Elysium Journal
Engineering Research & Management

P-ISSN: 2347-4408
E-ISSN: 2347-4734

ROBUST SCENE TEXT RECOGNITION WITH SPARSE BELIEF PROPAGATION

S.Priyadharsini[1], Mrs. J.Ranjani[2]

[1]*PG Student, Department of Computer Applications, Sri Sairam Engineering College, Chennai, Tamilnadu, India.*
[2]*Assistant Professor, Department of Computer Applications, Sri Sairam Engineering College, Chennai, Tamilnadu, India.*

Abstract- Content-based image analysis task requires text detection in natural scene images. The character candidates in text detection on minimizing regularized variations termed as Maximally Stable Extremely Regions (MSERs), which are extracted by using a fast and effective pruning algorithm. The single-link clustering algorithm groups character candidates into text candidates. In this grouping, candidates automatically learn the distance weights and clustering threshold. The text classifier used in this grouping performs two operations. They are elimination of non-text probabilities and identification of texts. MSER (Maximally Stable Extremely Regions) efficiently identifies the text and eliminate the non-text area. Forward and Backward algorithm is used to read the different orientation of the text. The retrieved text is converted into audio format by using the classic text to speech engine interface. The audio format which is very useful in various situation in the real world. Then convert the retrieved text into an audio format which is useful in various situations in the real world. Proposed system going to implement audio converted text in the reserved area which is a very useful application for the blind people.

Key Words – Scene Text Detection, Maximally Stable External Regions (MSERs), Single-Link Clustering, Distance Metric Learning.

1. INTRODUCTION

The text in image is exploited, which contain valuable information in many content based image and video applications. The exploitation of text in content-based web image search video information retrieval, and mobile based text analysis and recognition is discussed [1-3]. The robust text detection before the image is being recognized and retrieved is affected by various factors such as complex background, variations of font, size, color and orientation.

MSER-based methods have reported promising performance on the widely used ICDAR 2011 Robust Reading Competition database. However, several problems remain to be addressed. First, the MSERs algorithm detects a large number of repeating components and these repeating components are problematic for the latter character grouping algorithm. The most likely part in the image that corresponds to characters is to be removed before image processing. The accuracy and speed in existing methods for MSERs are still required improvement. Second, current approaches for text candidate's construction are categorized into two methods. They are rule based and clustering-based methods. Rule-based methods generally require tuning parameters by manual, which is time-consuming and error prune. The clustering-based method provides good performance but is complicated by incorporating post-processing stage followed by minimum spanning tree clustering.

In this paper, a robust and accurate MSER based scene text detection method is proposed to enhance the speed and accuracy. Initially, a fast and accurate MSERs pruning algorithm is designed from the hierarchical structure of MSERs and by acquiring simple features. There are two observations are made by using proposed algorithm. They are the number of character candidates are reduced and high accuracy is achieved.

Second, a novel self-training distance metric learning algorithm is proposed, which learns the distance weights and clustering threshold automatically. The single-link algorithm clusters the character candidates into text candidates with learned parameters. Third, character classifier estimates the posterior probabilities of text candidates to remove text candidates with high non-text probabilities. The non-text probability elimination helps to train the powerful text classifier for identification of text.

Finally, an accurate and robust text detection builds on the basis of fast and accurate MSERs pruning algorithm, novel self-training distance metric learning algorithm, character classifier. The text detection system is evaluated on the benchmark ICDAR 2011 Robust Reading Competition database (Challenge 2) and has achieved an f - measure of 76%, which is much higher than the current best performance of 71%. The experiments are carried out on multilingual, multi-orientation, street view and even born-digital (web and email) databases. The proposed method demonstrates the significant improvement over the existing methods.

The color variation is used to identify candidates. The blue bounding rectangles denotes text candidates. The green color denotes the character candidates and the red color indicates the other candidates in the image.

P-ISSN: 2347-4408
E-ISSN: 2347-4734

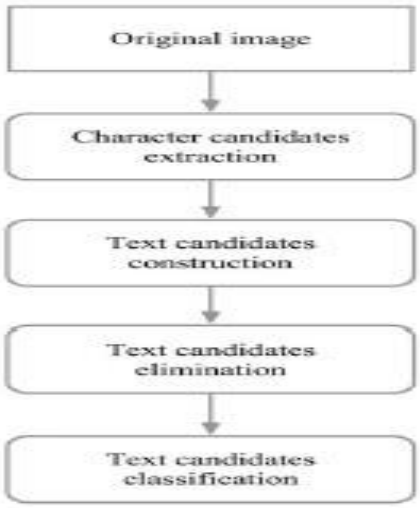

Fig.1 Flowchart of the proposed system and results after each step of the sample

2. EXISTING SYSTEM

The traditional methods of scene text detection is classified into three groups namely, sliding window based, connected component based and hybrid methods. The region based methods are termed as sliding window based methods [4], which are used to search the possible texts in the image. The multiple scaling process of image is slow in sliding window based method. Connected component based methods [5-9] Character candidates are grouped into text, additional checks needed to remove non-character candidates. Hybrid methods [10] Extracts connected component as character candidates.

2.1 DISADVANTAGES

- Slow process –Many iterations required to extract the text candidates.
- Less Accuracy in extracting the text.
- Time consuming due to iterations needed.
- High cost.

3. PROPOSED SYSTEM

The proposed system uses the MSER strategy to detect the text in natural sceneries [10]. Hierarchical structure of MSERs strategy is used to process the character candidates. The single-link clustering algorithm clusters the character candidates into text candidates. Character classifier is used to identify the text. Refer Fig.2 architecture diagram of proposed system.

3.1 ADVANTAGES

- Easy to extract text from image files
- MSER is useful to detect very small regions.
- Improve readability to blur images.
- MSER is better stereo matching and object recognition algorithm.
- MSER is adaptive to color Images, used to detect color regions.

Fig.2 Contextual Diagram of Proposed System

4. PROCESS INVOLVED IN CONVERTING TEXT TO SPEECH

The steps used to extract text from images and converted into speech format:

1. Capturing Images
2. Loading Images
3. Extracting text
4. Removing Noise
5. Identifying text
6. Separating text from images
7. Conversion of text to speech

P-ISSN: 2347-4408
E-ISSN: 2347-4734

4.1 CAPTURING IMAGES

The images are captured by using web camera or by using mobile. Also capturing the images from the video footage which is running sequentially, from the video footage used to capture particular area which having the text**.**

4.2 LOADING IMAGES

Captured Images were subjected to identification process the quality, visibility of the text were checked. From captured images select the area which having the text and then marking with a rectangular shape selection tool.

4.3 EXTRACTING TEXT

Extracting the text using fast and effective pruning algorithm and MSER algorithm. The single-link clustering algorithm groups the character candidates into text candidates. The distance weights and clustering threshold are learned automatically in single-link algorithm. .

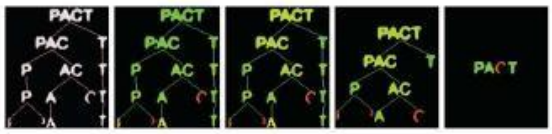

Fig.3 Hierarchical Structure of MSER Algorithm

At each step, the separated clusters are combined. The shortest links at any step causes the fusion of the two clusters. The elements of separated clusters are involved in the combination process. The method is also termed as nearest neighbor clustering.

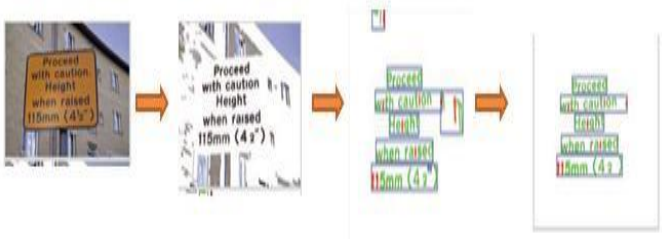

Fig.4 Process of extraction of text from images using rectangular tool

Let us consider the elements of clusters are X and Y. The distance between the sub set elements (x, y) and the linkage function D(x, y) between clusters X and Y-is related as follows:

D(X, Y) =min d(x, y)

If the text is arranged horizontally then single-link clustering algorithm is used to extracting the text. Multi orientation of text is not extracted by using this algorithm.

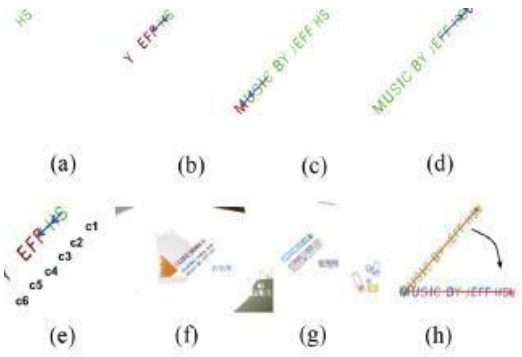

Fig.5 Forward and Backward algorithm is used to extract the Multi -orientation text.

4.4 REMOVING NOISE

With the basis of color combination and text based technique the noise and unwanted characters were removed to attain accuracy in text detection.

The noise can be removed by using the technique Leptonica and Tessa act to reduce noise and unwanted characters. Leptonica and Tessaract are the Library files.

4.5 IDENTIFYING TEXT

After removal of noise and unwanted content it makes easier to identifies the text exactly and then subjecting into analyzing of reading text, with the help of the techniques Leptonica and Tessract selected content were compared and text were identified as readable and editable.

4.6 SEPARATING TEXT FROM IMAGES

After identifying the text, text were shown in the separate area to differentiate the source text and resulting text.

P-ISSN: 2347-4408
E-ISSN: 2347-4734

4.7 CONVERSION OF TEXT TO SPEECH

Separated text is converted into speech by using the interface called Classic Text to Speech Engine. It supports multiple language and full-fledged text to speech engine for Android.

5. IMPLEMENTATION

Fig.6.1: Listing of Menus

The above figure has shown the listed of menus.

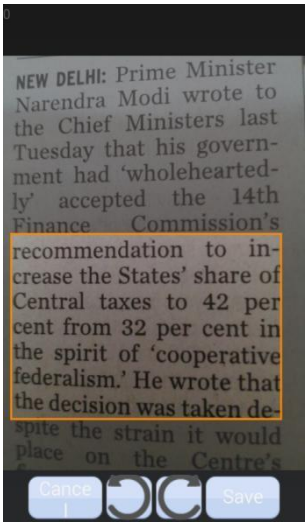

Fig.6.2 Cropping Text

Figure 6.2: Cropping the needed text area in an image.

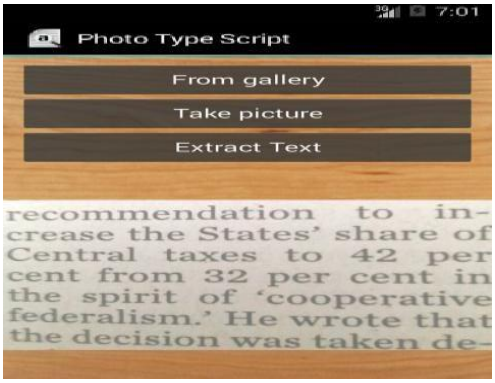

Fig.6.3: Extracting text from Images

Fig.6.3 has shown extracting text candidates from non-text candidates.

Fig.6.4: Conversion of Text To Speech

The above figure has shown initialization of Text-To- Speech Engine

Fig.6.5: Translation of text

P-ISSN: 2347-4408
E-ISSN: 2347-4734

Fig.6.5 Classic Text to Speech Engine provides Text translation in different languages.

6. RESULTS

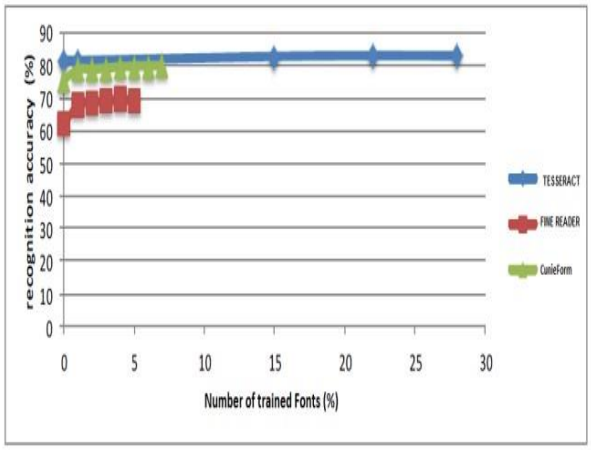

Fig .6: Performance Analysis Diagram

The above diagram has shown the performance analysis. It produces more accurate results and uses Tesseract, which is an accurate open source OCR engine available. In this application the words are trained using Tesseract library. This is an updated version of library whereas ABBYY FineReader and CunieForm are outdated versions.

7. CONCLUSION

In this paper presents a new MSER-based scene text detection method with several novel techniques. Proposed system is a fast and accurate MSERs pruning algorithm that enables us to detect most characters in low quality image. A novel self-training distance metric learning algorithm is learned distance weights and clustering threshold simultaneously; text candidates are constructed by clustering character candidates by the single-link algorithm using the learned parameters.

Text candidate corresponding to non-text candidate and eliminate text candidates with high non-text probability, which helps to build a more powerful text classifier. Finally, by integrating the above new techniques, to build a robust scene text detection system that exhibits superior performance over state-of-the-art methods on a variety of public databases.

Our text detection system provides the best result for both “Text Localization in Real Scenes” and “Text Localization in Born-Digital Images” in the ICDAR 2013 Robust Reading Competition.

8. REFERENCES

[1] H. Li, D. Doermann, and O. Kia, "Automatic text detection and tracking in digital video," *Image Processing, IEEE Transactions on*, vol. 9, pp. 147-156, 2000.

[2] J. J. Weinman, E. Learned-Miller, and A. R. Hanson, "Scene text recognition using similarity and a lexicon with sparse belief propagation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 31, pp. 1733-1746, 2009.

[3] X.-C. Yin, H.-W. Hao, J. Sun, and S. Naoi, "Robust vanishing point detection for MobileCam-based documents," in *Document Analysis and Recognition (ICDAR), 2011 International Conference on*, 2011, pp. 136-140.

[4] X. Chen and A. L. Yuille, "Detecting and reading text in natural scenes," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, 2004, pp. II-366-II-373 Vol. 2.

[5] B. Epshtein, E. Ofek, and Y. Wexler, "Detecting text in natural scenes with stroke width transform," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, 2010, pp. 2963-2970.

[6] C. Mancas-Thillou and B. Gosselin, "Color text extraction with selective metric-based clustering," *Computer Vision and Image Understanding*, vol. 107, pp. 97-107, 2007.

[7] C. Yi and Y. Tian, "Text string detection from natural scenes by structure-based partition and grouping," *Image Processing, IEEE Transactions on*, vol. 20, pp. 2594-2605, 2011.

[8] C. Yi and Y. Tian, "Localizing text in scene images by boundary clustering, stroke segmentation, and string fragment classification," *Image Processing, IEEE Transactions on*, vol. 21, pp. 4256-4268, 2012.

[9] C. Yi and Y. Tian, "Text extraction from scene images by character appearance and structure modeling," *Computer Vision and Image Understanding*, vol. 117, pp. 182-194, 2013.

[10] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," *Image and vision computing*, vol. 22, pp. 761-767, 2004.