Methods used in this paper are Sparse Coding as a Mapping-function and Feature learning method with Orthogonal matching pursuit for training. This method take time in producing result while feature vector is extracted.

In the paper Pan et. al [15] combined method for accurate text localization from natural scene images. Methods used in this paper are Graph cuts inference, Minimum spanning tree and hybrid form of the CC and Region based approach containing Conditional Random Field (CRF) model produce accurate and vigorous result but this process takes more time and design is very complex.

In the paper Huizhong Chen et.al [16] a novel text detection method is proposed. Methods used in this paper are Stroke width Transform (SWT), CC based Edge-enhanced and Maximally Stable Extrenal (MSER) produces simple and efficient result. In this methods pre-processing step is not perform so that the detection of text is fail because there is no focus on the detection and excessive blurriness is also present in the images or the scenes that we take from the road video so that the clear detection is not possible.

In the paper Huang et.al [17] proposed video scene text detection and localization method. Perform two steps, First, by using Log-Gabor filter map is construct and Second, calculate texture feature on every line of map. Methods used in this paper are Connected Component Analysis, Stroke Map and Harris Corner Detection produces vigorous result from video scene text detection in case of variation of text size but not produce suitable result in low contrast background.

In the paper Nobuo Ezaki et.al [18] present different methods for the text extraction based on different character size. Method used in this paper are Rule-based connected-component selections, Connected Component extraction. Combination of these methods is easy to design and produces good result but the accuracy of text detection of small text is low.

In the paper Gllavata et.al [19] .Method used in this paper are Edge detection, colour reduction technique and localization of text using projection profile and geometrical properties operate in Grey scale as well as Colour image but detection in low quality images is complex.

Hrishav Raj,et.al [8] proposed a method for the extraction of Devanagari text from the images scenes.in this method they suppose that the characters are broad and large, so that the capturing of text from the images by using linear structuring element of fixed length become easy. Morphological opening operation method with structuring element as line is used for text detection. There are different geometrical properties were used for finding the location of the overall text part associated to the detected text. Above discussed methods pre-processing, processing and post-processing are used in this method. But this method is used or we can say instructed for the detection of Devanagari text from the images.

Andrej Ikica, et.al [9] proposed a method for text detection by using Edge profile based detection with Canny edge map, Heuristic rules. In the Canny edge map method edge pixels of images are combined into the connected components. But there is a problem in detecting the character that are smaller than 5 pixels.

To solve the problem in canny edge map method heuristic rules have been proposed. These heuristic rules therefore frequently remove the true text regions and alternatively leave the non-text region undamaged. But there are also some problems, so to solve this problem a moderately improved strategy is proposed [9]. Some rules for the elimination of non-text candidate region which is same as [10] are used. Some new and improved methods are proved helpful in solving the problems related to this method.

### III. METHODOLOGY

For our comfort, we can divide the overall process of text detection and extraction from videos into the three main steps:

In the pre-processing stage we use some algorithm for the removal of affects created by the background, styles and sporadic light. After this stage we made image perfect for further processing. We perform processing and post-processing step after this stage.

In the processing stage takes the pre-processed image. In this stage we can use different types of text detection, extraction, segmentation and localization methods.

The post-processing stage is applied after the processing of data for the text recognition task.

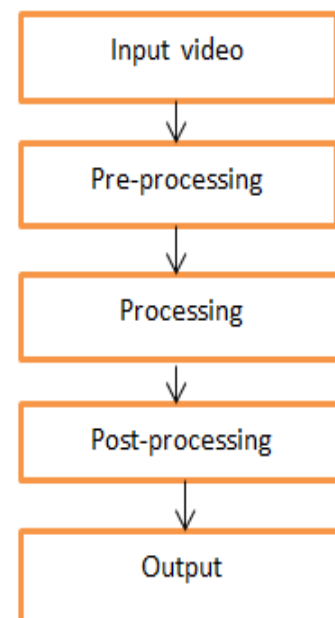


Figure 1. Text Detection and extraction process

### A. Pre-Processing Stage

This is the first step from where we start the processing of text from the video. Before beginning the overall process of text detection an extraction from videos, we first examine video images text that can be mixed with different types of noise or it can be blurred due to the disturbance in the camera. For the improvement, we should use different types of image pre-processing techniques. For the purpose of De-blurring following techniques are used: Blind de-convolution algorithm, Lucy Richardson algorithm, Wiener de-blurring techniques are used [3]. A wiener filter is selected from these techniques for the removing of noise present in the images. Sometimes, salt and pepper are also present in the images which are also a type of Impulse noise. This type of noise is also amoral the images. Salt means white dot on the black background and pepper means black dot on white background which reduces the performance of the system [4]. The standard median filter (SMF) is the non-linear filter used to remove noise present in the images. It is helpful removing the noise and has good computational efficiency which is not present in the linear filtering method.

Sometimes, we capture image which is not clear due to dark and lumpy lighting here also the problem of text extraction occur. To control the problem of lumpy lighting application of contrast enhancement is used which is predominantly prefer the histogram equalization method. This is the proper method for the process.

### B. Processing Stage

1) *Text Detection*: Text detection is the process in which we take video as an input and search for the text. It plays an important role in the text searching from the input data and arbitrates that it contains text or not.

There are three types of methods are present for the process of text detection.

Connected component based method works with the fix colours and spatial layout like certain size, style and shape. This method is not work properly when the colour of text and background are same. This method use bottom up approach.by grouping smaller region into the larger region until all regions are identified in the image.

Edge based method focus on the high contrast between the background and the text and the edge of the boundary are identified and merged. Text with large font size was not detected effectively by this method.

Texture based method is comparatively less sensitive to background colours. Grey-level concurrence matrix is involved in this method which is used to determine the qualities like contrast, homogeneity, dissimilarity and which are the results for feature extraction in texture based method [6].

2) *Text Localization*: In this process text regions amalgamate to construct the text objects and explain the tight bounds around the text objects. Text detection, text localization and text tracking is challenging and tough part of overall process [5].

3) *Text Tracking*: This process is specially and only for the video data. For the readability, text implanted in the video becomes visible more than thirty consecutive frames. This stage utilizes temporal occurrences of the same text object in multiple consecutive frames. This method helps in improving the result of text detection and localization stage. Speed of text extraction process is also improved by this process [5].

4) *Text Binarization*: Many times the process of Binarization can be applied before localization stage. This stage is the part of the segmentation which helps in segmentation of text object from the background in the bounded text objects. The result of this stage is the binary image. There are two types of pixels text pixels and the background pixels they are also appear in two different levels like white text on dark background or vice versa.

### C. Post-Processing Stage

The last stage of the text extraction is the character recognition. In this stage binary text is converted into the ASCII text [7].

Text detection and extraction process mainly consist of the following steps shown in the Fig 2.

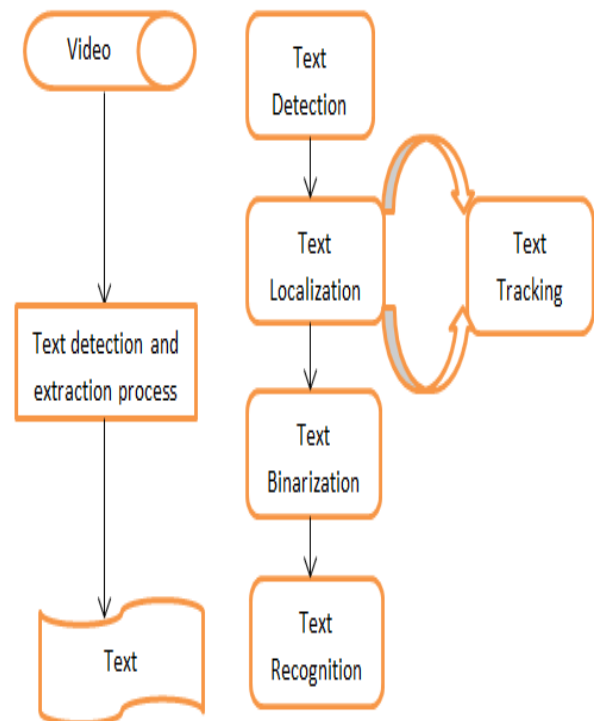


Figure 2. Architecture for text detection and extraction process

Table 1. Selected Methods

Sr.No.	Author	Year	Used Method	Accuracy	Advantages
1.	Hrishav raj et,al	2014	Connected Component(CC), Morphological operations, Canny Edge detection, Binarization	72.8%	Font Size, Style and direction of text does not affect
2.	Andrej et,al	2011	Heuristic rules, Edge profile based detection with Canny Edge map.	70.9%	Efficient, fast and simple

Precision rate is the factor on the basis of this we can check the performance of the different methods. Calculation of precision is

$$\text{precision} = \frac{\text{correctly detected}}{\text{correctly detected} + \text{false positive}} \times 100\%$$



(a)



(b)

Figure 3. Input images [20, 21]

#### IV. CONCLUSION

Detail study of the text extraction, detection, localization and tracking are covered in this paper. We have study the different type of present researches on the text extraction from images in the scene and apply these same methods on

the video images, each method have their advantages and disadvantages. Different papers proposed different methods, some of them improve or we can say modify the process while some develop new processes.

Keeping all these factors in the mind we design a process by combining the two previously discussed processes. Now-a-days vehicles number plate contains different styles, some of them in English, some in Hindi and some combination of both English and Hindi. This paper contains the process for both Hindi and English fonts. This type of process helps in different fields of security, helpful for car tracking system.

#### REFERENCES

- [1] Roberto Manduchi and James Coughlan, "Computer Vision without Sight", Communications of the ACM, Vol.55, no.1, January 2012.
- [2] Qixiang Ye and David Doermann, "Text Detection and Recognition in Imagery: A Survey", IEEE Transactions on Pattern Analysis Machine Intelligence, Vol.37, No.7, July 2015.
- [3] Sonia George, Noopa Jagdeesh, "A Survey on Text Detection and Recognition from Blurred Images", International Journal of Advanced Research Trends in Engineering and Technology(IJARTET), Vol. II, Special Issue X, pp. 1180-1184, March 2015.
- [4] Madhu S. Nair, K. Revathy, and Rao Tataavarti, "Removal of Salt-and Pepper Noise in Images: A New Decision-Based Algorithm", Proceedings of the International Multi-Conference of Engineers and Computer Scientists IMECS, Hong Kong, Volume I, March 2008
- [5] Julinda Gllavata, Ralph Ewerth and Bernd Freisleben, "A Robust Algorithm for Text Detection in Images", IEEE Proceedings of the 3rd International Symposium on Image and Signal Processing and Analysis (ISPA), Volume 2, pp.611 – 616, 2003.
- [6] 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 Region", Image Processing (ICIP), 18th IEEE conference on Image Processing, 2011.
- [7] Rodrigo Minetto, Nicolas Thome, Matthieu Cord, Neucimar J. Leite, Jorge Stolfi, "Snooper Text: A text detection system for automatic indexing of urban scene", Computer Vision and Image Understanding, journal homepage: www.elsevier.com/locate/cvui, 2013.
- [8] Hrishav raj, Rajib Ghosh, "Devanagari Text Extraction from Natural Scene Images", IEEE International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp.513-517, 2014.
- [9] Andrej Ikica, Peter Peer, "An improved edge profile based method for text detection in images of natural scenes", IEEE EUROCON - International Conference on Computer as a Tool (EUROCON), pp. 1-4, 2011.
- [10] N. Ezaki, M. Bulacu, L. Schomaker, "Text Detection from Natural Scene Images: Towards a System for Visually Impaired Persons", Int. Conf. on Pattern Recognition (ICPR 2004), vol. II, pp. 683-686.
- [11] International Journal of Engineering Research and General Science Volume 4, Issue 2, March-April, 2016 ISSN 2091-2730.
- [12] Wahyono, Munho Jeong and Kang-Hyun Jo, "Multi Language Text Detection Using Fast Stroke Width Transform", IEEE 21st

Korea- Japan Joint Workshop on Frontiers of Computer Vision (FCV), pp.1-4, 2015.

- [13] Ho Vu, Duongl and Quoc Ngoc, "A Feature Learning Method for Scene Text Recognition", IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), pp. 176 - 180, 2012.
- [14] C.P. Sumathi, N. Priya, "Analysis of an Automatic Text Content Extraction Approach in Noisy Video Images", International Journal of Computer Applications (0975 – 8887) Volume 69– No.4, May 2013
- [15] 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 Region", Image Processing (ICIP), 18th IEEE conference on Image Processing, 2011.
- [16] Xiaodong Huang, Huadong Ma, "Automatic Detection and Localization of Natural Scene Text in Video", International Conference on Pattern Recognition (ICPR 2010), IEEE Computer Society, pp.3216-3219, 2010.
- [17] Y. Pan, X. Hou, and C. L. Liu, "Text Localization in Natural Scene Images based on Conditional Random Field", Proc. IEEE International Conf. Doc. Anal. Recognition, pp. 6-10, 2009.
- [18] Nobuo Ezaki, Marius Bulacu, and Lambert Schomaker, "Text Detection from Natural Scene Images: Towards a System for Visually Impaired Persons", Proc. of 17th Int. Conf. on Pattern Recognition (ICPR), IEEE Computer Society, pp. 683-686, vol. II, 23-26 August, Cambridge, UK, 2004.
- [19] Julinda Gllavata, Ralph Ewerth and Bemd Freisleben, "A Robust Algorithm for Text Detection in Images", IEEE Proceedings of the 3rd International Symposium on Image and Signal Processing and Analysis (ISPA), Volume 2, pp.611 – 616, 2003.
- [20] V.Lakshmi Priya, K.Perumal, "Detection the Car Number Plate Using Segmentation", International Journal of Engineering and Computer Science ISSN: 2319-7242, Volume 3 Issue 10 October, 2014.
- [21] Ravi Theja, "Number Plate Detection on Indian Car Vehicles Using YOLOv2"  
<https://medium.com/@ravidesetty/number-plate-detection-on-indian-car-vehicles-using-yolov8c99e1a259f5>.

### Author's profile

*Ms. Namrata Choudhary* pursued Bachelors of Engineering form Lakshmi Narain College of Technology, Bhopal, M.P., India in 2017. She is currently pursuing M.Tech. from Lakshmi Narain College of Technology, Bhopal, M.P., India. She has presented paper in National Conference in 2015 & 2016 held in College.



*Mrs. Kirti Jain* pursued Bachelors of Engineering and M.Tech. She is currently pursuing Ph.D. and working as Associate Professor in Lakshmi Narain College of Technology. She has been published paper in International and National journals and conferences.