

# SURF BASED FAST AND EFFECTIVE PRUNING ALGORITHM FOR IMAGETEXT IDENTIFICATION

M.Manju<sup>1</sup>, K.Rajasundari<sup>2</sup>, S.U.Sowmiya<sup>3</sup>

<sup>1</sup>PG scholar Department of CSE, Francis Xavier Engineering College, Tamilnadu, India.

<sup>2,3</sup>AP, Department of CSE, Francis Xavier Engineering College, Tamilnadu, India.

**Abstract-**In today world text image retrieval place role due to it applications in content-based image analysis tasks but the variations of front, size, color, orientations. Text retrieval is difficult the existing methods do not provide effective text candidates construction algorithms. The original MSER built extract text detection method is proposed to identify text in ordinary scene image. This groups the text candidates using a single-link clustering algorithm. Texts are identified using the machine learning techniques. Evaluation results prove this method efficient and effective for text retrieval.

**Index Terms-** Maximally Stable Extremal Regions, Content-Based Image, Single-Link Clustering Algorithm, Speed up Robust feature.

## I. INTRODUCTION

Text in image contains valuable information. Additionally found a similar reduction when users choose no strokes of segmentation. Complex backgrounds and multiple texts appear. To overall system performance is observed. One character being segmented as multiple characters. Text character can be classified into various font sizes, color, and orientation. The regions based method text in the image and then use character extraction. To identify text extract character candidate from image candidate into text. Hybrid method based on character candidate by local linearization, non-character is eliminated and character can finally be grouped into text. In many image local linearization is stable over larger range of threshold in certain region. In variance transformation stability since only external region whose support is threshold is selected multi-scale detection. The conventional of all extremely regions can be enumerated in segment. Each and every character is segmented of an image.

## II. RELATED WORK

As described in the author proposed technique called image text detection. We are still using the text line alignment feature. The track of alignment is calculated

from the first pair of components and is repeatedly recalculated and used to normalize the increasing operation. The importance order ensures that the first twosome of components can be principal to compute the initial arrangement path. The advantage for the angle requirement is identified empirically. The Component distance and inception use our learned distance metrics in with the exception that top and bottom orientation feature values are set. The line between cancroids of the existing component and the optimal candidate should not interconnect with extra candidate. Scheduled the additional hand, if we have not found any component candidates satisfying the angle and distance requirements, the existing section is not inflatable and increasing on this trend is finished. The author proposed a technique called a method for text localization and recognition in real-world images.

## III. PROPOSED SYSTEM

Identifying repeating component each and every character is segmented. Text candidate are constructed by canny edge detection. Character candidate extraction in multiple thresholds of different region filter. Text mask, area, eccentricity. To identify many low quality Characters. Global features are extracted. System offer high speed efficiency. Text candidate to extract higher order of text and uses exhaustive search for pruning to overcome the above parameters. Read the image and then use the RGB color in gray scale images. The Character candidate extraction in MSER construction use canny edge detection. Picture of street view text image (or) any other text image.

## IV. SYSTEM DESIGN

The disconnected line connection based on the distance and strength of the pixels and groups of Text detection. The corresponding to a character segmented.

To identify Component distance and threshold use our learned distance metrics.

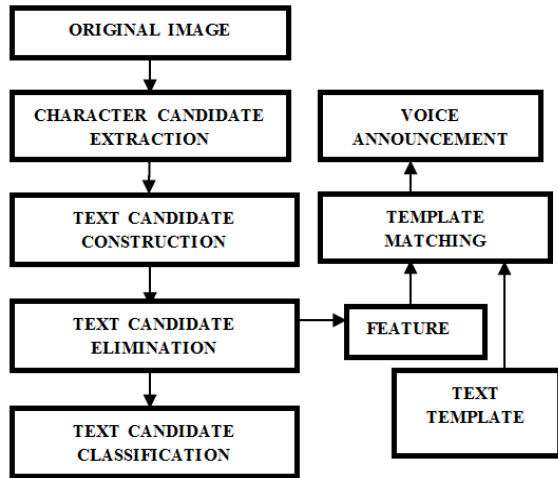


Fig.1 SYSTEM ARCHITECTURE

V. METHODOLOGY

The text image to be segmented into each and every individual character. To identify the text in edge detection strength is high. The sequential change of text image an extract in the character into text. The conditions to make easy it identification.

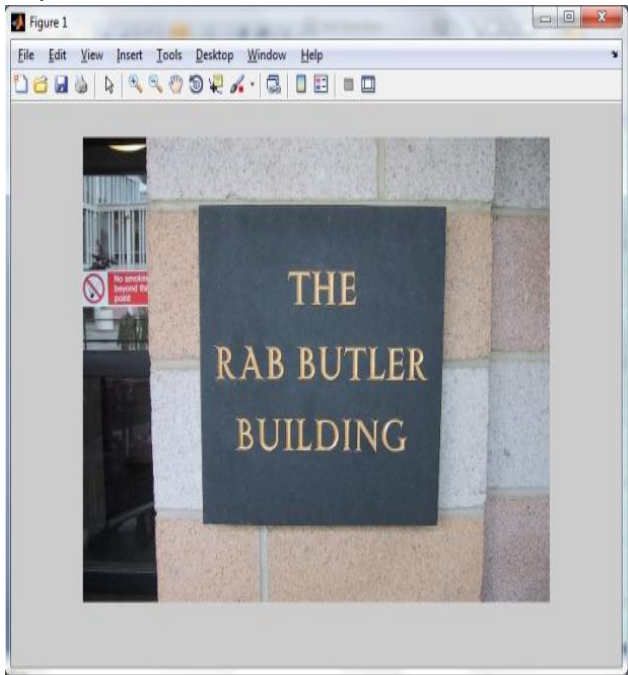


Fig.2 ORIGINAL IMAGE

5.1 CHARACTER CANDIDATE EXTRACTION

Character candidates are extracting using the MSERs algorithm.

(i) They are constructing during a process of trying many states.

(ii) The select regions are those that maintain unaffected shape over a huge set of region.

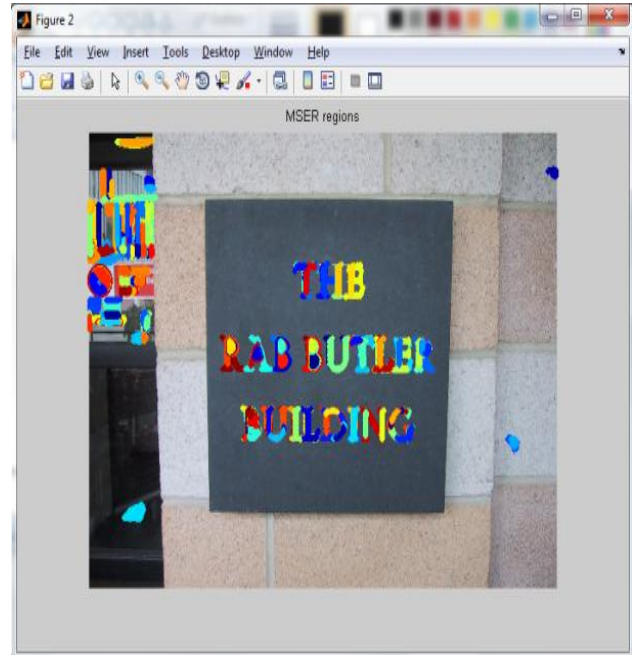
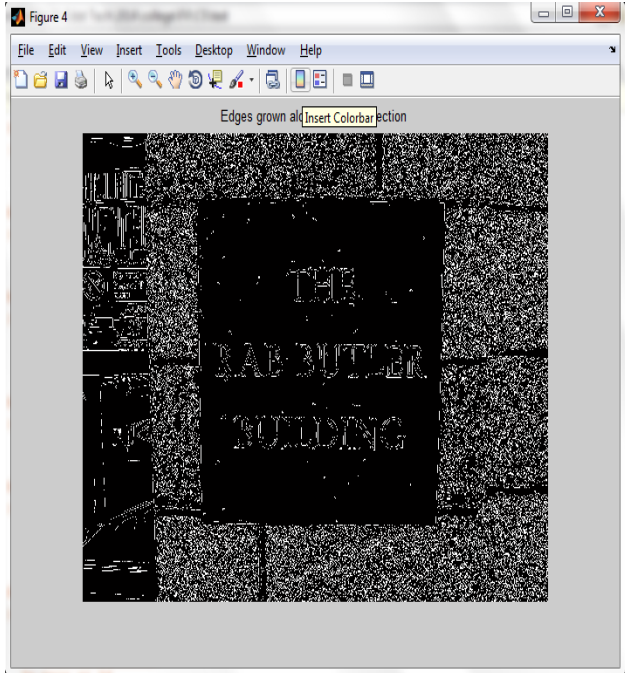


Fig.3 CHARACTER CANDIDATE EXTRACTION

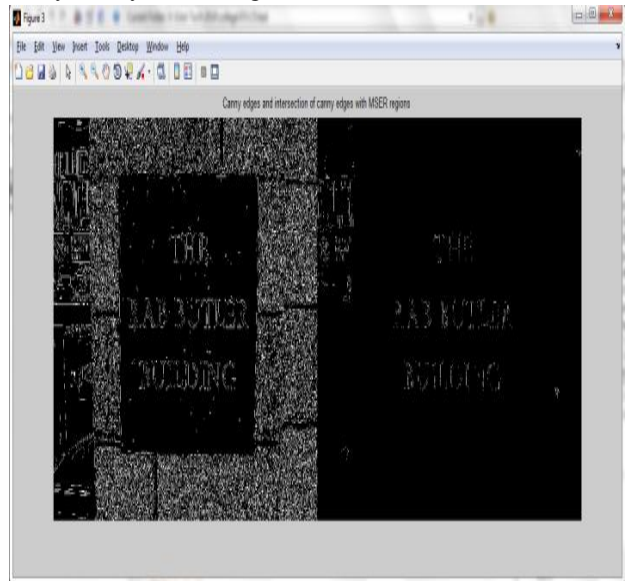
5.2 TEXT CANDIDATE CONSTRUCTION  
 Text can be constructing using by an edge detection technique. To read the image then identify text and clarify correct text in each and every individual character in the image. The text candidates edges are specify in the current direction. They should be identifying in the every single edge distance to be capture in the path and notify the text candidate from the image. The capital text candidate or small text candidate or broken text candidate are also identify in the text matching concept are include in the text using the template. They should not a particular language using in the different language. It contain edge detection do not have a frame edge. It's have an object edge. The edge detection can be clarify text are line of distance and strength of the pixels in another edge enhancement approach to be follow the concept of text construction.



**Fig.4 TEXT CANDIDATE CONSTRUCTION**

### 5.3 TEXT CANDIDATE ELIMINATION

The non-texts candidate only be identified using a canny binary to filtering the text for individual character.

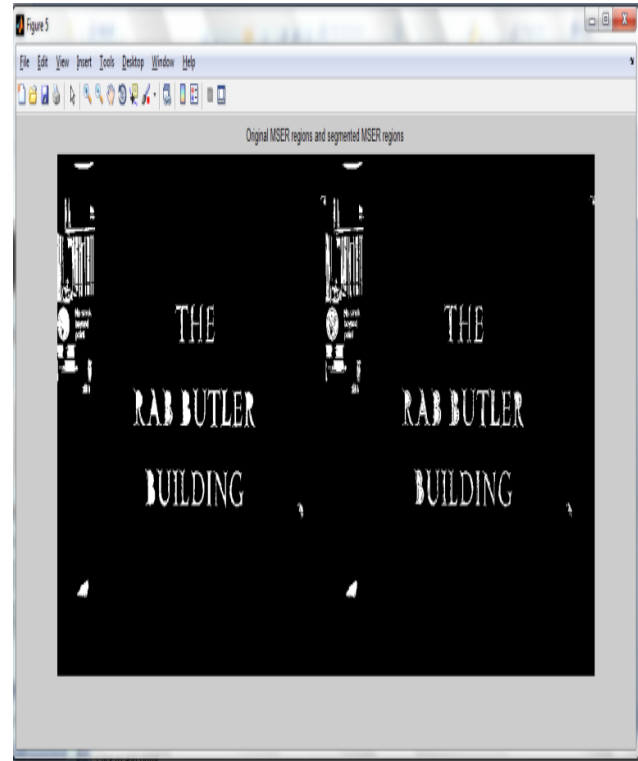


**Fig.5 TEXT CANDIDATE ELIMINATION**

### 5.4 TEXT CANDIDATE CLASSIFICATION

Template into text classification based on the template matching. The major concept is used by an ADABOOST classification technique. The text must be classify by the every character correct text to be selected in the template then to pair wise matching and true text are identified in

the classification. Finally correct text should be producing the text in screen and audio voice announcement.



**Fig.6 TEXT CANDIDATE CLASSIFICATION**

## VI. CONCLUSION

We propose a high-speed and accurate MSERs pruning algorithm that enable us to sense the majority characters even when the image is in low quality. We build a strong scene text detection method that exhibit higher performance over state-of-art method on a variety public database.

## VI. ACKNOWLEDGMENT

I thank my guide ,Ms K Rajasundari Assistant Professor, Department of CSE for her guidance and support. I also thank Ms. S.U.Sowmiya Assistant Professor for his support.

## REFERENCE

- [1].Zhong Y,Zhang,H and Jain.A(2000)'Automatic caption localization in compressed video ,'IEEE Trans. Pattern Anal. Mach. Intel.
- [2].Doermann .D, and Kia.O (2000) 'Automatic text detection and tracking in Digital video,'IEEE Trans. Image Process.

- [3].Weinman.J, and Learned-Miller's, and Hanson .A(2009)'Scene text recognition using Similarity Intel and a lexicon with sparse belief propagation,' IEEE Trans. Pattern Anal. Mach.
- [4].Yin.C-X, Hao.W-H, Sun.J, and Naoi.S(2011) 'Robust vanishing point detection for mobile cam-based documents,' in Proc. ICDAR, Beijing, .Shivakumara.P, Trung.P, and Tan.L .C, 'A laplacian approach to multi-oriented Text detection in video,' IEEE Trans. Pattern Anal. Mach. Intel.
- [5].Chen.J, and Yuille.A(2012) 'Detecting and reading text in natural scenes,' in Proc. IEEEConf. CVPR.
- [6].Lee.J-J, Lee.P-H, Lee.W-S, Yuille.A, and Koch.C (2011) 'Adaboostfor text Detection in natural scene,' inProc.
- [7].Kim.K, Jung.K, and Kim.J(2003) 'Texture-base approach for text detection in Images using support vector machines and continuously a412-419, Feb. 2011.daptive mean shift algorithm,' IEEE Trans. Pattern Anal .Mach. Intel.
- [8].Epshtein.B, Ofek.E, and Wexler.Y (2010) 'Detecting text in natural scenes with stroke width transform,' in Proc. IEEE Conf. CVPR, San Francisco,
- [9].Yi.CandTian.Y(2011)'Text string detection from natural scenes by structure- based partitue and grouping,' IEEE Trans. Image Process.
- [10].Mancas-Thillou.C, and Gosselin.B (2007) 'Color text extraction with selective metric-based clustering,' Comput. Vis. Image Und.