International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 03 Issue: 05 | May-2016 www.irjet.net p-ISSN: 2395-0072

Text Extraction from image using MSER approach

Geetanjali Adlinge[1], Shashikala Kashid[2], Tejasvini Shinde[3], Virendrakumar Dhotre[4]

[1234]*Student, Dept. of Computer Science & Engineering, FTC COER, Sangola, Maharashtra, India*

Abstract—*This document will propose all features to recognize text from image. Text recognization is very useful for computer vision, document analysis and license plate reading. The texts may also appear written on objects around us: a motto painted on a wall, arranging small stones on the sand, text on street signs. Environment text is a source of information in our daily life. In this document we are using Maximally Stable Extremal Region (MSER) algorithm. The MSER algorithm is able to detect almost all characters from image even when the image is in low quality.*

1.INTRODUCTION

Image processing is a method to perform some operations on an input image, in order to get an extract some useful information from it or to enhance the image. Image processing is a type of signal processing in which input is an image and output may be image or characteristics/features associated with that image. Nowadays, image processing is rapidly growing technology. It forms core research area within engineering and computer science disciplines too.

Image processing basically includes the following three steps:

- Importing the image;
- Analyzing and manipulating the image;
- Output in which result can be altered image or report that is based on image analysis.

As an important representation from human beings language, visual texts are widely used in our daily life. The texts may also appear written on objects around us: a motto painted on a wall, arranging small stones on the sand, text on street signs. Environment text is a source of information in our daily life. Recent advances in digital technology allow to take photos from a large number of mobile devices. As a result, the number of photos taken by users is increasing day by day. At the same time, we often have no annotations for images except those made by the device. Text in images provides information about semantics of the image. In this paper we discuss and use the MSER algorithm. The MSER algorithm is able to detect almost all characters from image even when the image is in low quality. The images contain important names, locations, date and time, brands of the products, street signs, which are helpful information to understand image. Finding best way to detect and extract different types of text from complex images is the first aim of the text recognition from images.

1.1Definitions of the terms of text retrieval

1) Document text: In document text the characters have clean background such as some scanning document pages.
2) Text texture: The texture consists of the pixels belong with the media text
3) Local background: The local background contains neighborhood area of the text texture. In local background there are a lots of text embedded in the images.

1.2Application of the text recognition

Today, more and more information is stored in various kinds of digital forms including the images. Performing the analysis of the input document is an important application of text reorganization. The document contains vast information to manage that information we use text recognition.

The text contained in image is powerful source of information. The text may appear on the objects around us: text on street signs and even text produced by arranging small stones on the sand, player's numbers and names, scores, logos, maps.

1.3Scientific validation

In document text processing, the problem is how to extract the text from arbitrary background. The text in images may have various sizes and fonts. We simply magnify the images to match the size of standard patterns and process the document text.

(A) Input (B) Output

2.EXISTING SYSTEM

The automatic understanding of textual information in image has gained increasing attention over the past decade, giving rise to various new computer vision challenges. The text recognition from image is still a challenging task due to the large variety of text appearance in images. Text in images can have different variations of the size, font style, distortion; it can have different contrast due to different lighting conditions. The problem of extracting the text from image involves the steps of (i) detection, (ii) localization, (iii) tracking, (iv) text extraction and enhancement, and (v) recognition using (OCR). Lluis Gomez and dimosthenis Karatzas defined Multi-script Text Extraction from Natural Scenes. In this paper, they argue that the text extraction problem could be posed as the detection of meaningful groups of regions. They present a method built around a perceptual organization framework that exploits collaboration of proximity and similarity laws to create text-group hypotheses. Fredrik Kristensen and W. James MacLean defined Real-Time Extraction of Maximally Stable Extremal Regions on an FPGA. In this paper MSER detector is used. In order to reach real time performance, both algorithmic and memory issue have been addressed. The union find algorithm which is heart of MSER detector. It is used to create linked regions that significantly decrease the time. V.Kalai selvan and M.Prakash defined Text Extraction from Image Using MSER Approach. In this paper maximally stable extremal region algorithm extract the connected components in images. Then obtained extracted connected components are partitioned into clusters so that we can generate candidate regions using an AdaBoost classifier which determines the adjacency relationship and cluster those connected components by using their pair wise relations.

3.PROPOSED WORK

Our approach consists of five steps which are shown in the figure 1. First MSER algorithm detects the Candidate Text Region within the image and plots the result. MSER detects almost all text regions from image but alongside it also detect some non-text regions which are filtered in next step. MSER uses two important properties to remove non text regions from image first is Geometric Properties and another is Stroke Width Variation Properties. After removing non text regions the individual characters are merged into word or text lines. After merging characters OCR function recognize the text within each bounding box.

s

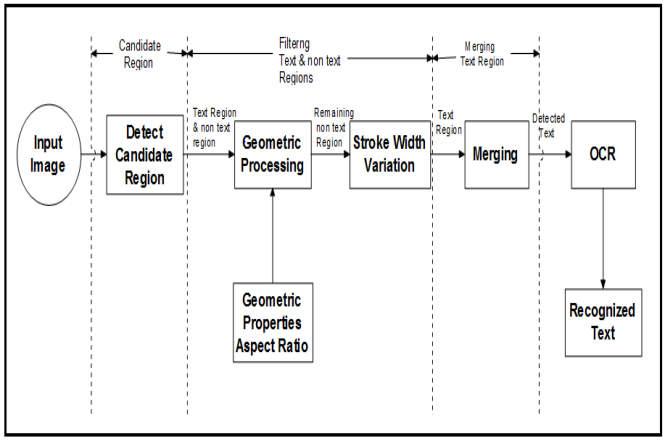

Fig -1: Architecture Diagram

The figure 2 shows Data Flow Diagram (DFD). In DFD the user gives input image to MSER algorithm then MSER detects text from that image and gives output to OCR. OCR function recognizes text within the bounding box and gives result to the user.

Step 1: Detect Candidate Text Regions

In this step the MSER algorithm works for finding text regions to find all the regions within the image and plot the results. There are many non-text regions detected alongside the text which are removed in third step. To detect text regions MSER first convert the color image into gray scale image.

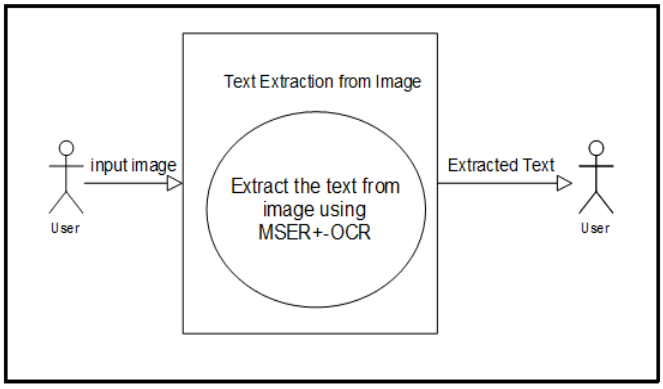

Fig -2: Data Flow Diagram

Step 2: Remove Non-Text Regions Based On Basic Geometric Properties.

The MSER algorithm chooses most of the text and it also detects many other stable regions in the image that are not text. We can use a rule-based approach to detect and remove non-text regions present in the image. The non text regions from image are filter by using geometric properties. Alternatively, we can use a machine learning approach to train a text vs. non-text classifier. Typically, a combination of the geometric properties and machine learning approach gives a better result.

Step 3: Remove Non-Text Regions Based On Stroke Width Variation

Another common metric used to make decision between text and non-text is stroke width. The stroke width variation properties are measure of the width of the curves and lines that make up a character. Text regions have little stroke width variation, whereas non-text regions have larger stroke width variations. To help understand how the stroke width can be used to remove non-text regions, estimate the stroke width of one of the detected MSER regions.

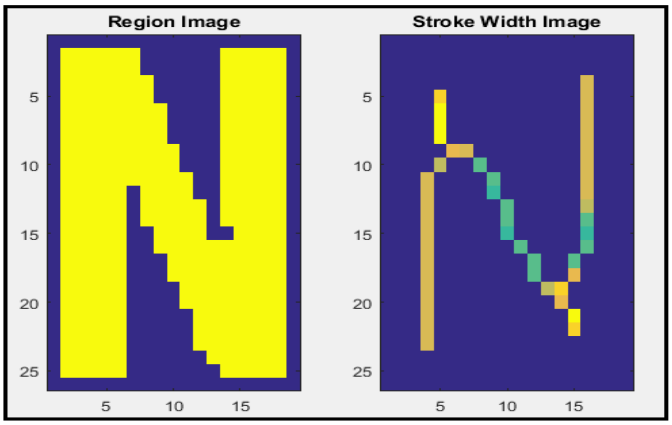

Fig -3: Region Image and Stroke Width Image

Step 4: Merge 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. For example, recognizing the string 'HELP' vs. the set of individual characters {'L','H','P','E'}, where the meaning of the word is lost without the correct ordering.

Step 5: Recognize Detected Text Using OCR

After detecting the text regions, use the OCR function to recognize the text within each bounding box as shown in Figure 4

Fig -4: Detected Text

4.THE TEXT DETECTION ALGORITHM

In this section, we describe the text detection algorithm that is MSER (Maximally Stable Extremal Region) algorithm. MSER is a method for text detection, blob detection in images. The MSER algorithm extracts number of co-variant regions from image. We first define the concept of stroke and then explain the stroke width transformation.

MSER is based on the idea of taking regions which stay nearly the same through a wide range of thresholds. All the pixels above or equal to a given threshold are black and all the pixels below a given threshold are white.

MSER uses two important properties to remove non text regions from image first is Geometric Properties and another is Stroke Width Variation Properties.

To use MSER algorithm we first summarized the common attributes of text as

a) Text in image always contains lots of edges;
b) The width of text is larger than height;
c) Text is bounded in size;
d) Text has special texture but this texture is irregular;

The flow chart of the algorithm is shown on Figure 3. The flowchart shows the actual flow of MSER algorithm. MSER algorithm detect all text regions alongside it also detect some non text regions. To remove that non text regions Geometric Properties are applied on that image. If Geometric Properties are unable to remove all non text regions then apply Stroke Width Variation on that image to remove remaining non text regions.

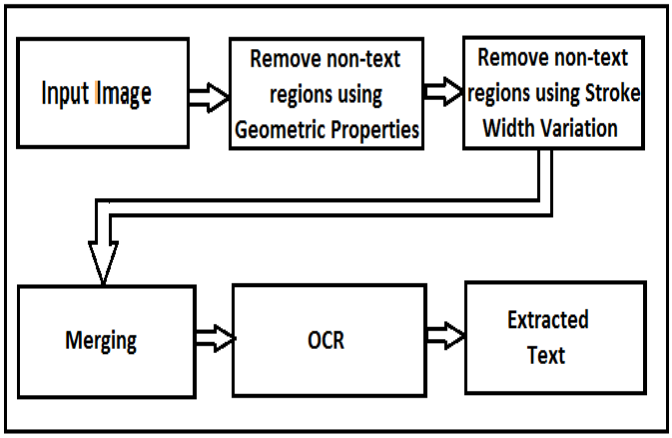

Fig -5: Flowchart of the algorithm

Geometric properties

MSER detects almost all text regions from image but alongside it also detect some non text regions. To remove those non text regions first we apply Geometric Properties on image. Geometric Properties detects the non text regions from image and remove those regions. The non text regions which are not removed in Geometric Properties for those regions we use the Stroke Width Variation Properties.

Stroke width variation properties

The Stroke Width Variation Properties are also used to remove the non text regions. Stroke Width is a measure of the curves and lines that make up character. Text regions have little stroke width variation, where as non text regions have larger variations. To remove the non text regions using stroke width we require thresholds. All the pixels above or equal to a given threshold are black and all the pixels below a given threshold are white.

Grouping letters into text lines

At this point, the individual detected letters are grouped into words or text lines. The grouping of letters carries more meaningful information than just the individual letter. For example, recognizing the string 'HELP' vs. the set of individual characters {'L','H','P','E'}, where the meaning of the word is lost without the correct ordering.

Result

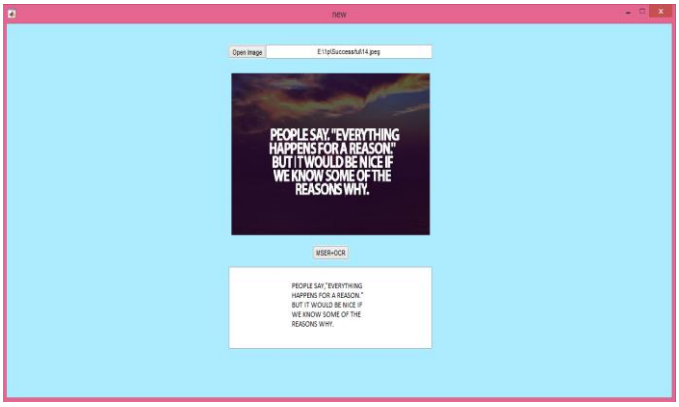

Fig -6: Result

5.CONCLUSION

In this paper, we have recognized a text from image using MSER algorithm. In this paper we have used five steps to recognize text. For recognizing the text we use two important properties first is Geometric Properties and second is Stroke Width Variation Properties. First MSER algorithm detects the text regions and then OCR identifies the characters. The MSER algorithm has good text recognition performance.

REFERENCES

[1] L. Gomez and D. Karatzas, "Multi-script text extraction from Natural scenes," in ICDAR, 2013.

[2] Y.-F. Pan, X. W. Hou, and C.-L. Liu, "Text localization in natural scene images based on conditional random field," in Proc. 10th Int. Conf.Document Analysis and Recognition (ICDAR'09), Barcelona Spain,2009,pp. 6–10.

[3] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust Wide Baseline Stereo from Maximally Stable Extremal Regions." in *In Proceedings of* the British Machine Vision Conference (BMVC), London, UK, sep 2002, pp. 384–393.

[4] .M. Lee, and A. Kankanhalli, Automatic Extraction of Characters in Complex Images, International Journal of Pattern Recognition Artificial Intelligence, 9 (1) (1995) 67-82.

[5] V. Wu, R. Manmatha, and E. Riseman, "Finding text in images," 20th Int. ACM Conf. Research and Development in Information Retrieval, pp. 3-12,1997.

[6] E. Ofek ,B. Epshtein, , and Y. Wexler, "Detecting text in natural scenes with stroke width transform," in Proc. IEEE Conf. Computer Vision. Pattern Recognition., Jun. (2010), pp. 2963–2970.

[7] G. Schroth, S. TsaiD. Chen, H. Chen ,R. Grzeszczuk, and B. Girod,"Robust text detection in natural images with edge-enhanced maximally stable extremal regions," in Proc. IEEE Int. Conference. Image Processing,(2011), pp. 2609–2612.

[8] O. Chum, J. Matas, T. Pajdla and U. Martin,, "Robust wide baseline stereo from maximally stable extremal regions," in Proc. Brit. Mach.Vision. Conference., (2002), pp. 384–393.

[9] A. Jain, B. Yu, "Automatic Text Location in Images and Video Frames", Pattern Recognition 31(12): 2055-2076 (1998)

[10] K. Subramanian, P. Natarajan, M. Decerbo, D. Castañòn, "Character-Stroke Detection for Text-Localization and Extraction", International Conference on Document Analysis and Recognition (ICDAR), 2005

[11] S. Akamatsu , A. Shio J. Ohya, , and, Recognizing Characters in Scene Images, IEEE Transactions on Pattern Analysis and Machine Intelligence, (1994) ,pp214-224.

[12] Y. Zhong, K. Karu, and A. K. Jain, "Locating text in complex color images," Pattern Recognition. Vol. 28, pp.1523-1535, October, 1995

[13] S. Park, J. Lee, J. Koo, O. Kwon, S. Hog, S, "Adaptivetracking algorithm based on direction field using ML estimation in angiogram", In IEEE Conference on Speech and Image. Technologies for Computing and Telecommunications. Vol. 2. 671-675 (1999).

[14] Y. Sun, "Automated identification of vessel contours in coronary arteriogramsby an adaptive tracking algorithm", IEEE Trans. on Med. Img. 8, 78-88 (1989).

[15] J. Canny, "A Computational Approach To Edge Detection", IEEE Trans. Pattern Analysis and Machine Intelligence, 8:679-714, 1986.